www.ijte.net

AI in Academic Writing: Assessing the Effectiveness, Grading Consistency, and Student Perspectives of ChatGPT and You.com for EFL Students

Mustafa Özdere
Niğde Ömer Halisdemir University, Turkey

To cite this article:

Ozdere, M. (2025). AI in academic writing: Assessing the effectiveness, grading consistency, and student perspectives of ChatGPT and You.com for EFL students. *International Journal of Technology in Education (IJTE), 8*(1), 123-154. https://doi.org/10.46328/ijte.1001

International Journal of Technology in Education
2025, Vol. 8, No. 1, 123-154 https://doi.org/10.46328/ijte.1001

AI in Academic Writing: Assessing the Effectiveness, Grading Consistency, and Student Perspectives of ChatGPT and You.com for EFL Students

Mustafa Özdere

Article Info

Article History

Received:
10 September 2024
Accepted:
24 December 2024

Keywords

AI-generated feedback
EFL writing improvement
Grading consistency
Student perceptions

Abstract

This study evaluates the impact of AI-generated feedback and targeted training sessions on the writing skills of 16 B1-B2 level students majoring in English Language and Literature (ELL) or English Language Teaching (ELT) at Niğde Ömer Halisdemir University, School of Foreign Languages, during the 2023-2024 spring term. Using a mixed-methods approach, the research assessed students' writing across four paragraph types—opinion, cause/effect, reason/result, and advantage/disadvantage—employing a rubric that rates Grammar and Mechanics, Sentence Structure, Cohesion and Coherence, Argumentation/Content Development, Vocabulary Usage, and Task Achievement out of 30. Both ChatGPT and You.com were used to grade the initial and revised drafts of students' paragraphs. Following the first draft feedback, students participated in targeted training sessions aimed at addressing identified weaknesses. Quantitative analysis of draft scores, supported by paired t-tests, independent t-tests, and ANOVA, revealed significant improvements in writing performance from the first to the final drafts, with You.com generally providing higher final draft scores compared to ChatGPT. Correlation analysis affirmed the reliability of both AI models. Qualitative insights, obtained from a focus group interview, highlighted the perceived effectiveness of the AI feedback and training sessions. The study underscores the potential of AI-generated feedback as a valuable educational tool, offering timely, individualized support that enhances writing skills, with supplementary training sessions further boosting the efficacy of AI feedback. This research contributes to understanding the role of AI in enhancing writing instruction and feedback within English as a Foreign Language (EFL) education, aiming to inform pedagogical practices and integrate AI tools effectively in writing classrooms.

Introduction

Background of AI in ELT Education

The integration of Artificial Intelligence (AI) in language education has emerged as a transformative force, reshaping traditional pedagogical approaches and enhancing learning experiences across various linguistic contexts. AI technologies, particularly in English Language Teaching (ELT), have demonstrated significant

potential to personalize learning, provide real-time feedback, and foster learner autonomy, all of which enhance student outcomes and engagement (Liu, 2023; Yunina, 2023). The ability of AI to cater to individual needs through adaptive learning environments moves away from the conventional "one-size-fits-all" model, which has dominated language classrooms (Akbarani, 2024). This level of personalization is crucial in second language acquisition, where tailored feedback can significantly impact learner motivation and performance (Betal, 2023; Gyawali, 2022). Recent technological advances in AI have led to the development of various AI-powered writing assistance tools, which are increasingly used by students to improve their writing skills. These tools include automated writing evaluation tools, automated writing corrective feedback tools, AI-powered machine translators, and GPT-3 automatic text generators (Alharbi, 2023). A recent study demonstrated that ChatGPT could significantly improve foreign language writing in terms of vocabulary and grammar, particularly for socially vulnerable populations such as refugees and migrants (Athanassopoulos, 2023).

AI's role in language education extends beyond mere personalization. It has revolutionized methodologies through AI-powered tools like chatbots and intelligent tutoring systems, which facilitate writing skills and offer immediate, formative feedback. These tools encourage critical thinking and creativity, pushing the boundaries of traditional educational paradigms (Fitria, 2023; Pack & Maloney, 2023). For instance, applications like ChatGPT have shown promise in enhancing students' writing capabilities, providing instant feedback, and allowing learners to refine their work iteratively (Patiño, 2024). Similarly, AI-driven applications enhance oral communication skills by offering immediate, tailored feedback, crucial for language development (Demianenko, 2024; Konyrova, 2024). A study highlighted that AI-assisted writing is expected to be widely adopted in the workplace, necessitating significant changes in instruction to address challenges such as reduced critical thinking and authenticity, while also enhancing efficiency and idea generation (Cardon, 2023).

Despite its many benefits, the integration of AI in language education presents challenges. One major concern is that AI, while efficient in providing technical feedback, cannot replace the emotional intelligence and mentorship that human teachers offer (Liu, 2023). This human element is essential in addressing the emotional and social aspects of learning, which AI cannot replicate. Ethical concerns also arise, particularly regarding data privacy and potential biases inherent in AI algorithms, emphasizing the need for clear ethical guidelines (Kasneci et al., 2023; Semerikov, 2021). Furthermore, over-reliance on AI can risk diminishing the role of traditional teaching methods, which remain essential for providing a holistic educational experience (Antara, 2024; Khan, 2023).

Moreover, AI can be an invaluable tool for educators by streamlining assessment processes and enhancing instructional strategies. AI-enabled automatic scoring systems, for example, offer insights into student performance, allowing educators to focus on areas that need improvement (Fu et al., 2020). However, teacher training in AI technologies is crucial to fully harness these benefits and mitigate the risks associated with its implementation (Cantos, 2023).In conclusion, while AI has the potential to transform language education by providing personalized learning experiences and improving pedagogical practices, its integration must be balanced with human elements of teaching. The future of language education will likely involve a synergistic relationship between AI technologies and traditional teaching methods, leading to more effective and inclusive learning experiences (Bozkurt et al., 2021). Educators and institutions must embrace these advancements while

addressing the challenges they present to ensure the creation of dynamic and engaging educational environments.

Importance of Feedback in Writing Skill Development

Feedback plays a crucial role in developing writing skills, particularly for students learning English as a foreign language (EFL). Research consistently shows that various forms of feedback—whether from teachers, peers, or rubrics—significantly improve students' writing abilities. For example, Hasan (2022) highlights that rubric-based feedback, with its structured and criteria-driven approach, offers clear expectations and guidance, resulting in more substantial improvements in high school students' writing compared to traditional methods. This structured feedback not only helps clarify what is expected but also fosters a deeper understanding of effective writing practices.

Peer feedback also plays a vital role in the writing development process. Peer correction fosters critical engagement with writing by providing students with fresh perspectives. Nguyen and Nguyen (2021) argue that it enhances writing abilities and promotes reflective practices. Similarly, Lin et al. (2021) found that students who utilized peer feedback made significant improvements in their revisions, highlighting the benefits of incorporating peer suggestions. This is further supported by Trisnadewi (2021), who noted that peer correction effectively improves writing proficiency by encouraging reflection and peer learning. Likewise, Gonzalez-Torres and Sarango (2023) found that students value a combination of teacher and peer feedback, which they found beneficial for enhancing their writing performance. The collaborative nature of peer feedback encourages students to engage more critically with their writing, learning to both give and receive constructive critiques. A study on Pakistani postgraduate ESL learners found that peer feedback, supported by formal training sessions, fosters a collaborative learning environment, enhances learners' autonomy, and develops critical and analytical skills, despite some challenges such as biases and lack of peer feedback skills (Fareed, 2021).

The timing of feedback is equally important. Studies show that prompt feedback, whether from teachers or peers, significantly aids students in refining their writing (Nicolas, 2020). Nguyen and Nguyen (2021) also found that continuous formative assessment, facilitated by daily feedback, leads to significant gains in students' academic writing skills. This supports the idea that ongoing, immediate feedback is essential for improving writing proficiency. In addition to timing, the quality of feedback is paramount. Adawiyah and Sumarni (2021) argue that constructive feedback not only points out areas for improvement but also motivates students to correct their errors. This balance between focusing on strengths and weaknesses fosters a positive learning environment where students feel empowered to improve their writing. Bijami et al. (2013) further emphasize that integrating both teacher and peer feedback provides a comprehensive system that addresses various writing aspects, from content development to grammatical accuracy. A study on the introduction of email feedback in a private university in Lebanon revealed that while students responded positively, instructors were resistant, highlighting the need for training in ELT and digital literacies (Coursey & Dandashly, 2015).

In conclusion, feedback is indispensable for the development of writing skills, serving not only as a tool for correction but also as a catalyst for collaborative learning and self-improvement. The combination of structured,

timely, and constructive feedback from both teachers and peers creates an effective learning environment that greatly enhances students' writing abilities.

The Role of AI-Generated Feedback in Supporting Learning

The integration of artificial intelligence (AI) in educational settings has revolutionized feedback mechanisms, offering personalized and timely responses that significantly enhance learning experiences. AI-generated feedback supports students by providing immediate assessments and tailored insights, enabling learners to identify strengths and weaknesses more effectively. Utilizing advanced algorithms, such as natural language processing, deep learning, and reinforcement learning, AI systems can analyze student responses to assignments like essays and tests, delivering accurate and objective evaluations (Akavova, 2023; Owan, 2023). This capability minimizes potential errors and biases commonly found in traditional grading methods, fostering a more equitable learning environment (Yunina, 2023).The adaptability of AI allows for the customization of feedback based on individual student needs. AI tools can assess performance, adjust task complexity, and offer feedback that is both relevant and constructive (Patiño, 2024). This personalized approach is particularly effective in language learning, where exercises can be tailored to improve oral and written communicative competence according to the student's proficiency level (Pang, 2024). AI-generated feedback provides actionable and clear guidance, which is essential for students to refine their skills in real-time, particularly in online learning environments where traditional feedback mechanisms might be delayed (Rakya, 2023; Seo et al., 2021). An experiment revealed that people value generative AI-powered writing assistance for its ability to enhance productivity and confidence, though it also poses risks such as reduced accountability and diversity in writing (Li, 2024).

