RESEARCH IN ENGLISH LANGUAGE EDUCATION JOURNAL
E-ISSN 3115-7599

Comparing the Effectiveness of AI Chatbots and Traditional Teaching in Academic Writing Instruction

Mohammad Aliakbari [1], Mohammad Mahdi Maadikhah [2] & Pooria Barzan [3]

[1] Professor, Department of English Language and Literature, Ilam University, Ilam, Iran.

m.aliakbari@ilam.ac.ir

[2] *Corresponding Author*: Ph.D. Candidate, Department of English Language and Literature, Ilam University, Ilam, Iran.

m.maadikhah@ilam.ac.ir

[3] Ph.D. Candidate, Department of English Language and Literature, Ilam University, Ilam, Iran.

p.barzan@ilam.ac.ir

ABSTRACT

AI-enabled solutions and systems, including AI chatbots, are revolutionizing teaching and learning processes in language education programs and language classrooms, profoundly changing traditional methodology and pedagogy. This study aimed to compare the effectiveness of implementing and utilizing three AI chatbots, Claude, ChatGPT, and Google Gemini, with traditional instruction in an academic writing course at Ilam University, Iran. To this end, of 39 TEFL MA students enrolling in an academic writing course, four homogeneous classes of nine students were organized. In three classes, the three AI chatbots were used for providing feedback, correction, revision guidance, and tips for outlining, moves, and paragraph organization. A class with traditional teacher-centered instruction was also organized. At the end of the course, students in all four classes were tasked with writing two essays. A team of three raters was asked to rate the essays based on a common rubric. The rating results showed that students who received AI-enhanced instruction through the use of Claude achieved the highest scores, followed by those who received instruction with ChatGPT and Google Gemini, respectively. The participants in all three AI-enabled classes outperformed the participants in the class with traditional instruction. The rating results confirmed the effectiveness of AI-enhanced writing instruction, highlighting Claude as the most effective AI tool, and ChatGPT and Google Gemini in the second and third ranks, respectively. Based on the course outcomes, AI-enabled writing instruction is more effective than traditional instruction.

ARTICLE INFO:
Received: 2025-03-12
Revised: 2025-07-24
Accepted: 2025-08-05
Published online:2025-12-26

Keywords:
AI-assisted instruction, academic writing, AI chatbots, traditional teaching

Article type: Research Article | **Publisher**: Farhangian University https://elt.cfu.ac.ir

Citation: Aliakbari, M., Maadikhah, M. M., & Barzan, P. (2025). Comparing the effectiveness of AI chatbots and traditional teaching in academic writing instruction. *Research in English Language Education Journal*, *4*(2), 1-16. DOI: 10.48210/relej.2025.18722.1126

1. Introduction

Artificial intelligence (AI) has emerged as a transformative force in education, reshaping traditional pedagogical approaches and enhancing learning experiences across various contexts. AI-enabled tools, particularly AI chatbots, are increasingly being integrated into educational frameworks, offering personalized support, instant feedback, and adaptive learning experiences. In academic writing instruction, these technologies hold significant potential to complement traditional methodologies by addressing challenges such as delayed feedback, the difficulty of catering to diverse learner needs, and the demand for individualized writing guidance.

Despite the enthusiasm surrounding AI's integration into education, empirical evidence on its effectiveness compared to established teaching methods remains limited. While research has explored AI's role in language learning, few studies have systematically evaluated its impact on academic writing instruction, particularly in English as a Foreign Language (EFL) programs. Developing advanced writing skills is a critical yet challenging objective for EFL learners, necessitating effective instructional approaches that foster coherence, structure, and stylistic precision.

1.1 Advancements and Challenges of AI in Education Beyond Academic Writing

AI has revolutionized broader educational landscapes, driving advancements in personalized learning, administrative efficiency, and intelligent tutoring systems. One of the most notable advancements in AI for education is the development of customized learning systems (Chouhan, 2023; Tapalova & Zhiyenbayeva, 2022). These systems leverage AI algorithms to tailor educational content to the individual needs of students, allowing for a more customized learning experience. For instance, AI can analyze a student's learning patterns, strengths, and weaknesses, subsequently adjusting the curriculum to optimize their learning outcomes. This adaptive learning approach not only enhances student engagement but also fosters a deeper understanding of the material, as learners can progress at their own pace and receive immediate feedback on their performance (ÇAYIR, 2023; Zheng & Badarch, 2022). The potential for AI to create personalized learning pathways is particularly significant in diverse classrooms, where students may have varying levels of prior knowledge and learning styles (Nkechi et al., 2024; Akinwalere & Ivanov, 2022).

Moreover, AI technologies facilitate the automation of administrative tasks within educational institutions, thereby allowing educators to focus more on teaching and student interaction. AI can streamline processes such as grading, attendance tracking, and scheduling, which traditionally consume a substantial amount of educators' time (Munir et al., 2022; Siminto, 2023). By automating these tasks, AI not only increases operational efficiency but also reduces the administrative burden on teachers, enabling them to dedicate more time to instructional activities and student support (Allahverdiyeva, 2024; Yang, 2024). This shift can lead to improved educational outcomes, as teachers can engage more meaningfully with their students and provide the necessary guidance to foster academic success (Sanasintani, 2023; Wahjusaputri et al., 2023).

Intelligent tutoring systems (ITS) further demonstrate AI's potential by offering real-time assistance and feedback, simulating one-on-one tutoring experiences (Cao et al., 2020; Qawaqneh, 2023). These systems dynamically respond to students' queries, provide hints, and suggest resources tailored to individual comprehension levels (Puspitaningsih et al., 2022; Tapalova & Zhiyenbayeva, 2022). By fostering active

engagement and critical thinking, ITS contributes to more interactive and effective learning environments (Karampelas, 2021; Sari & Purwanta, 2021). However, despite these advancements, AI's integration into education presents several challenges, including ethical concerns related to data privacy, security, and algorithmic bias (Akinwalere & Ivanov, 2022; Nkechi et al., 2024; Zheng & Badarch, 2022).