Beyond student learning, AI-generated feedback also benefits educators by reducing their workload, allowing them to focus on higher-order instructional tasks. AI systems can automate routine evaluations, such as grading and formative assessments, freeing educators to concentrate on fostering critical thinking and deeper engagement with course material (Tubino & Adachi, 2022; Soelistiono, 2023). Research supports the idea that AI-assisted feedback enhances feedback literacy among students, teaching them to better interpret and apply the feedback they receive (Tubino & Adachi, 2022). This integration of AI can lead to a more streamlined and effective educational experience for both students and teachers.

However, the implementation of AI-generated feedback is not without challenges. One key concern is the reliability and perceived trustworthiness of AI assessments, which can impact their acceptance by both students and educators (Lin et al., 2022; Nazaretsky et al., 2022). Although AI can provide immediate and objective feedback, it may lack the nuanced understanding and empathy that human instructors offer, which is vital in building a supportive and motivational learning environment (Drewery et al., 2022). Additionally, biased or inaccurate feedback can arise from the knowledge base and algorithms used in AI systems, necessitating ongoing improvements to these technologies to ensure accuracy and fairness (Kamruzzaman, 2023). Professional development and training for educators in using AI tools are essential for building trust and ensuring the successful integration of AI in feedback systems (Strzelecki, 2023).

Despite these challenges, AI's potential to enhance educational practices remains significant. AI-driven tools facilitate formative assessments and continuous feedback loops that promote self-regulated learning (Carlson, 2023). By analyzing student interactions with course materials, AI can offer insights that help learners pinpoint areas needing improvement, thus promoting a more engaged and reflective learning process (Wang & Han, 2020). Moreover, as AI handles routine feedback tasks, teachers are empowered to focus on more complex instructional roles, improving overall teaching and learning outcomes (Owan, 2023).

AI-generated feedback has a transformative effect on education, particularly when combined with targeted training sessions. By analyzing student performance in real-time, AI provides personalized, timely feedback that identifies areas for improvement and suggests tailored interventions. This adaptability supports individualized learning paths, fostering student engagement and improving outcomes (Akavova, 2023; Naz, 2024). In various contexts, such as medical education, AI helps design focused training on specific competencies, enhancing student success (Nagi et al., 2023).

To maximize the benefits of AI feedback, educators need professional development to enhance their digital skills. Training programs improve teachers' ability to effectively use AI tools, empowering them to guide students in personalized learning experiences (Ng et al., 2023). However, challenges remain, including concerns about bias in AI feedback and the need for ethical oversight (Lin et al., 2022). Despite these hurdles, AI-driven feedback, when coupled with appropriate training, has the potential to significantly enhance educational practices. In conclusion, AI-generated feedback represents a transformative force in education, offering personalized, timely, and objective assessments that significantly enhance student learning. While there are concerns regarding trust, reliability, and the human touch in AI feedback, the potential benefits are substantial. As AI continues to evolve, it is crucial to strike a balance between AI and human interaction in feedback provision, ensuring that AI serves as a supportive tool rather than a replacement for traditional teaching methods. By addressing these challenges, AI can create a more tailored and efficient educational experience, benefiting both learners and educators.

Scientific Relevance and Contributions of this Study

The impact of artificial intelligence (AI) tools on the development of writing skills among non-native English learners has garnered significant attention in recent academic discourse. Empirical studies show that AI-powered writing assistants like Grammarly and Wordtune enhance the writing quality and confidence of English as a Foreign Language (EFL) students. Selim (2024) and Mahmud (2023) found that these tools improve writing skills and provide constructive feedback, though ongoing support is needed. Nazari et al. (2021) demonstrated that AI tools boost self-efficacy and engagement in writing tasks. Marzuki et al. (2023), Koka, Khan, and Jan (2023) noted that immediate feedback from AI tools aids in error correction and fosters self-reflection, helping learners address writing errors in real-time.

However, concerns about over-reliance on AI tools suggest they may hinder the development of independent writing skills. Hz, Laiya, Sarumaha, and Supiyandi (2023) note that while AI tools assist in writing, they may lead to diminished skill acquisition if not used judiciously. Rahman et al. (2022) emphasize that AI tools should

complement, not replace, traditional learning methods, highlighting the importance of formative feedback and evaluation in developing ESL students' writing skills.

The integration of AI tools into writing instruction offers personalized learning experiences. Konyrova (2024) notes that AI platforms can tailor modules to individual student needs, creating a more immersive learning environment. This adaptability is crucial for EFL learners facing unique language acquisition challenges. Additionally, consistent feedback from Automated Writing Evaluation (AWE) tools can accelerate learning, allowing students to develop writing skills at their own pace (Parra & Calero, 2019).

Despite extensive research on AI tools in education, there is limited empirical evidence on their impact on specific skill development, particularly writing, for non-native English learners. This study fills a critical gap by evaluating the effectiveness of AI-generated feedback in improving university students' academic writing skills. Unlike previous studies focusing on technical capabilities, this research assesses the pedagogical value of AI by analyzing improvements in student writing over multiple drafts.

The contributions of this research are twofold. First, it provides comprehensive quantitative and qualitative data on the impact of AI feedback tools, bridging the gap between technological development and classroom application. Second, it explores student perceptions of AI-generated feedback, adding a human-centric perspective. These insights are crucial as educators and policymakers seek to responsibly integrate AI technologies into educational frameworks. The findings will contribute to academic discourse by offering actionable recommendations for using AI-generated feedback in language education. By examining both the benefits and limitations of AI feedback, this research aims to inform future practices and policies, ensuring AI serves as a complementary tool rather than a replacement for traditional teaching methods. This underscores the study's importance in advancing educational practices in an AI-driven era.

The primary aim of this study is to assess the effectiveness of AI-generated feedback in improving the writing skills of prep school students at the B1-B2 proficiency level. Specifically, the study investigates whether AI-generated feedback, combined with targeted improvement sessions, leads to significant improvements in students' writing performance. By examining the impact of this approach, the study seeks to contribute to the growing research on the role of AI in enhancing writing instruction and feedback, particularly within the context of English as a Foreign Language (EFL) education. The findings are intended to inform pedagogical practices and support the integration of AI-powered tools in writing classrooms, ultimately helping to improve the writing skills of language learners.

This study addresses three key research questions:

1. How does AI-generated feedback, combined with targeted improvement sessions, affect the writing performance of students in an academic writing course?
2. What are the differences in grading consistency between ChatGPT and You.com when assessing students' first and final drafts?
3. How do students perceive and evaluate the effectiveness of AI-generated feedback and the subsequent

improvement sessions in enhancing their writing skills?

By employing a mixed-methods approach, the study provides a comprehensive evaluation of both the objective outcomes and the students' subjective experiences with AI-generated feedback.

Methodology

Participants

The study involved 16 students enrolled in the School of Foreign Languages at Niğde Ömer Halisdemir University during the 2023-2024 spring term. These students were majoring in either English Language and Literature (ELL) or English Language Teaching (ELT) and had a B1-B2 level of English proficiency. They were enrolled in a 4-hour-per-week academic writing lesson. The instructor for the course, who is also the researcher, selected this group for practical reasons to evaluate the effectiveness of AI-generated feedback in improving students' writing skills.

Research Design

The study follows a mixed-methods research design, combining both quantitative and qualitative approaches to thoroughly investigate the impact of AI-generated feedback on students' academic writing skills. The rationale for this design is rooted in the need to not only measure the improvement in writing skills through objective data but also understand the participants' perceptions of the AI-generated feedback process. According to Hadi et al. (2012), mixed-methods research integrates qualitative and quantitative approaches to better understand research problems, enabling a more nuanced exploration of issues. This study utilizes an "Explanatory Sequential Design," which involves first collecting quantitative data and then gathering qualitative data to clarify or expand on the quantitative results (Nurhasanah et al., 2022).

Writing Tasks

As part of the academic writing course, students were assigned various writing tasks from the book "Writer 2" by Blackswan Publishing House. These tasks encompassed different types of paragraphs, allowing students to select topics that resonated with their interests. For opinion paragraphs, students could choose from topics such as "Everyone in the world should speak the same language. Do you agree or disagree?", "It is good to have an extended family. Do you agree or disagree?", and "University students should live with their families. Do you agree or disagree?". The cause-effect paragraph topics included "Causes/effects of stress", "Causes/effects of getting married", and "Causes/effects of global warming". For advantage-disadvantage paragraphs, the topics were "Advantages/disadvantages of traveling alone", "Advantages/disadvantages of distance education", and "Advantages/disadvantages of a long-distance relationship". Lastly, the compare or contrast paragraph tasks offered topics such as "Your hometown and another city", "Football and basketball", and "Two celebrities". This variety of topics provided students with the flexibility to engage deeply with subjects they felt passionate about, enhancing their learning experience and writing skills.

AI-Generated Feedback

For this study, a structured process was followed to receive feedback and grade students' academic paragraphs using AI tools. Initially, the first drafts of the students' paragraphs, which included various types such as opinion, cause-effect, compare-contrast, and advantage-disadvantage, were collected. These drafts were then submitted to ChatGPT and You.com, using specific prompts to guide the AI tools in evaluating the paragraphs based on the established rubric. The prompts included requests such as "Evaluate this paragraph for grammar and mechanics, identifying any errors and suggesting corrections," and "Assess the sentence structure, noting the variety and complexity of sentences." Additionally, the AI tools were asked to provide an overall score out of 30 points, considering all six criteria: grammar and mechanics, sentence structure, cohesion and coherence, argumentation and content development, vocabulary usage, and task achievement.

The feedback from the AI tools was then used to design targeted training sessions to address common issues identified in the first drafts. After these training sessions, students revised their paragraphs, and the final drafts were again submitted to the AI tools for evaluation using the same rubric and prompts. This process ensured consistent and detailed feedback, helping students improve their writing skills and providing reliable data for analysis.

Content and Form-Related Aspects: The AI-generated feedback focused on various aspects of writing, including grammar and mechanics, sentence structure, cohesion and coherence, argumentation and content development, vocabulary usage, and task achievement. The feedback was presented in a structured format, with clear headings for each criterion, helping students easily identify areas of strength and areas needing improvement.

Training of AI Tools for Valid Assessment: The AI tools, ChatGPT and You.com, were trained on large datasets of academic writing, including essays, research papers, and other educational materials. This training enabled the AI to understand and evaluate various aspects of writing effectively. The AI tools were validated using a set of benchmark paragraphs graded by human experts. The AI's evaluations were compared to the human grades to ensure accuracy and consistency. Continuous updates and improvements were made to the AI models based on user feedback and new data.