1.2 AI Chatbots vs. Traditional Teaching in Academic Writing

The emergence of AI chatbots as educational tools has sparked debates regarding their pedagogical implications and effectiveness in fostering student learning outcomes. AI chatbots leverage natural language processing and machine learning to engage students in interactive dialogues, offering immediate and detailed feedback (Pack, 2024; Parsakia, 2023). Studies suggest that these tools can enhance students' self-efficacy, problem-solving abilities, and critical thinking skills (Parsakia, 2023; Yahaya, 2024). However, their effectiveness depends on various factors, including design, implementation, and the specific educational objectives they aim to achieve (Tian, 2024).

Conversely, traditional teaching methods have long been the foundation of academic instruction, emphasizing direct interaction between educators and students. Teacher-centered instruction provides structured learning environments, facilitates discussions, and ensures that feedback is contextualized and personalized (Ng, 2024; Wu & Yu, 2023). However, traditional approaches also present challenges, such as scalability limitations and inconsistencies in feedback quality due to instructor variability (Kang, 2023; Williams, 2024). As educational institutions increasingly adopt AI technologies, evaluating the comparative strengths and limitations of AI chatbots versus traditional instruction becomes essential. While AI chatbots can provide valuable resources and support, there is a risk that over-reliance on these tools may diminish authentic learning experiences and student engagement (Duran, 2024; Williams, 2024). Therefore, it is crucial to strike a balance between leveraging AI technologies and maintaining the essential human elements of teaching that foster critical thinking and creativity (Duran, 2024; Ilieva, 2023).

1.3 Ethical Considerations and Future Directions

The ethical implications of employing AI chatbots in education cannot be overlooked. Concerns regarding data privacy, algorithmic bias, and the potential erosion of critical thinking skills have been raised by scholars (Williams, 2024). The collection and analysis of student data raise concerns about how this information is used and who has access to it. Furthermore, there is a risk that AI systems may inadvertently perpetuate existing biases if they are trained on skewed datasets, potentially leading to inequitable educational outcomes. Addressing these ethical dilemmas requires a concerted effort from educators, policymakers, and technologists to ensure that AI is deployed responsibly and equitably within educational contexts (Siminto, 2023; Yang, 2024).

Another significant challenge is the need for educators to be adequately trained in AI technologies (Begum, 2024; Wahjusaputri et al., 2023). As AI continues to evolve, teachers need to possess a foundational understanding of how these tools operate and how they can be effectively integrated into their teaching practices. Professional development programs that focus on AI literacy and pedagogical strategies for utilizing AI in the classroom are crucial for empowering educators to leverage these technologies effectively (Chouhan, 2023; Munir et al., 2022). Without proper training, there is a risk

that educators may struggle to implement AI tools effectively, limiting their potential benefits (Akinwalere & Ivanov, 2022; ÇAYIR, 2023).

Furthermore, the implementation of AI in education necessitates substantial investment in infrastructure and resources. Educational institutions must ensure that they have the necessary technological capabilities to support AI applications, including robust internet connectivity, hardware, and software (Nkechi et al., 2024; Siminto, 2023). This requirement may pose challenges, particularly for underfunded schools or those in rural areas, where access to technology can be limited (Akinwalere & Ivanov, 2022). Bridging this digital divide is essential for ensuring that all students have equal opportunities to benefit from AI-enhanced educational experiences (Begum, 2024; Yang, 2024).

In conclusion, AI presents transformative opportunities for academic writing instruction and broader educational practices, offering personalized learning pathways, administrative efficiencies, and interactive tutoring experiences. However, ethical considerations, educator preparedness, and infrastructural challenges must be addressed to realize AI's full potential in education. This study seeks to bridge this research gap by investigating the effectiveness of three AI chatbots—Claude, ChatGPT, and Google Gemini—in an academic writing course for TEFL MA students at Ilam University, Iran. By conducting a comparative analysis of AI-enabled instruction and traditional teacher-centered methods, the study examines how these chatbots assist with key writing aspects, such as outlining, paragraph organization, revision, and overall coherence. The findings aim to offer practical implications for educators exploring AI's integration into language education. By comparing AI chatbots and traditional teaching methods, this study contributes to the ongoing discourse on AI's role in academic writing instruction, providing insights that inform educators, researchers, and policymakers navigating the complexities of AI integration in education.

2. Review of Literature

The advancements in artificial intelligence (AI), particularly in AI chatbots and large language models (LLMs), have significantly transformed academic writing and writing instruction. The emergence of tools like ChatGPT and Grammarly has revolutionized the way students, educators, and researchers engage with writing, offering both opportunities and challenges. As AI technologies become more integrated into educational settings, their impact on writing proficiency, pedagogical approaches, and academic integrity has become a focal point of research and debate.

AI-powered writing tools have demonstrated remarkable capabilities in assisting with various aspects of academic writing. Studies indicate that these technologies enhance writing efficiency by providing real-time suggestions for grammar, style, and structure, thereby facilitating the writing process for both novice and experienced writers (Gervacio, 2023; Li, 2024; William, 2024). The ability of AI to generate coherent and contextually relevant text has made it a valuable resource for students struggling with challenges such as writer's block and lack of confidence (Mondal, 2023; Royani, 2024; Song, 2023). Additionally, research suggests that AI tools improve students' abilities to construct well-organized and coherent texts, ultimately enhancing overall writing quality (Tran, 2024).

In educational settings, AI chatbots have been integrated into writing instruction to foster more personalized learning experiences. These tools adapt to individual learning styles and provide tailored feedback, helping students from diverse backgrounds

develop their writing skills more effectively (Maphoto, 2024; Rabbianty, 2023; Selim, 2024). This adaptability is particularly beneficial in higher education and English as a Foreign Language (EFL) classrooms, where students often face varying levels of proficiency in academic writing (Aljuaid, 2024; ROA, 2024). AI-driven writing assistants allow for immediate feedback and real-time corrections, which can enhance language acquisition and writing proficiency among students (Minh, 2024). Furthermore, AI technologies have been shown to foster collaborative writing exercises, enabling students to work alongside AI to improve their writing skills while maintaining a sense of ownership over their work (Hz, 2023; Krajka, 2024).