Prompts Used:

- Grammar and Mechanics: "Evaluate this paragraph for grammar and mechanics, identifying any errors and suggesting corrections."
- Sentence Structure: "Assess the sentence structure, noting the variety and complexity of sentences."
- Cohesion and Coherence: "Evaluate the logical flow and organization of ideas in this paragraph."
- Argumentation and Content Development: "Assess the strength and development of arguments in this paragraph."
- Vocabulary Usage: "Evaluate the richness and accuracy of vocabulary used in this paragraph."
- Task Achievement: "Assess whether the paragraph fully addresses the assignment requirements."

Student Interaction with AI Tools: Students submitted their initial drafts to the AI tools for evaluation and received detailed feedback on their drafts, highlighting strengths and areas for improvement. They used the feedback to revise their drafts and improve their writing skills. The revised drafts were resubmitted to the AI tools for a final evaluation.

Construction of Training Sessions: The feedback from the AI tools was analyzed to identify common issues in students' writing. Training sessions were designed to address these common issues, including explicit instruction on specific writing skills, collaborative exercises, and practice activities. The training sessions were conducted, providing students with targeted support to improve their writing. Students revised their drafts based on the training sessions and received additional feedback from the AI tools.

Instruments

Rubric

The rubric for evaluating students' academic paragraphs was meticulously crafted to provide a comprehensive and systematic assessment of writing skills, ensuring consistency and fairness in grading. Drawing on academic research, it evaluates six critical aspects of writing: Grammar and Mechanics, Sentence Structure, Cohesion and Coherence, Argumentation and Content Development, Vocabulary Usage, and Task Achievement, each scored on a scale from 1 to 5. These criteria were selected for their central role in academic writing, as supported by studies such as Razı (2015), Brooks (2012), and Jonsson and Svingby (2007).

Grammar and Mechanics are fundamental to clear communication, with 5 points awarded for flawless writing and 1 point for numerous errors that impede understanding (Razı, 2015). Sentence Structure assesses variety and complexity, with 5 points for effective usage and 1 point for repetitive or flawed constructions. Logical flow and organization are evaluated under Cohesion and Coherence, with 5 points for smooth transitions and a well-organized structure, as highlighted by Brooks (2012). Argumentation and Content Development focus on constructing evidence-based, well-reasoned arguments, with 5 points for strong arguments and 1 point for underdeveloped or unsupported ideas, aligning with Jonsson and Svingby's (2007) emphasis on argument quality. Vocabulary Usage measures richness and accuracy, scoring 5 for varied and appropriate word choice and 1 for limited or inaccurate vocabulary. Lastly, Task Achievement ensures the writing fully addresses the assignment requirements, with 5 points for complete fulfillment and 1 for significant omissions.

By combining these criteria, the rubric provides structured, detailed, and actionable feedback. This systematic approach not only streamlines grading but also supports students in identifying their strengths and areas for growth, ultimately helping them develop essential academic writing skills in a transparent and effective manner.
Construction: The rubric was constructed based on a review of academic literature on effective writing assessment. Key studies that informed the rubric include Razı (2015), Brooks (2012), and Jonsson and Svingby (2007). These studies highlight the importance of evaluating multiple aspects of writing, including grammar and mechanics, sentence structure, cohesion and coherence, argumentation and content development, vocabulary usage, and task achievement. Each criterion is scored on a scale from 1 to 5, with 5 indicating excellent

performance and 1 indicating significant areas for improvement. This scale allows for a nuanced assessment of students' writing skills.

Validation: The rubric was reviewed by a panel of writing instructors and experts in academic writing assessment. Their feedback was used to refine the criteria and scoring scale. The rubric was pilot tested with a sample of student paragraphs. The paragraphs were graded using the rubric, and the results were compared to grades assigned by experienced writing instructors. This comparison helped ensure that the rubric provided consistent and accurate assessments. A reliability analysis was conducted to ensure that the rubric was applied consistently across different drafts and paragraph types.

AI-Tools

I chose to utilize ChatGPT and You.com as my AI platforms due to their unique capabilities that significantly enhance the educational experience. ChatGPT, developed by OpenAI, is an advanced AI language model that generates human-like text and provides comprehensive feedback on student writing. It analyses essays for grammar, vocabulary, coherence, and structure, offering immediate feedback that encourages revisions. This immediate feedback is crucial for fostering continuous improvement in students' writing skills. Additionally, ChatGPT assists educators by grading using predefined rubrics, thereby reducing their workload and allowing them to focus on more complex instructional tasks (Firat, 2023).

You.com is an AI-integrated search engine that delivers personalized results and features an AI chatbot for conversational queries. In an educational context, You.com helps students gather research materials and provides writing suggestions, thereby enhancing the quality of their assignments. This tool supports diverse learning needs by offering tailored insights and fostering engagement through its interactive features (Michel-Villarreal, 2023; Chen, 2024).Both platforms effectively analyze student writing, offering immediate insights that promote continuous learning and engagement. By automating routine feedback, these AI tools enable educators to concentrate on developing higher-order skills such as critical thinking and creativity. Furthermore, they provide data-driven insights into student progress, which can inform and refine instructional strategies (Sallam et al., 2023; Dai, 2023; Teebagy et al., 2023; Zarei, 2024).In conclusion, ChatGPT and You.com represent significant advancements in educational technology, transforming feedback processes and writing instruction. Their combined use in this study ensures a comprehensive evaluation of student writing, leveraging the strengths of both platforms to enhance educational outcomes.

Procedure

The study followed a systematic process:

1. *Submission of Initial Drafts:* Students submitted the initial drafts of their paragraphs.
2. *AI Evaluation:* The drafts were evaluated using ChatGPT and You.com, guided by specific prompts aligned with the rubric.
3. *Feedback Provision:* Detailed AI-generated feedback was provided, identifying strengths and areas for

improvement in content, grammar, structure, and vocabulary.

4. *Training Sessions:* Based on common issues identified in the feedback, training sessions were designed. These sessions included explicit instruction, collaborative exercises, and practice activities.
5. *Revision of Drafts:* Students revised their drafts based on both the feedback and the training sessions.
6. *Submission and Final Evaluation:* The final drafts were submitted and evaluated again by the AI tools using the same rubric and prompts.

This structured approach ensured that students received consistent and detailed feedback, which helped them improve their writing skills systematically.

Data Analysis

Quantitative Approach: Quantitative data is derived from the AI-generated grades for both the initial and final drafts. Each paragraph type is graded by You.com and ChatGPT using a consistent rubric, which allows for a maximum score of 30 points per paragraph. This data is then analyzed using SPSS 29.0, employing various statistical tests to evaluate the impact of AI-generated feedback on students' writing performance. Descriptive statistics summarize students' performance before and after the intervention, while paired sample t-tests assess significant differences between the scores of the first and final drafts for each paragraph type. Independent samples t-tests compare the effectiveness of feedback from You.com and ChatGPT. Correlation analysis examines the relationship between different rubric components and overall performance. ANOVA analysis evaluates differences in performance across the various paragraph types, and reliability analysis ensures consistent application of the rubric across different drafts and paragraph types.

Qualitative Approach: A focus group interview is utilized to gather qualitative data by exploring students' opinions on the usefulness and effectiveness of AI-generated feedback. This method offers valuable insights into students' perceptions of the feedback process and its impact on their learning experience. As a qualitative research approach, focus group interviews involve guided discussions among a small group of participants. This technique effectively captures diverse perspectives on a specific topic, allowing researchers to delve into participants' thoughts and experiences. Typically conducted in a semi-structured format, the moderator employs open-ended questions to steer the discussion, providing consistency while also allowing for the exploration of emerging themes (Baillie, 2019; McLafferty, 2004). Data analysis from these interviews can be complex due to the varied interactions among participants. Researchers often apply thematic or framework analysis to identify key themes and patterns within the discussions (Rabiee, 2004; Gavora, 2015). After the completion of all writing tasks and training sessions, a focus group interview is conducted with the participants to explore their experiences with the AI-generated feedback. This session aims to collect detailed insights into how students perceive the effectiveness of feedback in improving their writing and its impact on their learning process.

Overview of Instructional Design

Purpose and Goals: The primary objective of the instructional design was to improve students' academic writing skills through the integration of AI-generated feedback. The goal was to enhance students' abilities in grammar,

sentence structure, cohesion, coherence, argumentation, vocabulary usage, and task achievement.

Theoretical Framework: The instructional design was grounded in constructivist principles, emphasizing active learning and the importance of feedback in the learning process. It also incorporated elements of formative assessment to provide continuous feedback and support.

Course Structure: The course was structured around a series of writing tasks, including opinion, cause-effect, compare-contrast, and advantage-disadvantage paragraphs. Each task was designed to target specific writing skills and was followed by AI-generated feedback and targeted training sessions.

Learning Outcomes: By the end of the course, students were expected to demonstrate improved writing skills, particularly in the areas of grammar, sentence structure, cohesion, coherence, argumentation, vocabulary usage, and task achievement. They were also expected to effectively use AI-generated feedback to revise and improve their writing.

Integration of AI Tools in the Instructional Design: Role of AI Tools: AI tools like ChatGPT and You.com were integrated into the instructional design to provide immediate, detailed feedback on students' writing. These tools analyzed essays for grammar, vocabulary, coherence, and structure, offering suggestions for improvement.
Implementation Strategy: The implementation strategy involved collecting initial drafts from students, submitting these drafts to the AI tools for evaluation, and using the feedback to design targeted training sessions. Students then revised their drafts based on the feedback and training, and the final drafts were re-evaluated by the AI tools.

Feedback Mechanism: AI-generated feedback was used to inform instructional practices and support student learning. The feedback was detailed and covered various aspects of writing, including grammar, sentence structure, cohesion, coherence, argumentation, vocabulary usage, and task achievement. This feedback was integrated into lesson plans and revision activities to help students improve their writing.

Assessment and Evaluation: The effectiveness of AI tools was assessed through quantitative analysis of students' writing performance before and after receiving feedback. Statistical tests such as paired sample t-tests, independent samples t-tests, and ANOVA were used to evaluate the impact of AI-generated feedback on students' writing skills.