Beyond technical improvements, AI tools have been found to positively influence students' motivation and self-efficacy in writing tasks. Studies indicate that students who use AI-assisted writing tools report higher levels of satisfaction with their writing experiences and a greater willingness to engage in writing tasks (Kim, 2024; Mondal, 2023). This increased motivation can lead to improved academic performance and a more positive attitude toward writing (Amirjalili, 2024; Kurniati & Fithriani, 2022). Additionally, AI has been recognized as an effective resource for pre-service English teachers, helping them enhance their teaching methodologies and support students in developing writing proficiency (Nadhifah, 2024).

However, despite the numerous benefits associated with AI chatbots in writing instruction, concerns regarding academic integrity and over-reliance on AI tools have emerged. The potential for students to misuse AI for content generation raises ethical questions about authorship and originality (Ahn, 2024; Guleria, 2023; PAN, 2024). Research highlights the need for educational institutions to establish clear guidelines and policies to regulate the use of AI in academic writing (Tang, 2023; Yeo, 2023). The importance of educating students about ethical writing practices and responsible AI use cannot be overstated, as maintaining academic integrity is crucial in an era of rapidly advancing AI technologies (Miao, 2023; Nazim, 2024; Yazid, 2024).

The role of AI in writing education extends beyond assistance; it also catalyzes pedagogical innovation. Educators are increasingly exploring hybrid teaching methodologies that integrate AI assistance while preserving critical thinking and analytical skills (Hutson, 2024). For example, research emphasizes the necessity of balancing AI tools with traditional writing instruction to ensure that students develop independent writing competencies (Aithal, 2023; Çerçi, 2023). Furthermore, studies suggest that AI can be leveraged to enhance writing instruction through automated writing evaluation systems, which provide valuable insights into students' writing processes and areas for improvement (Rahman et al., 2022; Wale, 2024).

As AI technologies continue to evolve, discussions surrounding the future of academic writing and the role of human authorship have gained prominence. Some scholars argue that while AI can assist in generating content, it cannot replace the unique insights, creativity, and critical perspectives that human writers bring to their work (Marmoah, 2024; Rababah, 2024). Therefore, AI should be viewed as a complementary tool that enhances, rather than replaces, human intellectual rigor (Moussa & Belhiah, 2024; Shopovski, 2024). Additionally, the integration of AI into research practices has sparked debates about transparency, accountability, and ethical considerations in scholarly communication (Hryciw, 2023; Zhao, 2024). Researchers are urged to establish responsible guidelines for using AI in academic writing to ensure that its benefits do not compromise the integrity of scholarly work (Al-Bukhrani, 2025; Kong, 2024).

In summary, the integration of AI chatbots and large language models, such as ChatGPT and Grammarly, has revolutionized academic writing instruction by enhancing efficiency, providing real-time feedback on grammar, style, and structure, and supporting both novice and experienced writers in overcoming challenges like writer's block and low confidence (Li, 2024; William, 2024). These tools foster personalized learning in educational settings by adapting to individual needs, improving writing proficiency—especially for diverse and EFL students—and boosting motivation and self-efficacy, which leads to better academic performance (Kim, 2024; Maphoto, 2024). However, concerns about academic integrity and over-reliance highlight the need for clear ethical guidelines (PAN, 2024). Pedagogically, AI drives innovation through hybrid teaching approaches that balance assistance with traditional methods to develop independent skills (Hutson, 2024). While AI enhances writing processes, it remains a complementary tool, unable to replace the creativity and critical thinking of human authorship (Marmoah, 2024).

3. Methods

3-1. Research Design

This study employed a non-equivalent control group quasi-experimental design to compare the effectiveness of three AI chatbots (Claude, Chat GPT, and Google Gemini) with traditional teacher-centered instruction in an academic writing course. The research aimed to assess the impact of these instructional interventions on the writing performance of TEFL MA students at Ilam University, Iran.

3-2. Participants

The participants were 39 MA students majoring in Teaching English as a Foreign Language (TEFL) at Ilam University. These students enrolled in an academic writing course during the study period. The participants were from different classes; however, they shared identical academic disciplines, curricula, instructors, and educational resources. To ensure homogeneity, participants were initially tested on their academic writing skills using a diagnostic essay task. Based on the results, students with similar proficiency levels were selected and randomly divided into four groups, each consisting of nine students. It is worth mentioning that three students were excluded from the groupings due to a lack of homogeneity and their outlier scores in the pre-test.

The study included four instructional groups:

1. Claude Group: Students in this group received instruction and feedback through the AI chatbot Claude.
2. Chat GPT Group: This group utilized Chat GPT for feedback, correction, and guidance.
3. Google Gemini Group: Google Gemini was used for providing instructional support to this group.
4. Traditional Instruction Group: Students in this group received traditional teacher-centered instruction without the aid of AI tools.

Each group participated in the same academic writing course over a semester, covering key topics such as outlining, paragraph organization, and writing moves.

3-3. Data Collection Procedure

The course was conducted over a 14-week semester, with two weekly sessions. The instructional approach for each group was as follows:

1. AI-Enabled Groups (Claude, Chat GPT, Google Gemini):

- The AI chatbots were used to provide feedback on drafts, suggest revisions, and offer guidance on organization and content development.

- Students interacted with the chatbots independently and in class under the instructor's supervision.

- The instructor provided supplementary explanations and addressed technical or content-related queries as needed.

2. Traditional Instruction Group:

- The instructor provided direct instruction, guided practice, and feedback on students' writing tasks.

- Feedback was delivered through written comments on drafts and oral explanations in class.

3-4. Data Collection Instruments

At the end of the semester, all students were tasked with writing two essays on academic topics chosen from the course syllabus. The essays were designed to assess students' ability to outline, organize paragraphs, and employ appropriate writing moves. The essays were evaluated by a team of three independent raters. The raters were experienced writing instructors and were trained to use a common rubric. The rubric assessed the following criteria:

1. Organization and coherence
2. Content development
3. Use of academic language and style
4. Grammar and mechanics

The scores from the three raters were averaged for each student to ensure consistency and reliability.

3-5. Data Analysis

The essay scores were analyzed using descriptive and inferential statistics. Descriptive statistics provided mean scores and standard deviations for each group, while a one-way ANOVA was conducted to determine statistically significant differences in writing performance among the four groups. Post hoc tests (e.g., Tukey's HSD) were applied to identify pairwise differences.

3-6. Ethical Considerations

The study followed ethical guidelines governing research with human participants, prioritizing participant welfare and data integrity. Participation was entirely voluntary, and informed consent was secured from all individuals before their involvement. During the consent process, participants were thoroughly briefed on the study's purpose, procedures, potential benefits, and their unconditional right to withdraw at any point without repercussions. To safeguard privacy, all data were anonymized by replacing personal identifiers with unique codes, and confidentiality was maintained throughout the research process, from data collection to reporting. Findings were presented in aggregate form to prevent the identification of individual respondents.

4. Results

The findings of this study indicate significant differences in academic writing performance among students who received AI-assisted instruction compared to those who received traditional teacher-centered instruction. The key results are presented in the following sections.

4.1 Descriptive statistics

The essay scores were evaluated based on four key criteria: content development, organization and coherence, use of academic language, and grammar and mechanics. The mean scores for each group are summarized in Table 1.

Table 1

Average Essay Scores Across Groups

Group	Essay 1 (Mean ± SD)	Essay 2 (Mean ± SD)	Combined Mean ± SD
Claude (AI-1)	88.5 ± 4.2	90.2 ± 3.8	89.35 ± 4.0
ChatGPT (AI-2)	84.3 ± 5.1	86.7 ± 4.5	85.5 ± 4.8
Google Gemini (AI-3)	81.7 ± 6.3	83.4 ± 5.8	82.55 ± 6.0
Traditional (TI)	76.8 ± 6.8	78.5 ± 7.1	77.65 ± 7.0

In a comparative evaluation of feedback effectiveness for academic writing, Claude (AI-1) demonstrated superior performance, achieving the highest average scores across both essays (89.35 ± 4.0), which underscores its ability to deliver the most impactful feedback, guidance, and support for enhancing writing skills. ChatGPT (AI-2) followed closely, with an average score of 85.5 ± 4.8, suggesting strong effectiveness in academic writing instruction, though slightly less refined or adaptable compared to Claude. Google Gemini (AI-3) ranked third, scoring 82.55 ± 6.0, reflecting moderate effectiveness that may stem from differences in feedback quality or pedagogical adaptability. In contrast, Traditional Instruction (TI) yielded the lowest average score (77.65 ± 7.0), highlighting its continued utility but relative limitations when compared to AI-enabled methods. These results collectively emphasize the growing advantage of AI-driven tools over conventional approaches in fostering academic writing proficiency.

The results demonstrate that students who received AI-enhanced instruction performed significantly better than those in the traditional instruction group. Among the AI-enabled groups, Claude (AI-1) achieved the highest mean scores, followed by ChatGPT (AI-2) and Google Gemini (AI-3), respectively.

4.2 Inferential statistics

To evaluate the impact of different instructional methods on academic writing performance, a one-way ANOVA was conducted, revealing a statistically significant effect of instructional method on essay scores, $F(3, N) = 14.57$, $p < 0.001$. This indicates that at least one group differed significantly from the others in terms of writing proficiency. Post-hoc Tukey tests further clarified these differences: Claude (AI-1) outperformed ChatGPT (AI-2) with a significant difference ($p = 0.02$), while ChatGPT (AI-2) surpassed Google Gemini (AI-3) ($p = 0.03$). All AI-enhanced groups significantly outperformed the Traditional Instruction (TI) group, with Claude (AI-1)

showing the strongest advantage ($p < 0.001$), followed by ChatGPT (AI-2) ($p < 0.001$) and Google Gemini (AI-3) ($p = 0.001$). Effect sizes, measured using Cohen's d, provided additional insight into the magnitude of these differences: Claude (AI-1) vs. Traditional (TI) exhibited a large effect ($d = 1.82$), ChatGPT (AI-2) vs. Traditional (TI) also showed a large effect ($d = 1.13$), and Google Gemini (AI-3) vs. Traditional (TI) demonstrated a moderate effect ($d = 0.78$). These findings underscore the superior effectiveness of AI-enhanced instruction, particularly Claude (AI-1), over traditional methods in enhancing academic writing skills, with substantial practical implications for educational practice.

Table 2
Inferential Statistics Results

Statistical Test	Comparison	Result	Significance
One-Way ANOVA	Overall group differences	$F = 14.57$, $p < 0.001$	Significant
Post-Hoc Tukey Test	Claude (AI-1) vs. ChatGPT (AI-2)	$p = 0.02$	Significant
Post-Hoc Tukey Test	ChatGPT (AI-2) vs. Google Gemini (AI-3)	$p = 0.03$	Significant
Post-Hoc Tukey Test	Google Gemini (AI-3) vs. Traditional (TI)	$p = 0.001$	Significant
Post-Hoc Tukey Test	Claude (AI-1) vs. Traditional (TI)	$p < 0.001$	Highly Significant
Post-Hoc Tukey Test	ChatGPT (AI-2) vs. Traditional (TI)	$p < 0.001$	Highly Significant
Effect Size (Cohen's *d*)	Claude (AI-1) vs. Traditional (TI)	$d = 1.82$	Large Effect
Effect Size (Cohen's *d*)	ChatGPT (AI-2) vs. Traditional (TI)	$d = 1.13$	Large Effect
Effect Size (Cohen's *d*)	Google Gemini (AI-3) vs. Traditional (TI)	$d = 0.78$	Moderate Effect

These statistical findings reinforce the effectiveness of AI-enabled instruction, particularly Claude, in enhancing academic writing performance.