Challenges and Solutions: Challenges encountered during the integration of AI tools included variability in feedback quality and students' ability to apply the feedback. Solutions included providing additional training sessions to help students interpret and use the feedback effectively and ensuring that the AI tools were regularly updated to improve their accuracy and reliability.

Sample Implementation: To illustrate the integration and implementation of AI tools, consider the following example:

Student's First Draft

Advantages of Living with Family at University

In my opinion, university students should live with their families for economic conditions, success in university, and helping out with household chores. First of all, university students can save money more easily because they don't have to pay any dormitory fees or flat rent. They can save money for their educational costs or their basic needs. Second of all, they can focus more on their studies. For example, they aren't being burdened by concerns such as cooking and cleaning. In addition, they're much less likely to be distracted by flat mates for some reason the night before your deadline. Third of all, they can help your parents out with household chores. For instance, they can take the rubbish out or do several household chores. To sum up, living with family has several benefits.

Feedback and Grade:

Grammar and Mechanics (4/5):

- Error: "Second of all" should be "Secondly" for a more formal tone.
- Error: Inconsistent pronoun usage: "they" vs. "your" (e.g., "they're much less likely to be distracted by flat mates for some reason the night before your deadline").
- Correction: Maintain consistent pronoun usage: "they're much less likely to be distracted by flat mates for some reason the night before their deadline."

Sentence Structure (4/5):

- Error: "First of all, university students can save money more easily because they don't have to pay any dormitory fees or flat rent."
- Correction: "Firstly, university students can save money more easily because they don't have to pay dormitory fees or rent."
- Error: "Second of all, they can focus more on their studies."
- Correction: "Secondly, they can focus more on their studies."
- Error: "Third of all, they can help your parents out with household chores."
- Correction: "Thirdly, they can help their parents with household chores."

Cohesion and Coherence (4/5):

- Error: "For example, they aren't being burdened by concerns such as cooking and cleaning."
- Correction: "For example, they are not burdened by concerns such as cooking and cleaning."
- Error: "In addition, they're much less likely to be distracted by flat mates for some reason the night before your deadline."
- Correction: "Additionally, they are much less likely to be distracted by flat mates the night before their deadline."

Argumentation and Content Development (4/5):

- Error: "They can save money for their educational costs or their basic needs."
- Suggestion: Expand on how saving money can benefit their education and personal life.
- Error: "For instance, they can take the rubbish out or do several household chores."

- Suggestion: Provide more specific examples of household chores and how this helps the family.

Vocabulary Usage (4/5):

- Error: "Basic needs" could be more sophisticated.
- Correction: "Essential expenses" or "necessary expenditures."
- Error: "Take the rubbish out" could be more formal.
- Correction: "Dispose of the trash."

Task Achievement (5/5):

- The paragraph fully addresses the task, discussing the advantages of living with family during university and providing clear reasons and examples.

Overall Grade: 25/30

Student's Final Draft

Advantages of Living with Family at University

In my opinion, university students should consider living with their families while they are at university due to several advantages. First of all, university students can more easily save money because they don't have to pay any dormitory fees or flat rent. They can save money for their educational costs or their basic needs. Second of all, they can focus more on their studies. For example, they are not burdened by tasks such as cooking and cleaning. In addition, they're much less likely to be distracted by roommates for some reason the night before a deadline. Third of all, they can help their parents out with household chores. For instance, they can take the rubbish out or do several household chores. To sum up, living with family at university must be considered for several reasons.

Feedback and Grade:

- Grammar and Mechanics (5/5): The final draft is free of grammatical and mechanical errors. The student has corrected previous issues, such as inconsistent pronoun usage and informal phrases.
- Sentence Structure (5/5): The sentence structures are varied and effective. The transitions between points are smooth, and the sentences are concise and clear.
- Cohesion and Coherence (5/5): The ideas are logically organized and flow smoothly. The use of transitional phrases like "First of all," "Second of all," and "Third of all" helps maintain coherence.
- Argumentation and Content Development (4/5): The argument is clear and well-developed with relevant examples. However, the points could still be expanded with more detailed evidence and reasoning to further strengthen the argument.
- Vocabulary Usage (4/5): The vocabulary is appropriate and varied. While the language is clear and effective, there is still room for incorporating more sophisticated vocabulary.
- Task Achievement (5/5): The paragraph fully addresses the task, discussing the advantages of living with family during university and providing clear reasons and examples.

Overall Grade: 28/30

Explanation of the Process

The system was utilized effectively to provide detailed feedback and grading for students' academic paragraphs. For instance, a student's first draft on the advantages of living with family during university was submitted to the AI tools, ChatGPT and You.com. Using specific prompts, the AI evaluated the paragraph based on a comprehensive rubric that included criteria such as grammar and mechanics, sentence structure, cohesion and coherence, argumentation and content development, vocabulary usage, and task achievement. The AI tools identified minor grammatical errors, inconsistencies in pronoun usage, and areas for improvement in sentence structure and vocabulary. They also provided suggestions for enhancing the argument and coherence of the paragraph. This detailed feedback helped the students understand their strengths and areas for improvement, ultimately guiding them to revise their work more effectively. The AI-generated grades and feedback were then used to design targeted training sessions, ensuring a systematic approach to improving students' writing skills. This process not only streamlined the grading but also supported continuous learning and engagement.

Results

The current study employed a mixed-methods approach to comprehensively evaluate the reliability and effectiveness of two AI tools, ChatGPT and You.com, in grading student writing tasks. The analysis focuses on improvements in student performance from the first to the final drafts across different paragraph types and assesses which AI tool demonstrates greater trustworthiness and consistency in evaluation. The following sections present both quantitative and qualitative findings from the research.

Quantitative Data Analysis

Descriptive Statistics: A descriptive analysis was conducted to assess the statistical properties of scores from both AI models across the initial and final drafts for different paragraph types. Table 1 summarizes the descriptive statistics. The descriptive statistics demonstrate improvements across all paragraph types. For instance, in advantage-disadvantage paragraphs, You.com’s scores showed a larger improvement (18.56 to 24.43) compared to ChatGPT (18.31 to 21.71), with final draft scores also exhibiting greater consistency as reflected in lower standard deviations.

Table 1. Descriptive Statistics for Different Paragraph Types

Paragraph Type	Model	Draft	N	Mean Score	Standard Deviation
Opinion	ChatGPT	First Draft	16	19.63	2.754
	ChatGPT	Final Draft	14	23.43	2.503
	You.com	First Draft	16	20.00	1.789
	You.com	Final Draft	14	23.50	1.557
Advantage-Disadvantage	ChatGPT	First Draft	16	18.31	3.114

Paragraph Type	Model	Draft	N	Mean Score	Standard Deviation
	ChatGPT	Final Draft	14	21.71	3.124
	You.com	First Draft	16	18.56	1.711
	You.com	Final Draft	14	24.43	1.453
Cause-Effect	ChatGPT	First Draft	16	19.63	1.93
	ChatGPT	Final Draft	14	24.39	2.69
	You.com	First Draft	16	17.69	3.36
	You.com	Final Draft	14	22.46	3.38
Compare-Contrast	ChatGPT	First Draft	16	17.50	2.34
	ChatGPT	Final Draft	14	21.79	1.67
	You.com	First Draft	16	19.88	1.20
	You.com	Final Draft	14	24.71	1.20

Paired Sample t-Test: The paired samples t-test was performed to determine the significance of differences between the first and final draft scores. Table 2 summarizes the results. Highly significant improvements ($p < .001$) were observed for all paragraph types. Notably, in the advantage-disadvantage paragraphs, You.com demonstrated a greater mean improvement (-5.87) compared to ChatGPT (-3.40), indicating a more pronounced impact on student performance.

Table 2. Paired Sample t-Test for Different Paragraph Types

Paragraph Type	Model	Mean Difference	t-value	df	Sig. (2-tailed)
Opinion	You.com	-3.857	-11.72	13	< .001
	ChatGPT	-3.357	-9.833	13	< .001
Advantage-Disadvantage	You.com	-5.87	-14.34	13	< .001
	ChatGPT	-3.40	-10.15	13	< .001
Cause-Effect	You.com	-4.77	-11.14	13	< .001
	ChatGPT	-4.76	-9.17	13	< .001
Compare-Contrast	You.com	-5.143	-22.26	13	< .001
	ChatGPT	-4.714	-13.93	13	< .001

Independent Samples t-Test: To examine differences between You.com and ChatGPT, an independent samples t-test was conducted. Table 3 displays the results. The results indicate no significant difference between the two AI models for most paragraph types, except for the final drafts of advantage-disadvantage and compare-contrast paragraphs. The results indicate a significant difference in the final draft scores for advantage-disadvantage paragraphs ($p = .007$), with You.com assigning higher scores compared to ChatGPT. This suggests that while both models facilitated improvements, You.com's grading was more favorable for this paragraph type.

Table 3. Independent Samples t-Test for Different Paragraph Types

Paragraph Type	Draft	Model	Mean Score	t-value	df	Sig. (2-tailed)
Opinion	First Draft	You.com	20.00	0.326	30	.651
		ChatGPT	19.63			
	Final Draft	You.com	23.50	0.464	26	.928
		ChatGPT	23.43			
Advantage-Disadvantage	First Draft	You.com	18.56	0.281	30	.781
		ChatGPT	18.31			
	Final Draft	You.com	24.43	2.948	26	.007
		ChatGPT	21.71			
Cause-Effect	First Draft	You.com	17.69	2.00	30	.055
		ChatGPT	19.63			
	Final Draft	You.com	22.46	1.6	24	.122
		ChatGPT	24.39			
Compare-Contrast	First Draft	You.com	19.88	22.34	30	< .001
		ChatGPT	17.50			
	Final Draft	You.com	24.71	5.32	26	< .001
		ChatGPT	21.79			

Correlation Analysis: Correlation analysis was performed to assess the relationship between the scores from the two AI models. Table 4 presents the results.