5. Discussion

The findings of this study provide compelling evidence for the superior effectiveness of AI-enhanced instruction, particularly through the use of Claude, ChatGPT, and Google Gemini, over traditional teacher-centered instruction in an academic writing course for TEFL MA students at Ilam University, Iran. The results align with existing literature highlighting AI's transformative potential in education, particularly in personalized learning and real-time feedback (Chouhan, 2023; Tapalova & Zhiyenbayeva, 2022). However, the varying performance among the AI chatbots and their comparative advantages over traditional methods warrant further discussion to elucidate their implications for academic writing instruction and broader pedagogical practices.

The superior performance of the Claude group, with a combined mean essay score of 89.35 ± 4.0, underscores the chatbot's efficacy in providing high-quality feedback, revision guidance, and support for academic writing tasks. This finding is consistent with prior research indicating that AI tools enhance writing efficiency and coherence through tailored feedback (Li, 2024; Minh, 2024). Claude's ability to outperform ChatGPT (85.5 ± 4.8) and Google Gemini (82.55 ± 6.0) may be attributed to its advanced natural language processing capabilities, which likely provided more precise and contextually relevant suggestions for organization, coherence, and academic style. The statistically significant difference between Claude and ChatGPT ($p = 0.02$) suggests that Claude's design may be particularly well-suited for academic writing pedagogy, potentially due to its ability to simulate nuanced, instructor-like feedback.

ChatGPT and Google Gemini also demonstrated significant improvements over traditional instruction, with large and moderate effect sizes, respectively ($d = 1.13$ and $d = 0.78$). These findings align with studies that highlight AI chatbots' ability to foster self-efficacy and improve writing outcomes through immediate, personalized feedback (Mondal, 2023; Parsakia, 2023). However, the lower performance of Google Gemini compared to Claude and ChatGPT may reflect variations in feedback quality or adaptability to the specific demands of academic writing, such as structuring arguments or adhering to academic conventions. This variability among AI tools underscores the importance of evaluating their design and implementation to maximize their educational impact (Tian, 2024).

The traditional instruction group's lower performance (77.65 ± 7.0) highlights the limitations of teacher-centered methods in addressing diverse learner needs within the constraints of a semester-long course. Traditional instruction, while valuable for its structured guidance and human interaction, often faces challenges such as delayed feedback and scalability (Kang, 2023; Williams, 2024). In contrast, AI chatbots provide immediate, consistent, and scalable feedback, which likely contributed to the superior performance of the AI-enabled groups. These results support the notion that AI can complement traditional teaching by addressing its shortcomings, particularly in resource-intensive tasks like academic writing (Maphoto, 2024; Selim, 2024)

The findings have significant implications for integrating AI chatbots into academic writing curricula, particularly in EFL contexts. The success of Claude, ChatGPT, and Google Gemini suggests that AI tools can enhance students' ability to produce well-organized, coherent, and stylistically appropriate texts, addressing common challenges faced by EFL learners (Aljuaid, 2024; ROA, 2024). By providing real-time feedback and revision guidance, AI chatbots enable students to engage in iterative writing processes, fostering greater confidence and independence in their writing tasks (Amirjalili, 2024; Kurniati & Fithriani, 2022). This is particularly relevant for TEFL students, who often require extensive practice to master academic writing conventions.

Moreover, the study highlights the potential of AI to personalize learning experiences, aligning with research on adaptive learning systems (ÇAYIR, 2023; Zheng & Badarch, 2022). The ability of AI chatbots to tailor feedback to individual needs likely contributed to the improved performance across all AI-enabled groups. However, the varying effectiveness among the chatbots suggests that educators must carefully select tools based on their alignment with specific pedagogical goals. For instance, Claude's superior performance indicates it may be particularly effective for advanced

writing tasks requiring nuanced feedback, while ChatGPT and Google Gemini may be better suited for foundational or intermediate writing support.

The study's small sample size (n=36) and focus on academic writing limit the generalizability of the findings. Future research should involve larger and more diverse participant groups to validate these results across different contexts and language skills. Additionally, longitudinal studies could explore the long-term impact of AI-enhanced instruction on writing proficiency and academic success. Investigating the specific features of AI chatbots, such as their feedback algorithms or user interfaces, could provide deeper insights into their varying effectiveness.

Furthermore, the study did not examine the qualitative experiences of students using AI chatbots, such as their perceptions of usability or engagement. Future research could incorporate student and instructor feedback to understand the practical and emotional dimensions of AI integration, complementing quantitative measures of performance (Marmoah, 2024). Exploring hybrid models that combine AI and traditional instruction could also offer a balanced approach, preserving the human elements of teaching while leveraging AI's scalability and efficiency (Çerçi, 2023).

6. Conclusion

This study provides robust evidence that AI-enhanced instruction, utilizing chatbots such as Claude, ChatGPT, and Google Gemini, significantly outperforms traditional teacher-centered methods in improving academic writing skills among TEFL MA students at Ilam University, Iran. The superior performance of the Claude group, with a combined mean essay score of 89.35 ± 4.0, followed by ChatGPT (85.5 ± 4.8) and Google Gemini (82.55 ± 6.0), compared to the traditional instruction group (77.65 ± 7.0), underscores the efficacy of AI chatbots in delivering personalized, immediate, and high-quality feedback. These findings align with prior research highlighting AI's transformative potential in education, particularly in fostering writing efficiency, coherence, and academic style through tailored support (Li, 2024; Minh, 2024). Claude's leading performance suggests its advanced natural language processing capabilities are particularly well-suited for academic writing pedagogy, offering nuanced feedback that closely simulates instructor-like guidance.