Table 4. Correlation Analysis for Different Paragraph Types

Paragraph Type	Draft	Pearson Correlation	Sig. (2-tailed)
Opinion	Draft 1 (You.com) – Final Draft (You.com)	.638	.014
	Final Draft (You.com) – Final Draft (ChatGPT)	.770	.001
	Draft 1 (ChatGPT) – Final Draft (ChatGPT)	.868	< .001
Advantage-Disadvantage	Draft 1 (You.com) – Final Draft (You.com)	.828	< .001
	Final Draft (You.com) – Final Draft (ChatGPT)	.751	.002
	Draft 1 (ChatGPT) – Final Draft (ChatGPT)	.884	< .001
Cause-Effect	Draft 1 (You.com) – Final Draft (You.com)	.741	.004
	Final Draft (You.com) – Final Draft (ChatGPT)	.774	.002
	Draft 1 (ChatGPT) – Final Draft (ChatGPT)	.914	< .001
Compare-Contrast	Draft 1 (You.com) – Final Draft (You.com)	.701	.005

Paragraph Type	Draft	Pearson Correlation	Sig. (2-tailed)
	Final Draft (You.com) – Final Draft (ChatGPT)	.884	< .001
	Draft 1 (ChatGPT) – Final Draft (ChatGPT)	.812	< .001

The correlation analysis shows strong positive relationships between the scores assigned by You.com and ChatGPT across all paragraph types, indicating consistent grading between the two models.

ANOVA Analysis: ANOVA was conducted to assess the statistical properties of scores assigned by the two AI models. Table 5 presents the results.

Table 5. ANOVA Results for Different Paragraph Types

Paragraph Type	Draft	Sum of Squares	df	Mean Square	F	Sig. (2-tailed)
Opinion	Draft 1	45125	1	45125	13.048	.001
	Final Draft	60036	1	60036	28.270	< .001
Advantage-Disadvantage	Draft 1	0.05	1	0.05	0.079	.078
	Final Draft	51571	1	51571	8.691	.007
Cause-Effect	Draft 1	30031	1	30031	4.001	.055
	Final Draft	24038	1	24038	2.572	.122
Compare-Contrast	Draft 1	45125	1	45125	130.48	< .001
	Final Draft	60036	1	60036	282.70	< .001

The results indicate significant differences in the scores assigned by the two AI models for most paragraph types, particularly in the final drafts. The results reveal a significant difference in final draft scores for advantage-disadvantage paragraphs ($F = 8.691$, $p = .007$), highlighting the varying effects of AI-generated feedback.

Reliability Analysis: The reliability analysis assesses the internal consistency of the scores provided by You.com and ChatGPT. The Cronbach's Alpha values for different paragraph types are presented in Table 6. The high Cronbach's Alpha values indicate high reliability and consistency in the grading provided by both AI models across all paragraph types.

Table 6. Reliability Analysis for Different Paragraph Types

Paragraph Type	Cronbach's Alpha
Opinion	.892
Advantage-Disadvantage	.868
Cause-Effect	.929
Compare-Contrast	.852

Interpretation of Quantitative Data Findings

The findings of this study demonstrate that both ChatGPT and You.com are effective in grading student writing tasks, with significant improvements observed in student performance from the first to the final drafts across all paragraph types. The process involved grading the first draft, delivering focused teaching sessions based on AI-generated feedback, and then grading the final draft. Paired sample t-tests confirmed that these improvements were statistically significant ($p < .001$), with You.com showing a slightly greater impact on scores for advantage-disadvantage paragraphs. Independent t-tests revealed a notable difference in final draft scores for advantage-disadvantage paragraphs ($p = .007$), where You.com consistently assigned higher scores than ChatGPT. Correlation analysis indicated strong positive relationships between the scores provided by both AI models, reflecting consistent grading practices. Furthermore, ANOVA results highlighted significant differences in specific paragraph types, emphasizing the varying effectiveness of each tool. Finally, the reliability analysis, with high Cronbach's Alpha values, underscores the consistency of grading provided by both AI models. Statistical tests confirmed the consistency between the grades, identified the better-performing AI model, and assessed the improvement from the first to the final draft.

Qualitative Data Analysis

Based on the qualitative data collected from the focus group interview, the following table outlines the major and minor themes that emerged regarding students' experiences with AI-generated feedback. Table 7 provides a structured overview of the insights gained from the discussions.

Table 7. Themes from Focus Group Interviews on AI-Generated Feedback

Theme Type	Theme	Description
Major Themes	Usefulness of AI Feedback	Students found the AI feedback to be beneficial in identifying areas for improvement.
	Clarity and Application of Feedback	The feedback was clear and actionable, allowing students to apply suggestions effectively.
Minor Themes	Effectiveness of Improvement Sessions	Sessions designed to address AI feedback were seen as helpful in enhancing writing skills.
	Level of Personalization and Interaction	Students noted the feedback lacked personalization and interactive elements.
	Challenges in Implementing Feedback	Some students faced difficulties in understanding and applying certain feedback.
	Influence on Confidence and Writing Skills	The feedback positively influenced students' confidence and writing abilities.

This table captures the essence of the qualitative findings, highlighting both the strengths and areas for improvement in the use of AI-generated feedback in academic writing.

Discussion

Interpretation of Quantitative Data Findings

This section presents the quantitative findings from the study, focusing on key metrics such as the effectiveness of AI-generated feedback and consistency in grading. The following analysis interprets these results in the context of existing research on AI in education, highlighting significant trends and variations. By examining these quantitative insights, we aim to provide a comprehensive understanding of the impact of AI tools on student learning outcomes, emphasizing key alignments and discrepancies with previous studies. The findings of this study align with existing research on the positive effects of AI-generated feedback on the writing skills of English as a Foreign Language (EFL) students. Consistent with Alzahrani and Alotaibi (2024), our study demonstrates substantial improvements in students' writing skills following AI intervention. Similarly, Al-Raimi (2024) and Muslimin (2024) reported significant enhancements in writing proficiency and overall performance, respectively, reinforcing the broader role of AI in language acquisition.

Our results also corroborate the findings of Nazari et al. (2021), who highlighted the benefits of AI tools like Grammarly in improving students' self-efficacy and engagement. The personalized feedback provided by AI, as noted by Hwang et al. (2023), significantly improved writing quality, particularly in terms of cohesion and grammatical accuracy. This is in line with our observation of improved scores from the first to the final drafts across all paragraph types.Furthermore, Selim (2024) and Pariyanto and Tungka (2024) confirmed that AI tools positively influence writing quality and student confidence, and that automated feedback leads to measurable improvements in writing quality. Our study supports these findings, particularly in the context of advantage-disadvantage paragraphs, where You.com showed a slightly greater impact on scores compared to ChatGPT.

Despite the positive outcomes, our study acknowledges the concerns raised by other researchers regarding the use of AI-generated feedback. Drewery et al. (2022) and William (2024) highlighted potential negative impacts on students' critical thinking and creativity, suggesting that over-reliance on AI tools might limit deep engagement with the writing process. Storey (2023) warned that AI-generated content could undermine cognitive skill development, particularly in advanced academic programs. Mun (2024) raised concerns about the reliability of AI feedback, noting that over-reliance on AI tools might hinder the development of independent writing skills. This aligns with our observation that while AI tools are effective, they should complement, rather than replace, traditional teaching methods. Akgün and Greenhow (2021) and Yan (2023) emphasized the ethical implications and potential for academic dishonesty associated with AI tools, which educators must address to ensure a supportive learning environment.

Al-Othman (2024) and Rahman et al. (2022) pointed out that AI tools might create cognitive barriers by overwhelming students and that the quality of AI feedback can vary. Our study underscores the importance of careful evaluation and integration of AI tools to enhance, rather than undermine, the development of essential cognitive skills.In conclusion, while AI-generated feedback offers substantial benefits for improving EFL students' writing skills, it is crucial to address concerns about its impact on critical thinking, creativity, and student engagement. Educators must thoughtfully integrate AI tools into their teaching practices to ensure they enhance

cognitive skill development without undermining it.

Interpretation of Qualitative Data Findings

This section presents student perspectives on AI-generated feedback, focusing on major themes like clarity and application, and minor themes such as implementation challenges and impact on writing skills. The following section interprets these findings considering existing research on AI in language learning, highlighting key alignments and discrepancies.

Usefulness of AI Feedback: Many students found AI-generated feedback particularly helpful in identifying specific areas for improvement in their writing. For instance, A**, G**, and Ö** appreciated how the AI pinpointed grammatical errors and structural issues, noting, "I found the AI feedback very helpful, especially for identifying specific grammar mistakes and problems with the structure of my writing." These observations align with the literature, which emphasizes AI's ability to offer clear, detailed, and contextually relevant feedback, particularly in language learning (Jacobsen, 2023). Research indicates that students who receive AI-generated feedback find it more readable and detailed compared to traditional instructor feedback, which can sometimes lack the same level of specificity (Escalante, 2023).

E** valued the objectivity of the feedback, stating, "I liked the AI feedback because it was fair and gave me ideas I might not have thought of." This reflects findings that AI feedback can often be more impartial and less subject to human biases, offering a balanced evaluation of strengths and weaknesses (Lin et al., 2022). G** felt motivated by the feedback, highlighting its dual focus on both strengths and areas for development: "The AI feedback was very encouraging as it highlighted the strengths in my writing as well as the areas for improvement." This mirrors research showing that AI feedback not only supports students in correcting their errors but also boosts their confidence by acknowledging their achievements (Naz, 2024).

Clarity and Application of Feedback: The integration of AI-generated feedback in English Language Teaching (ELT) has been widely recognized for its potential to enhance students' writing skills. Research indicates that AI feedback is often clearer and more specific than traditional instructor feedback, which can lead to improved writing outcomes (Escalante, 2023). This clarity was validated by students in this study, though some found it overwhelming. For instance, D** noted, "The AI feedback was a bit overwhelming at first because it covered so many areas at once," highlighting a challenge with the detailed nature of AI responses. Similarly, H** mentioned that the feedback was sometimes too general, making application difficult, but added, "The improvement sessions helped clarify some of these points and gave me a better understanding." This aligns with research that emphasizes the importance of contextual accuracy in AI-generated feedback, which should be tailored to individual learners' needs (Song, 2023).

While AI feedback can enhance writing instruction, it must be paired with guidance to ensure its effectiveness. M, for example, struggled with applying the feedback, saying, "I found the AI feedback somewhat helpful, but I struggled with implementing the suggestions. The improvement sessions were crucial for me." This aligns with

findings that educators need to support students in developing feedback literacy to interpret and apply AI-generated suggestions effectively (Tubino & Adachi, 2022). These improvement sessions were key in addressing the challenges students faced in translating AI feedback into actionable steps, reinforcing research that AI tools must be integrated into a broader pedagogical framework.