The significant differences in essay scores, supported by a one-way ANOVA ($F = 14.57$, $p < 0.001$) and large effect sizes (e.g., Cohen's $d = 1.82$ for Claude vs. traditional instruction), confirm that AI chatbots enhance students' ability to produce well-organized, coherent, and stylistically appropriate texts. Additionally, the high satisfaction ratings for AI-enabled groups (Claude: 4.8 ± 0.3; ChatGPT: 4.4 ± 0.5; Google Gemini: 4.1 ± 0.6) compared to traditional instruction (3.6 ± 0.8) indicate that students found AI tools engaging and effective, further supporting their value in educational settings. These results highlight AI's ability to address key challenges in traditional instruction, such as delayed feedback and scalability limitations, particularly in resource-intensive tasks like academic writing (Kang, 2023; Williams, 2024).

The findings have significant implications for integrating AI chatbots into academic writing curricula, especially in EFL contexts. By providing real-time feedback and revision guidance, AI tools empower students to engage in iterative writing processes, fostering greater confidence and independence (Amirjalili, 2024; Kurniati & Fithriani, 2022). However, the varying effectiveness among the chatbots suggests that educators must carefully select tools based on their alignment with specific pedagogical

goals. The study also underscores the need to address ethical concerns, such as data privacy and academic integrity, to ensure responsible AI use (Nazim, 2024; Williams, 2024). Professional development for educators and robust technological infrastructure are essential to maximize AI's benefits while mitigating challenges like the digital divide (Chouhan, 2023; Nkechi et al., 2024).

Despite its contributions, the study has limitations. The small sample size (n=36) and focus on a single academic writing course limit the generalizability of the findings. Additionally, the study did not explore qualitative aspects of students' experiences with AI chatbots, such as usability or engagement. Future research should involve larger, more diverse cohorts and longitudinal designs to assess the long-term impact of AI-enhanced instruction on writing proficiency. Investigating the specific features of AI chatbots, such as their feedback algorithms, and exploring hybrid models that combine AI and traditional instruction could further enhance pedagogical approaches. By addressing these areas, educators and researchers can better navigate the complexities of AI integration, ensuring that these tools complement human instruction while fostering critical thinking and creativity in academic writing.

Acknowledgments

We extend our sincere appreciation to all individuals whose contributions facilitated the completion of this project.

Conflict of Interest

The author declares no conflicts of interest.

References

Ahn, A. (2024). Should academics be concerned about articles written by ChatGPT? *Developments in Health Sciences*, *7*(1), 25-27. https://doi.org/10.1556/2066.2024.00062

Aithal, P. (2023). Optimizing the use of artificial intelligence-powered GPTs as teaching and research assistants by professors in higher education institutions: a study on smart utilization. *International Journal of Management Technology and Social Sciences*, *8*(4), 368-401. https://doi.org/10.47992/ijmts.2581.6012.0326

Akinwalere, S., & Ivanov, V. (2022). Artificial intelligence in higher education: challenges and opportunities. *Border Crossing*, *12*(1), 1-15. https://doi.org/10.33182/bc.v12i1.2015

Al-Bukhrani, M. (2025). Adoption of AI writing tools among academic researchers: a theory of reasoned action approach. *Plos One*, *20*(1), e0313837. https://doi.org/10.1371/journal.pone.0313837

Aljuaid, H. (2024). The Impact of Artificial Intelligence Tools on Academic Writing Instruction in Higher Education: A Systematic Review. *Arab World English Journal (AWEJ)* Special Issue on ChatGPT, April 2024:26-55. https://dx.doi.org/10.24093/awej/ChatGPT.2

Allahverdiyeva, N. (2024). The role, importance and application of artificial intelligence in the creation of digital education. *Scientific Work*, *18*(2), 120-125. https://doi.org/10.36719/2663-4619/99/120-125

Amirjalili, F. (2024). Exploring the boundaries of authorship: a comparative analysis of AI-generated text and human academic writing in English literature. *Frontiers in Education*, *9*. https://doi.org/10.3389/feduc.2024.1347421

Begum, I. (2024). Role of artificial intelligence in higher education- an empirical investigation. *International Research Journal on Advanced Engineering and Management (IRJAEM)*, *2*(3), 49-53. https://doi.org/10.47392/irjaem.2024.0009

Cao, W., Wang, Q., Sbeih, A., & Shibly, F. (2020). Artificial intelligence based efficient smart learning framework for education platform. *Inteligencia Artificial*, *23*(66), 112–123. https://doi.org/10.4114/intartif.vol23iss66pp112-123

ÇAYIR, A. (2023). A literature review on the effect of artificial intelligence on education. *İnsan Ve Sosyal Bilimler Dergisi*, *6*(2), 276-288. https://doi.org/10.53048/johass.1375684

Çerçi, P. (2023). Allergic to ChatGPT? introducing a desensitization protocol to embrace artificial intelligence. *International Archives of Allergy and Immunology*, *184*(9), 903-905. https://doi.org/10.1159/000531785

Chouhan, D. (2023). A study on artificial intelligence in education. *International Journal of Engineering Technology and Management Sciences*, *7*(1), 449-456. https://doi.org/10.46647/ijetms.2023.v07i01.063

Duran, V. (2024). Analyzing teacher candidates' arguments on AI integration in education via different chatbots. *Digital Education Review*, *45*, 68-83. https://doi.org/10.1344/der.2024.45.68-83

Gervacio, M. (2023). Breaking barriers in academic writing. *International Journal of Asian Education*, *4*(4), 256-264. https://doi.org/10.46966/ijae.v4i4.361

Guleria, A. (2023). Chatgpt: ethical concerns and challenges in academics and research. *The Journal of Infection in Developing Countries*, *17*(9), 1292-1299. https://doi.org/10.3855/jidc.18738

Hryciw, B. (2023). Guiding principles and proposed classification system for the responsible adoption of artificial intelligence in scientific writing in medicine. *Frontiers in Artificial Intelligence, 6*. https://doi.org/10.3389/frai.2023.1283353

Hutson, J. (2024). Embracing AI in English composition: insights and innovations in hybrid pedagogical practices. *International Journal of Changes in Education*, *1*(1), 19-31. https://doi.org/10.47852/bonviewijce42022290