Effectiveness of Improvement Sessions: The improvement sessions were consistently praised for their role in translating AI feedback into actionable steps. Feedback-based training sessions are crucial in improving writing skills within ELT contexts. Students have specifically praised these sessions for their role in making AI feedback actionable. For instance, N** found the sessions highly valuable, stating, "The improvement sessions provided additional support and allowed me to improve my understanding of the feedback." Similarly, İ highlighted the effectiveness of the structured approach, noting, "The improvement sessions helped me focus on specific areas and gave me a clear way to revise my drafts." Studies have shown that students who engage with AI feedback during writing sessions are more likely to enhance their writing skills and self-correction abilities (Tang, 2023). However, Liu (2023) emphasizes that while automated corrective feedback can be beneficial, it must be contextualized within the learning environment to maximize its impact.

Level of Personalization and Interaction: Many students have expressed a clear preference for personal interaction during feedback sessions, emphasizing the importance of human engagement in the learning process. Y** described the AI-generated feedback as "somewhat impersonal" and appreciated the opportunity to discuss it in real-time during improvement sessions: "The improvement sessions were helpful because they let me discuss things in real-time." E** echoed this sentiment, highlighting the immediacy and clarity of in-person interactions: "The improvement sessions allowed me to ask questions and get quick feedback on how to fix the AI's suggestions." Konyrova (2024) highlights that AI's ability to analyze extensive data allows for a more nuanced understanding of each student's writing patterns, leading to tailored feedback that addresses individual needs. This personalized approach can significantly boost student engagement and motivation, as learners perceive the feedback as relevant to their specific challenges. Nonetheless, balancing AI feedback with human interaction remains a challenge, as some studies suggest that students might respond more positively to feedback from instructors compared to AI-generated comments (Drewery et al., 2022).

Challenges in Implementing Feedback: While AI-generated feedback can provide valuable insights, some students have found it challenging to implement effectively, revealing key limitations in its application. Ec** noted that the AI feedback often missed the context of her arguments, explaining, "The AI feedback was helpful, but it sometimes missed the context of my arguments." This reflects a broader issue in AI-based systems, as studies have shown that while AI excels at correcting surface-level errors, it struggles with understanding context. For instance, Gyawali (2022) explains that while AI can effectively analyze grammar and syntax, it often fails to grasp the deeper meaning of a text, leading to feedback that can feel disconnected from the writer's intended message. Similarly, As** found that the feedback, while useful as a starting point, did not always address the more subtle aspects of his writing. He shared, "The AI feedback was a good start, but it sometimes missed the details of my writing." This resonates with findings from Priyantin (2021), who notes that AI feedback, despite its efficiency, is often perceived as lacking in depth and fails to address the complexities of student writing, particularly when it

comes to style, tone, and argument development.

Influence on Confidence and Writing Skills: The positive impact of AI feedback and improvement sessions on students' confidence and writing skills is evident through student reflections. A**, for instance, highlighted how the process boosted his confidence, stating, "Overall, I feel more confident about my writing now." This aligns with research emphasizing the role of feedback in fostering self-efficacy. Herwiana (2021) notes that a supportive feedback environment, such as the improvement sessions, can create a sense of assurance in students, allowing them to feel more confident in their abilities. Similarly, Z** observed noticeable improvements in her writing because of these sessions, sharing, "The improvement sessions helped me use feedback more effectively, and I believe my writing has improved." Her experience highlights the effectiveness of formative feedback, a crucial factor in writing development. As Gul et al. (2023) emphasize, feedback from instructors can help students identify and correct their errors, ultimately improving their writing skills. Z**'s reflection is further validated by Nguyen and Nguyen (2021), who found that continuous, formative feedback leads to significant enhancements in academic writing.

In conclusion, while AI-generated feedback offers numerous benefits for EFL writing instruction, its implementation requires careful consideration. By focusing on feedback clarity and usefulness, ensuring effective integration into improvement sessions, personalizing interactions, addressing implementation challenges, and fostering student confidence, educators can effectively leverage AI tools to enhance writing skills among EFL learners. The integration of AI-generated feedback into English as a Foreign Language (EFL) writing lessons has attracted significant attention for its potential to improve writing skills and learning outcomes. However, several factors must be addressed to ensure its effectiveness, including the quality and applicability of the feedback, the effectiveness of improvement sessions, personalization and interaction, challenges in implementation, and its impact on students' confidence and writing abilities.

Conclusion

The findings from this study underscore the transformative potential of AI-generated feedback in enhancing the writing skills of EFL students. By providing clear, detailed, and contextually relevant feedback, AI tools can significantly improve students' writing proficiency, self-efficacy, and engagement. These benefits are particularly valuable in the context of English Language Teaching (ELT), where personalized and immediate feedback can address individual learning needs and promote more effective language acquisition. However, the study also highlights important considerations for the effective integration of AI in teaching. Educators must be mindful of the potential drawbacks, such as the risk of over-reliance on AI tools, which can impede the development of critical thinking and creativity. It is crucial to balance AI feedback with human interaction to ensure a holistic learning experience that fosters both cognitive and affective skills.

Practical implications of these findings include the need for structured improvement sessions that help students interpret and apply AI feedback effectively. Training educators to use AI tools proficiently is essential, as is developing students' feedback literacy. This dual approach ensures that both teachers and students can maximize

the benefits of AI-generated feedback. Additionally, AI tools can make personalized feedback more scalable and accessible, especially in large classes or remote learning environments. This scalability ensures that all students receive timely and relevant feedback, regardless of class size or learning setting. The use of AI in education raises several ethical considerations that must be addressed to ensure a supportive and fair learning environment. Potential biases in AI algorithms must be identified and mitigated to ensure fair and unbiased feedback. Protecting student data is paramount, and measures must be taken to ensure data privacy and security. Over-reliance on AI tools can hinder the development of independent writing skills and critical thinking, highlighting the need for a balanced approach. Additionally, the potential for academic dishonesty must be addressed, and strategies to promote ethical use of AI should be implemented. Transparency in the use of AI tools and their feedback generation processes is essential to build trust and understanding among students and educators.

By addressing these practical implications and ethical considerations, educators can effectively leverage AI tools to enhance writing skills among EFL learners while ensuring a balanced and ethical approach to AI integration in education. This approach will ensure that AI tools support, rather than replace, the development of essential cognitive and writing skills, ultimately enhancing students' overall learning experience. Future research should continue to explore the optimal integration of AI tools in educational settings, ensuring that they complement rather than replace traditional teaching methods.

Limitations of the Study

The study has several limitations that impact its findings. Incomplete student submissions and the small number of students reduce the comprehensiveness of the data, potentially skewing the results. Additionally, variability in how students interpreted and applied AI feedback complicates the assessment of its overall effectiveness. The small sample size of 16 participants further limits the generalizability of the findings. A notable limitation is the inconsistency between the grades provided by the two AI models, ChatGPT and You.com. It was observed that You.com generally provided higher grades compared to ChatGPT. This discrepancy could be attributed to the different algorithms utilized by each platform. Therefore, it's important to consider that a human touch in evaluation may be more beneficial than solely relying on automated grades. Without a human evaluator to validate or cross-check the AI-generated scores, concerns arise about the accuracy and reliability of the feedback. Another issue is that students primarily focused on correcting grammar and vocabulary errors, neglecting higher-order thinking skills such as cohesion, argumentation, and content development, which limited the depth of their revisions.

Individual differences in learning styles also affected how students benefited from the feedback, with some finding it more effective than others. The short-term nature of the study restricted insights into the long-term effects of AI feedback, and the technological limitations of the AI tools meant they struggled to address more complex writing issues. Additionally, there is a risk of over-reliance on AI feedback, which could hinder the development of self-editing and peer-review skills. Addressing these limitations and incorporating a human evaluator in future studies will enhance understanding of the full potential and effectiveness of AI feedback in improving writing skills.

Suggestions for Further Research

To expand on the findings of this study, several areas for further research are suggested. Increasing the sample size would enhance the generalizability of the results and provide a broader understanding of the effectiveness of AI-generated feedback and improvement sessions. Longitudinal studies could assess the long-term impact of these interventions on students' writing skills and academic performance.

Future research could also explore which specific components of AI feedback are most effective, helping to refine AI tools for better addressing student needs. Examining the influence of individual learning styles on feedback reception could lead to more personalized interventions. Additionally, integrating AI feedback with peer reviews and instructor feedback could create a more comprehensive writing development program.

Addressing challenges in applying AI feedback is another key area for study, as well as investigating the risk of over-reliance on AI, to ensure students continue developing self-editing and critical thinking skills. These research directions would further improve the effectiveness of AI-based writing instruction. In conclusion, AI-generated feedback presents numerous opportunities to improve EFL students' writing skills by offering clear, immediate, and personalized feedback. However, concerns about its impact on critical thinking, creativity, and student engagement must be addressed. By adopting a balanced approach that combines AI feedback with human guidance, educators can harness the advantages of AI tools while mitigating potential drawbacks. This approach will ensure that AI tools support, rather than replace, the development of essential cognitive and writing skills, ultimately enhancing students' overall learning experience. Future research should continue to explore the optimal integration of AI tools in educational settings, ensuring that they complement rather than replace traditional teaching methods.

References

Adawiyah, B., & Sumarni, B. (2021). The EFL lecturers' feedback on pre-service teachers' writing performance in online instruction. Borneo Educational Journal (Borju), 3(2), 33-47. https://doi.org/10.24903/bej.v3i2.781

Akavova, A. (2023). Adaptive learning and artificial intelligence in the educational space. E3S Web of Conferences, 451, 06011. https://doi.org/10.1051/e3sconf/202345106011

Akbarani, R. (2024). Use of artificial intelligence in English language teaching. International Journal of English Learning and Applied Linguistics (IJELAL), 4(1), 14-23. https://doi.org/10.21111/ijelal.v4i1.10756

Akgün, S., & Greenhow, C. (2021). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. AI and Ethics, 2(3), 431-440. https://doi.org/10.1007/s43681-021-00096-7

Alharbi, W. (2023). AI in the foreign language classroom: A pedagogical overview of automated writing assistance tools. Education Research International, 2023, 1-15. https://doi.org/10.1155/2023/4253331

Al-Othman, A. (2024). EFL students develop cognitive and metacognitive self-regulated writing strategies using automated feedback: A case study. Theory and Practice in Language Studies, 14(5), 1525-1536. https://doi.org/10.17507/tpls.1405.26

Al-Raimi, M. (2024). Utilizing artificial intelligence tools for improving writing skills: Exploring Omani EFL learners' perspectives. Forum for Linguistic Studies, 6(2), 1177. https://doi.org/10.59400/fls.v6i2.1177

Alzahrani, F., & Alotaibi, H. (2024). The impact of artificial intelligence on enhancing EFL writing skills among high school students. Journal of Educational and Human Sciences, 34, 226-240. https://doi.org/10.33193/JEAHS.34.2024.482

Antara, I. (2024). Utilizing artificial intelligence to evaluate students' competency in writing simple past tense. Stilistika Jurnal Pendidikan Bahasa Dan Seni, 12(2), 126-140. https://doi.org/10.59672/stilistika.v12i2.3691

Athanassopoulos, S. (2023). The use of ChatGPT as a learning tool to improve foreign language writing in a multilingual and multicultural classroom. Advances in Mobile Learning Educational Research, 3(2), 818-824. https://doi.org/10.25082/amler.2023.02.009

Baillie, L. (2019). Exchanging focus groups for individual interviews when collecting qualitative data. Nurse Researcher, 27(2), 15-20. https://doi.org/10.7748/nr.2019.e1633

Betal, A. (2023). Enhancing second language acquisition through artificial intelligence (AI): Current insights and future directions. Journal for Research Scholars and Professionals of English Language Teaching, 7(39). https://doi.org/10.54850/jrspelt.7.39.003

Bijami, M., Kashef, S., & Nejad, M. (2013). Peer feedback in learning English writing: Advantages and disadvantages. Journal of Studies in Education, 3(4), 91. https://doi.org/10.5296/jse.v3i4.4314

Bozkurt, A., Karadeniz, A., Bañeres, D., & Rodríguez, M. (2021). Artificial intelligence and reflections from the educational landscape: A review of AI studies in half a century. Sustainability, 13(2), 800. https://doi.org/10.3390/su13020800

Brooks, G. (2012). Assessment and academic writing: A look at the use of rubrics in the second language writing classroom. Kwansei Gakuin University Humanities Review, 17, 227-240. https://core.ac.uk/download/pdf/143638458.pdf

Cantos, K. (2023). Artificial intelligence in language teaching and learning. Ciencia Latina Revista Científica Multidisciplinar, 7(4), 5629-5638. https://doi.org/10.37811/cl_rcm.v7i4.7368

Cardon, P. (2023). The challenges and opportunities of AI-assisted writing: Developing AI literacy for the AI age. Business and Professional Communication Quarterly, 86(3), 257-295. https://doi.org/10.1177/23294906231176517

Carlson, M. (2023). Utilizing OpenAI's GPT-4 for written feedback. TESOL Journal, 15(2). https://doi.org/10.1002/tesj.759

Chen, H. (2024). Educational applications of ChatGPT: Ethical challenges and countermeasures. English Language Teaching and Linguistics Studies, 6(3), 100. https://doi.org/10.22158/eltls.v6n3p100

Coursey, C., & Dandashly, N. (2015). Digital literacies and generational micro-cultures: Email feedback in Lebanon. English Language Teaching, 8(11), 216. https://doi.org/10.5539/elt.v8n11p216

Dai, Y. (2023). ChatGPT and its impact on research supervision: Insights from Australian postgraduate research students. Australasian Journal of Educational Technology, 39(4), 74-88. https://doi.org/10.14742/ajet.8843

Demianenko, O. (2024). Impact of artificial intelligence in foreign language teaching: Forming and developing students' skills. Věda a Perspektivy, 4(35), 191-199. https://doi.org/10.52058/2695-1592-2024-4(35)-

191-199

Drewery, D., Woodside, J., & Eppel, K. (2022). Artificial intelligence and résumé critique experiences. Canadian Journal of Career Development, 21(2), 28-39. https://doi.org/10.53379/cjcd.2022.338

Escalante, J. (2023). AI-generated feedback on writing: Insights into efficacy and ENL student preference. International Journal of Educational Technology in Higher Education, 20(1). https://doi.org/10.1186/s41239-023-00425-2

Fareed, U. (2021). Peer feedback on writing skills: Perceptions of Pakistani ESL postgraduate students. Pakistan Journal of Educational Research, 4(1). https://doi.org/10.52337/pjer.v4i1.169

Firat, M. (2023). What ChatGPT means for universities: Perceptions of scholars and students. Journal of Applied Learning and Teaching, 6(1), 1-22. https://doi.org/10.37074/jalt.2023.6.1.22

Fitria, T. (2023). Artificial intelligence (AI) technology in OpenAI ChatGPT application: A review of ChatGPT in writing English essays. ELT Forum Journal of English Language Teaching, 12(1), 44-58. https://doi.org/10.15294/elt.v12i1.64069

Fu, S., Gu, H., & Yang, B. (2020). The affordances of AI-enabled automatic scoring applications on learners' continuous learning intention: An empirical study in China. British Journal of Educational Technology, 51(5), 1674-1692. https://doi.org/10.1111/bjet.12995

Gavora, P. (2015). An analysis of interaction patterns in the focus group interview. Acta Technologica Dubnicae, 5(3), 11-23. https://doi.org/10.1515/atd-2015-0068

Gonzalez-Torres, P., & Sarango, C. (2023). Effectiveness of teacher and peer feedback in EFL writing: A case of high school students. International Journal of Learning Teaching and Educational Research, 22(4), 73-86. https://doi.org/10.26803/ijlter.22.4.5

Gul, A., Saleem, Z., & Zahra, N. (2023). Teachers' oral feedback practices in promoting academic writing skills of ESL learners. Journal of Social Sciences Review, 3(1), 302-309. https://doi.org/10.54183/jssr.v3i1.157

Gyawali, Y. (2022). Artificial intelligence in English language teaching: Navigating the future with emerging perspectives. Journal of Language and Linguistics in Society, 26, 21-27. https://doi.org/10.55529/jlls.26.21.27

Hadi, M., Alldred, D., Closs, S., & Briggs, M. (2012). Mixed-methods research in pharmacy practice: Basics and beyond (Part 1). International Journal of Pharmacy Practice, 21(5), 341-345. https://doi.org/10.1111/ijpp.12010

Hasan, A. (2022). Effect of rubric-based feedback on the writing skills of high school graders. Journal of Innovation in Educational and Cultural Research, 3(1), 49-58. https://doi.org/10.46843/jiecr.v3i1.52

Herwiana, S. (2021). Strengths and weaknesses of collaborative writing and peer feedback in an EFL intensive reading and writing coursework. Pioneer Journal of Language and Literature, 13(1), 105. https://doi.org/10.36841/pioneer.v13i1.944

Hwang, W., Nurtantyana, R., Purba, S., Hariyanti, U., Indrihapsari, Y., & Surjono, H. (2023). AI and recognition technologies to facilitate English as a foreign language writing for supporting personalization and contextualization in authentic contexts. Journal of Educational Computing Research, 61(5), 1008-1035. https://doi.org/10.1177/07356331221137253

Hz, B. I. R., Laiya, R. E., Sarumaha, M. S., & Supiyandi. (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

Jacobsen, L. (2023). The promises and pitfalls of ChatGPT as a feedback provider in higher education: An exploratory study of prompt engineering and the quality of AI-driven feedback. https://doi.org/10.31219/osf.io/cr257

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. Educational Research Review, 2(2), 130-144. https://www.uidaho.edu/-/media/UIdaho-Responsive/Files/provost/IR/Assessment/Assessment-Resources/3E/scoring_rubrics_validity.pdf?la=en&hash=485A51B8A5A2FBC57C28BB5732500F516210A4D3

Kamruzzaman, M. (2023). AI- and IoT-assisted sustainable education systems during pandemics, such as COVID-19, for smart cities. Sustainability, 15(10), 8354. https://doi.org/10.3390/su15108354

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., … & Nerdel, C. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. https://doi.org/10.35542/osf.io/5er8f

Khan, W. (2023). Examining the transformative role of artificial intelligence in language skill enhancement: A case study of BS English students in Okara, Pakistan. ABBDM, 3(1), 190-196. https://doi.org/10.62019/abbdm.v3i1.88

Koka, N. A., Khan, A. S., & Jan, N. (2023). Exploring the attitudes and perceptions of foreign language lecturers on artificial intelligence-driven writing evaluation and feedback tools for improving undergraduate writing skills. Journal of Southwest Jiaotong University, 58(5). https://doi.org/10.35741/issn.0258-2724.58.5.6

Konyrova, L. (2024). The evolution of language learning: Exploring AI's impact on teaching English as a second language. Eurasian Science Review, 2(2), 133-138. https://doi.org/10.63034/esr-42

Konyrova, L. (2024). The evolution of language learning: Exploring AI's impact on teaching English as a second language. Eurasian Science Review, 2(2), 133–138. https://doi.org/10.63034/esr-42

Li, Z. (2024). The value, benefits, and concerns of generative AI-powered assistance in writing. https://doi.org/10.1145/3613904.3642625

Lin, J., Sha, L., Gašević, D., & Chen, G. (2022). Establishing trustworthy artificial intelligence in automated feedback. https://doi.org/10.35542/osf.io/5efxn

Lin, Z., Song, X., Guo, J., & Feng, W. (2021). Peer feedback in translation training: A quasi-experiment in an advanced Chinese–English translation course. Frontiers in Psychology, 12. https://doi.org/10.3389/fpsyg.2021.631898

Liu, B. (2023). Chinese university students' attitudes and perceptions in learning English using ChatGPT. International Journal of Education and Humanities, 3(2), 132-140. https://doi.org/10.58557/ijeh.v3i2.145

Liu, Y. (2023). A comparison of automated corrective feedback and traditional corrective feedback: A review study. The Educational Review USA, 7(9), 1365-1368. https://doi.org/10.26855/er.2023.09.024

Mahmud, F. (2023). Investigating EFL students' writing skills through artificial intelligence: Wordtune application as a tool. Journal of Language Teaching and Research, 14(5), 1395–1404. https://doi.org/10.17507/jltr.1405.28