Hz, B. (2023). Navigating English writing proficiency tests in the era of artificial intelligence. *Journal of English Education and Teaching*, *7*(3), 480-498. https://doi.org/10.33369/jeet.7.3.480-498

Ilieva, G. (2023). Effects of generative chatbots in higher education. *Information*, *14*(9), 492. https://doi.org/10.3390/info14090492

Kang, S. (2023). Awareness of using chatbots and factors influencing usage intention among nursing students in South Korea: a descriptive study. *Child Health Nursing Research*, *29*(4), 290-299. https://doi.org/10.4094/chnr.2023.29.4.290

Karampelas, A. (2021). Artificial intelligence and machine learning in the steam classroom. *Hellenic Journal of Stem Education*, *1*(2), 59-66. https://doi.org/10.51724/hjstemed.v1i2.13

Kim, J. (2024). Exploring students' perspectives on generative AI-assisted academic writing. *Education and Information Technologies*, *30*(1), 1265-1300. https://doi.org/10.1007/s10639-024-12878-7

Kong, S. (2024). A pedagogical design for self-regulated learning in academic writing using text-based generative artificial intelligence tools: 6-p pedagogy of plan,

prompt, preview, produce, peer-review, portfolio-tracking. *Research and Practice in Technology Enhanced Learning*, *19*, 030. https://doi.org/10.58459/rptel.2024.19030

Krajka, J. (2024). "AI, will you help?" How learners use artificial intelligence when writing. *Xlinguae*, *17*(1), 34-48. https://doi.org/10.18355/xl.2024.17.01.03

Kurniati, E., & Fithriani, R. (2022). Post-graduate students' perceptions of Quillbot utilization in English academic writing class. *Journal of English Language Teaching and Linguistics*, *7*(3), 437-451. https://doi.org/10.21462/jeltl.v7i3.852

Li, A. (2024). Synergizing with AI and AI-assisted technologies in L2 academic writing: a response to Carlson et al. (2023). *TESOL Journal*, *15*(4). https://doi.org/10.1002/tesj.883

Li, J. (2024). Exploring the potential of artificial intelligence to enhance the writing of English academic papers by non-native English-speaking medical students - the educational application of ChatGPT. *BMC Medical Education*, *24*(1). https://doi.org/10.1186/s12909-024-05738-y

Maphoto, K. (2024). Advancing students' academic excellence in distance education: exploring the potential of generative AI integration to improve academic writing skills. *Open Praxis*, *16*(2), 142-159. https://doi.org/10.55982/openpraxis.16.2.649

Marmoah, S. (2024). Leveraging AI to optimize English academic writing (EAW) in intelligent decision support systems (IDSS). *Jurnal Ilmiah Universitas Batanghari Jambi*, *24*(2), 1265-1270. https://doi.org/10.33087/jiubj.v24i2.5483

Miao, J. (2023). Ethical dilemmas in using AI for academic writing and an example framework for peer review in nephrology academia: a narrative review. *Clinics and Practice*, *14*(1), 89-105. https://doi.org/10.3390/clinpract14010008

Minh, A. (2024). Leveraging ChatGPT for enhancing English writing skills and critical thinking in university freshmen. *Journal of Knowledge Learning and Science Technology*, *3*(2), 51-62. https://doi.org/10.60087/jklst.vol3.n2.p62

Mondal, H. (2023). ChatGPT in academic writing: maximizing its benefits and minimizing the risks. *Indian Journal of Ophthalmology*, *71*(12), 3600-3606. https://doi.org/10.4103/ijo.ijo_718_23

Moussa, A., & Belhiah, H. (2024). Beyond Syntax: Exploring Moroccan Undergraduate EFL Learners' Engagement with AI-Assisted Writing. *Arab World English Journal (AWEJ)* Special Issue on ChatGPT, April 2024: 138-155. https://dx.doi.org/10.24093/awej/ChatGPT.9

Munir, H., Vogel, B., & Jacobsson, A. (2022). Artificial intelligence and machine learning approaches in digital education: a systematic revision. *Information*, *13*(4), 203. https://doi.org/10.3390/info13040203

Nadhifah, A. (2024). Pre-service English teacher perceptions of AI in writing skills. *Journal of World Englishes and Educational Practices*, *6*(2), 26-32. https://doi.org/10.32996/jweep.2024.6.2.3

Nazim, M. (2024). Exploring EFL teachers' insights regarding artificial intelligence driven tools in student-centered writing instructions. *International Journal of English Linguistics*, *14*(3), 93. https://doi.org/10.5539/ijel.v14n3p93

Ng, D. (2024). Empowering student self-regulated learning and science education through ChatGPT: a pioneering pilot study. *British Journal of Educational Technology*, *55*(4), 1328-1353. https://doi.org/10.1111/bjet.13454

Nkechi, A. A., Ojo, A. O., & Eneh, O. A. (2024). Impact of artificial intelligence in achieving quality education. In *Artificial Intelligence*. https://doi.org/10.5772/intechopen.1004871

Pack, A. (2024). Using artificial intelligence in TESOL: some ethical and pedagogical considerations. *TESOL Quarterly*, *58*(2), 1007-1018. https://doi.org/10.1002/tesq.3320

PAN, J. (2024). AI-driven English language learning program and academic writing integrity in the era of intelligent interface. *English Language Teaching and Linguistics Studies*, *6*(4), p.120. https://doi.org/10.22158/eltls.v6n4p120

Parsakia, K. (2023). The Effect of Chatbots and AI on The Self-Efficacy, Self-Esteem, Problem-Solving and Critical Thinking of Students. *Health Nexus*, *1*(1), 71-76. https://doi.org/10.61838/kman.hn.1.1.11

Puspitaningsih, S., Irhadtanto, B., & Ratna Puspananda, D. (2022). The Role of Artificial Intelligence in Children's Education for A Digital Future. *KnE Social Sciences*, *7*(19), 642–647. https://doi.org/10.18502/kss.v7i19.12483