Marzuki, W., Widiati, U., Rusdin, D., Darwin, & Indrawati, I. (2023). The impact of AI writing tools on the content and organization of students' writing: EFL teachers' perspective. Cogent Education, 10(2), Article 2236469. https://doi.org/10.1080/2331186X.2023.2236469

McLafferty, I. (2004). Focus group interviews as a data collecting strategy. Journal of Advanced Nursing, 48(2), 187-194. https://doi.org/10.1111/j.1365-2648.2004.03186.x

Michel-Villarreal, R. (2023). Challenges and opportunities of generative AI for higher education as explained by ChatGPT. Education Sciences, 13(9), 856. https://doi.org/10.3390/educsci13090856

Mun, C. (2024). EFL learners' English writing feedback and their perception of using ChatGPT. STEM Journal, 25(2), 26-39. https://doi.org/10.16875/stem.2024.25.2.26

Muslimin, A. (2024). Evaluating CAMI AI across SMART stages: Students' achievement and perceptions in EFL writing instruction. Online Learning, 28(2). https://doi.org/10.24059/olj.v28i2.4246

Nagi, F., Salih, R., Alzubaidi, M., Shah, H., Alam, T., Shah, Z., … & Househ, M. (2023). Applications of artificial intelligence (AI) in medical education: A scoping review. https://doi.org/10.3233/shti230581

Naz, I. (2024). Exploring the feasibility and efficacy of ChatGPT-3 for personalized feedback in teaching. The Electronic Journal of E-Learning, 22(2), 98-111. https://doi.org/10.34190/ejel.22.2.3345

Nazaretsky, T., Ariely, M., Cukurova, M., & Alexandron, G. (2022). Teachers' trust in AI-powered educational technology and a professional development program to improve it. British Journal of Educational Technology, 53(4), 914-931. https://doi.org/10.1111/bjet.13232

Nazari, N., Shabbir, M., & Setiawan, R. (2021). Application of artificial intelligence-powered digital writing assistant in higher education: Randomized controlled trial. Heliyon, 7(5), e07014. https://doi.org/10.1016/j.heliyon.2021.e07014

Ng, D., Leung, J., Su, J., Ng, R., & Chu, S. (2023). Teachers' AI digital competencies and twenty-first century skills in the post-pandemic world. Educational Technology Research and Development, 71(1), 137-161. https://doi.org/10.1007/s11423-023-10203-6

Nguyen, T., & Nguyen, H. (2021). Teachers' beliefs about the impact of students' peer correction at the post-writing stage on letter writing. International Journal of Science and Management Studies (IJSMS), 4(3), 120-133. https://doi.org/10.51386/25815946/ijsms-v4i3p112

Nicolas, H. (2020). Student determinants in the licensure examination for agriculturists of a state college in the Philippines. International Journal of Educational Sciences, 28(1-3). https://doi.org/10.31901/24566322.2020/28.1-3.1117

Nurhasanah, N., Samsudin, A., Linuwih, S., & Sutrisno, A. (2022). Is Sipoert an innovative learning? Implementation and identification of students' conception on simple harmonic motion. Indonesian Journal on Learning and Advanced Education (Ijolae), 5(1), 78-91. https://doi.org/10.23917/ijolae.v5i1.19567

Owan, V. (2023). Exploring the potential of artificial intelligence tools in educational measurement and assessment. Eurasia Journal of Mathematics Science and Technology Education, 19(8), em2307. https://doi.org/10.29333/ejmste/13428

Pack, A., & Maloney, J. (2023). Using generative artificial intelligence for language education research: Insights from using OpenAI's ChatGPT. TESOL Quarterly, 57(4), 1571-1582. https://doi.org/10.1002/tesq.3253

Pang, T. (2024). Artificial intelligence use in feedback: A qualitative analysis. Journal of University Teaching and

Learning Practice, 21(06). https://doi.org/10.53761/40wmcj98

Pariyanto, & Tungka, Novalita. (2024). Enhancing writing skills of EFL learners through automated feedback: An empirical investigation. Jurnal Ilmiah Spectral, 10, 028-042. https://doi.org/10.47255/gymtcf78

Parra G., L., & Calero S., X. (2019). Automated writing evaluation tools in the improvement of the writing skill. International Journal of Instruction, 12(2), 209–226. https://doi.org/10.29333/iji.2019.12214a

Patiño, A. (2024). The improvement of oral communicative competence in English through artificial intelligence. Latam Revista Latinoamericana de Ciencias Sociales y Humanidades, 5(1). https://doi.org/10.56712/latam.v5i1.1850

Priyantin, S. (2021). The limitations of AI feedback in writing instruction: A critical analysis. Journal of Educational Technology & Society, 24(3), 45-56.

Rabiee, F. (2004). Focus-group interview and data analysis. Proceedings of the Nutrition Society, 63(4), 655-660. https://doi.org/10.1079/pns2004399

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

Rakya, Z. (2023). Exploring the impact of artificial intelligence (AI) on learner-instructor interaction in online learning (literature review). International Journal of Emerging Multidisciplinaries Computer Science & Artificial Intelligence, 2(1). https://doi.org/10.54938/ijemdcsai.2023.02.1.236

Razı, S. (2015). Development of a rubric to assess academic writing incorporating plagiarism detectors. SAGE Open, 5(2), 1–13. https://doi.org/10.1177/2158244015590162

Sallam, M., Salim, N., Barakat, M., & Al-Tammemi, A. (2023). ChatGPT applications in medical, dental, pharmacy, and public health education: A descriptive study highlighting the advantages and limitations. Narra J, 3(1), e103. https://doi.org/10.52225/narra.v3i1.103

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

Semerikov, S. (2021). AI-assisted language education: Critical review. Освітній Вимір, 4, 1-7. https://doi.org/10.31812/ed.623

Seo, K., Tang, J., Roll, I., Fels, S., & Yoon, D. (2021). The impact of artificial intelligence on learner–instructor interaction in online learning. International Journal of Educational Technology in Higher Education, 18(1). https://doi.org/10.1186/s41239-021-00292-9

Smith, K., Burns, C., O'Neill, C., Duggan, J., Winkelman, N., Wilkie, M., … & Coughlan, E. (2022). How to coach: A review of theoretical approaches for the development of a novel coach education framework. International Journal of Sports Science & Coaching, 18(2), 594-608. https://doi.org/10.1177/17479541221136222

Soelistiono, S. (2023). Educational technology innovation: AI-integrated learning system design in AILS-based education. Influence International Journal of Science Review, 5(2), 470-480. https://doi.org/10.54783/influencejournal.v5i2.175

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

Storey, V. A. (2023). AI technology and academic writing. International Journal of Adult Education and Technology, 14(1), 1-15. https://doi.org/10.4018/ijaet.325795

Strzelecki, A. (2023). Students' acceptance of ChatGPT in higher education: An extended unified theory of acceptance and use of technology. Innovative Higher Education, 49(2), 223-245. https://doi.org/10.1007/s10755-023-09686-1

Suraworachet, W., Zhou, Q., & Cukurova, M. (2023). Impact of combining human and analytics feedback on students' engagement with, and performance in, reflective writing tasks. International Journal of Educational Technology in Higher Education, 20(1). https://doi.org/10.1186/s41239-022-00368-0

Tang, W. (2023). Unlocking second language students' potential: ChatGPT's pivotal role in English for academic purposes writing success. Proceedings of the 4th International Conference on Language, Literature, and Education (pp. 694-706). https://doi.org/10.2991/978-2-38476-126-5_79

Teebagy, S., Colwell, L., Wood, E., Yaghy, A., & Faustina, M. (2023). Improved performance of ChatGPT-4 on the OKAP exam: A comparative study with ChatGPT-3.5. https://doi.org/10.1101/2023.04.03.23287957

Trisnadewi, K. (2021). The use of peer correction to improve student writing skills at STMIK STIKOM Indonesia. Yavana Bhasha Journal of English Language Education, 4(1), 73. https://doi.org/10.25078/yb.v4i1.2213

Tubino, L., & Adachi, C. (2022). Developing feedback literacy capabilities through an AI automated feedback tool. Ascilite Publications, e22039. https://doi.org/10.14742/apubs.2022.39

Vattøy, K. (2020). Feedback for self-regulated learning: The role of human interaction. Assessment & Evaluation in Higher Education, 45(5), 685-698. https://doi.org/10.1080/02602938.2020.1717631

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

Wang, D., & Han, H. (2020). Applying learning analytics dashboards based on process-oriented feedback to improve students' learning effectiveness. Journal of Computer Assisted Learning, 37(2), 487-499. https://doi.org/10.1111/jcal.12502

Ward-Horner, J. & Sturmey, P. "The Effects of Behavioral Skills Training on Instructor and Learner Behavior Across Responses and Skill Sets." *Research in Developmental Disabilities*, vol. 33, no. 2, pp. 458-466, 2012. doi: 10.1016/j.ridd.2011.10.016.

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

Yan, D. (2023). How ChatGPT's automatic text generation impacts learners in a L2 writing practicum: An exploratory investigation. https://doi.org/10.35542/osf.io/s4nfz

Yunina, O. (2023). Artificial intelligence tools in foreign language teaching in higher education institutions. The Modern Higher Education Review, (8). https://doi.org/10.28925/2617-5266.2023.85

Zaker, A. (2016). Literature and creativity in an ELT context. Asian TEFL Journal of Language Teaching and Applied Linguistics, 1(2). https://doi.org/10.21462/asiantefl.v1i2.20

Zarei, M. (2024). ChatGPT, a friend or a foe in medical education: A review of strengths, challenges, and opportunities. Shiraz E-Medical Journal, 25(6). https://doi.org/10.5812/semj-145840

Zhao, Y. (2022). The impact of AI writing tools on the content and organization of students' writing: EFL

teachers' perspective. Information & Communications Technology in Education. https://www.tandfonline.com/doi/pdf/10.1080/2331186X.2023.2236469

Zou, B. (2023). An investigation into artificial intelligence speech evaluation programs with automatic feedback for developing EFL learners' speaking skills. SAGE Open, 13(3). https://doi.org/10.1177/21582440231193818Stoneman, R. (2008). *Alexander the Great: A life in legend*. Yale University Press.

Author Information

Mustafa Özdere

https://orcid.org/0000-0001-7535-9024

Niğde Ömer Halisdemir University

School of Foreign Languages

Turkey

Contact e-mail: *mozdere@gmail.com*