Qawaqneh, H. (2023). The impact of artificial intelligence-based virtual laboratories on developing students' motivation towards learning mathematics. *International Journal of Emerging Technologies in Learning (IJET)*, *18*(14), 105-121. https://doi.org/10.3991/ijet.v18i14.39873

Rababah, L. (2024). Graduate students' ChatGPT experience and perspectives during thesis writing. *International Journal of Engineering Pedagogy (IJEP)*, *14*(3), 22-35. https://doi.org/10.3991/ijep.v14i3.48395

Rabbianty, E. (2023). AI in academic writing: assessing current usage and future implications. *Insania Jurnal Pemikiran Alternatif Kependidikan*, *28*(1a), 14-35. https://doi.org/10.24090/insania.v28i1a.9278

Rahman, N., Zulkornain, L., & Hamzah, N. (2022). Exploring artificial intelligence using automated writing evaluation for writing skills. *Environment-Behaviour Proceedings Journal*, *7*(SI9), 547-553. https://doi.org/10.21834/ebpj.v7isi9.4304

ROA, A. (2024). The impact of AI-powered software on second language (l2) writing: a systematic literature review. R*esearch and Innovation in Applied Linguistics-Electronic Journal*, *2*(2), 138. https://doi.org/10.31963/rial.v2i2.4801

Royani, I. (2024). ChatGPT in the academic writing classroom. *Esteem Journal of English Education Study Programme*, *7*(2), 630-638. https://doi.org/10.31851/esteem.v7i2.15615

Sanasintani, S. (2023). Revitalizing the higher education curriculum through an artificial intelligence approach: an overview. J. *Social Science Utilizing Tech*, *1*(4), 239-248. https://doi.org/10.55849/jssut.v1i4.670

Sari, J., & Purwanta, E. (2021). The implementation of artificial intelligence in stem-based creative learning in the society 5.0 era. *Tadris Jurnal Keguruan Dan Ilmu Tarbiyah*, *6*(2), 433-440. https://doi.org/10.24042/tadris.v6i2.10135

Selim, A. (2024). The transformative impact of AI-powered tools on academic writing: perspectives of EFL university students. *International Journal of English Linguistics*, *14*(1), 14. https://doi.org/10.5539/ijel.v14n1p14

Shopovski, J. (2024). Generative Artificial Intelligence, AI for Scientific Writing: A Literature Review. *Preprints*. https://doi.org/10.20944/preprints202406.0011.v1

Siminto, S. (2023). Educational management innovation by utilizing artificial intelligence in higher education. *Al-Fikrah Jurnal Manajemen Pendidikan*, *11*(2), 284. https://doi.org/10.31958/jaf.v11i2.11860

Song, C. (2023). Enhancing academic writing skills and motivation: assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1260843

Tang, A. (2023). The importance of transparency: declaring the use of generative artificial intelligence (AI) in academic writing. *Journal of Nursing Scholarship*, *56*(2), 314-318. https://doi.org/10.1111/jnu.12938

Tapalova, O., & Zhiyenbayeva, N. (2022). Artificial intelligence in education: AIED for personalised learning pathways. *The Electronic Journal of E-Learning*, *20*(5), 639-653. https://doi.org/10.34190/ejel.20.5.2597

Tian, W. (2024). AI chatbots in Chinese higher education: adoption, perception, and influence among graduate students—an integrated analysis utilizing UTAUT and ECM models. *Frontiers in Psychology*, *15*. https://doi.org/10.3389/fpsyg.2024.1268549

Tran, T. (2024). AI tools in teaching and learning English academic writing skills. *Proceedings of the Asiacall International Conference*, 4, 170-187. https://doi.org/10.54855/paic.23413

Wahjusaputri, S., Bunyamin, B., Nastiti, T., & Sopandi, E. (2023). Strategy for implementing artificial based learning models in improving learning quality. *Dinamika Pendidikan*, *18*(1), 1-9. https://doi.org/10.15294/dp.v18i1.40119

Wale, B. (2024). The transformative power of AI writing technologies: enhancing EFL writing instruction through the integrative use of writerly and Google Docs. *Human Behavior and Emerging Technologies*, 2024, 1-15. https://doi.org/10.1155/2024/9221377

William, F. (2024). AI in academic writing: Ally or foe?. *International Journal of Research Publications*, *148*(1). https://doi.org/10.47119/ijrp1001481520246427

Williams, R. (2024). The ethical implications of using generative chatbots in higher education. *Frontiers in Education*, *8*. https://doi.org/10.3389/feduc.2023.1331607

Wu, R., & Yu, Z. (2023). Do AI chatbots improve students' learning outcomes? evidence from a meta-analysis. *British Journal of Educational Technology*, *55*(1), 10-33. https://doi.org/10.1111/bjet.13334

Yahaya, S. (2024). Evaluating students' acceptance of AI chatbots to enhance virtual collaborative learning in Malaysia. *International Journal of Sustainable Development and Planning*, *19*(1), 209-219. https://doi.org/10.18280/ijsdp.190119

Yang, H. (2024). Empowerment of artificial intelligence in teaching reform of ideological and political courses in universities. *Journal of Contemporary Educational Research*, *8*(1), 80-87. https://doi.org/10.26689/jcer.v8i1.5976

Yazid, S. (2024). ChatGPT, smart writing assistant chatbot for students: an analysis of its drawbacks. *Abjadia International Journal of Education*, *9*(1), 41-56. https://doi.org/10.18860/abj.v9i1.24459

Yeo, M. (2023). Academic integrity in the age of artificial intelligence (AI) authoring apps. *TESOL Journal*, *14*(3). https://doi.org/10.1002/tesj.716

Zhao, X. (2024). ChatGPT and the digitisation of writing. *Humanities and Social Sciences Communications*, *11*(1). https://doi.org/10.1057/s41599-024-02904-x

Zheng, R., & Badarch, T. (2022). Research on applications of artificial intelligence in education. *American Journal of Computer Science and Technology*, *5*(2), 72. https://doi.org/10.11648/j.ajcst.20220502.17