Potentials and Implications of ChatGPT for ESL Writing Instruction

Dr. Karim Ibrahim and Dr. Robert Kirkpatrick
Gulf University for Science and Technology

Abstract

The release of ChatGPT has marked the dawn of a new information revolution that will transform how people communicate and make meaning. However, to date, little is known about the implications of ChatGPT for L2 composition instruction. To address this gap, the present study uses a systematic review design to synthesize available research on the educational potentials of ChatGPT as an instructional assistant, outline the implications of these potentials for L2 writing instruction, and discuss their practical applications. The findings, based on a meta-analysis of 42 research articles, demonstrate that ChatGPT can enhance L2 writing instruction by boosting learners' motivation, automating instructional tasks, and offering instantaneous, personalized feedback to learners. These findings have important implications for harnessing the instructional potential of generative AI in L2 writing classes.

Keywords: artificial intelligence, artificial neural network, L2 writing, AI applications, chatbots

Introduction

The public release of ChatGPT in November 2022 marked the dawn of a new information literacy revolution in which generative AI may come to dominate the generation, classification, and dissemination of knowledge in information and communication technologies (ICT). The impressive abilities of ChatGPT to interact verbally with users, perform complex natural language processing tasks, and generate coherent, intelligible, and humanlike texts (Lund & Wang, 2023; Radford et al., 2018; Radford et al., 2019; Solaiman et al., 2019) attracted 100 million users within two months of its release, making it the fastest-growing Internet app ever (Hu, 2023). These revolutionary abilities have transformational effects on education as they facilitate the automation of various mundane tasks that are integral to education (Ibrahim, 2023).

A sub-field of education that has been deeply influenced by the rise of generative AI is second language (L2) writing/composition instruction, perhaps because of the wide overlap between the cognitive skills taught in writing classes and generative AI capabilities. The unprecedented capabilities of ChatGPT to process and generate texts hold both great potentials and challenges to L2 writing instruction, but most of the existing research on the implications of ChatGPT for L2 writing has emphasized ethical challenges to traditional L2 writing assessments (Cotton et al., 2024; Francke & Alexander, 2019; Gao et al., 2023; Haque et al., 2022; King & ChatGPT, 2023; Khalil & Er, 2023; Susnjak, 2022; Yeadon et al., 2022). As a result, little is known about the potential of ChatGPT to support L2 writing instruction and development.

To address this gap in the literature and shed light on the practical implications of this promising new technology, the present study will review available research on the potentials and implications of ChatGPT for L2 writing instruction. It is hoped that this review can underline the instructional benefits of ChatGPT for L2 writing, provide guidelines for integrating ChatGPT in L2 writing classes, and outline essential areas and questions for future research in this new field. These findings can help L2 writing instructors harness the instructional potentials of ChatGPT and integrate AI literacies in their classes. The paper uses a systematic review approach to summarize and synthesize research findings on the applications of ChatGPT for L2 writing instruction. Research articles were gathered through an extensive search on Web of Science, EBSCOHost, Google Scholar, and arXiv databases using the following keywords: "ChatGPT and L2 writing," "ChatGPT and writing instruction," "ChatGPT and automatic feedback," "generative AI and instruction automation," and "ChatGPT and academic writing motivation." The search results generated 58 relevant research articles, but 16 articles were excluded over methodological issues; thus, 42 articles were finally included in the literature review.

Literature Review

ChatGPT

The generative pre-trained transformer model (GPT) is an auto-regressive large language model (LLM) equipped with a chatbot, ChatGPT, that allows Internet users to interact verbally with the model via a virtual agent called "Assistant" (Radford et al., 2018). GPT is a machine-learning neural network trained to model natural language systems (MacNeil et al., 2022) and perform natural language processing (NLP) tasks, including writing articles, composing music, answering questions, and generating computer code (OpenAI,

n.d.-b). GPT was trained using an innovative machine-learning approach that combines unsupervised pre-training (Radford et al., 2018) and reinforcement learning from human feedback approaches (Christiano et al., 2017) on a massive corpus of one billion words of web texts (Radford et al., 2019). This unique training approach allows the model to recognize patterns in the training data and develop different NLP skills without task-specific training (Brown et al., 2020). This approach and rich training data allowed GPT to grow into a 175-billion-parameter model that can generate fluent, coherent, and humanlike texts on various topics (Ibrahim, 2023). The latest version of the model, GPT-4, is multimodal and can interact with users verbally and interpret audiovisual content (OpenAI, n.d.-b).

ChatGPT and L2 Education

As an intelligent chatbot, ChatGPT can leverage its natural language processing capabilities, deep learning dynamics, and extensive training to support L2 education (Huang & Li, 2023; Hong, 2023; Kohnke et al., 2023; Rahman & Watanobe, 2023). ChatGPT can be used as an intelligent L2 tutor to offer L2 learners extended L2 practice opportunities (Rahman & Watanobe, 2023) and corrective feedback on their L2 output (Rudolph et al., 2023). ChatGPT can use its machine-learning capabilities to analyze L2 learners' performance and personalize their L2 learning trajectories to match their individual learning needs (Huang & Li, 2023). These potentials would boost autonomous L2 learning and learners' engagement (Qadir, 2022). ChatGPT could also offer L2 educators rich L2 learning resources that support autonomous and collaborative L2 instruction and learning, including practice activities, grammar worksheets, and vocabulary illustration (Rahman & Watanobe, 2023).

A few exploratory studies have demonstrated the potential of ChatGPT to support L2 education (Rakhmonov & Kurbonova, 2023; Shaikh et al., 2023; Young et al., 2023). Rakhmonov and Kurbonova (2023) investigated the L2 learning potentials of ChatGPT with 20 L2 learners and instructors. Analysis of survey and interview data revealed that ChatGPT could generate personalized L2 learning content that matches learners' needs. Similarly, Shaikh et al. (2023) explored the usability of ChatGPT in L2 learning contexts with 10 Norwegian ESL learners. The researchers engaged the participants in ChatGPT-mediated ESL learning activities and collected their responses via a questionnaire. Data analysis revealed that the participants perceived ChatGPT to be a valuable and accessible resource for L2 practice. In a similar vein, Young et al. (2023) investigated the appropriateness of ChatGPT output for ESL education. Using a readability matrix analysis, the researchers analyzed the appropriateness of English dialogues generated by ChatGPT for ESL learners. Data analysis revealed that ChatGPT-generated content is deemed appropriate for the learning needs of ESL learners.

ChatGPT as an L2 Writing Instructional Assistant

ChatGPT can support L2 writing instruction and enhance teachers' efficiency by automating instructional planning, materials design, assessment, and course management (Qadir, 2022; Rudolph et al., 2023). Firstly, AI can help writing instructors reduce their administrative and planning workload so they can spend more time designing L2 writing opportunities by automating instructional design tasks (Baidoo-anu & Ansah, 2023). They could use ChatGPT to design course outlines and plans (Rahman & Watanobe, 2023) tailored to specific L2 writing objectives and goals (Baskara & Mukarto, 2023); generate specific, well-articulated, and measurable intended learning outcomes based on Bloom's taxonomy for specific L2 writing sub-skills (Sridhar et al., 2023); and create custom-designed lesson plans that follow specific pedagogical

approaches or methods (Hong, 2023). For instance, in an exploratory study, Sridhar et al. (2023) examined the potential of GPT-4 to support instructional design by authoring high-quality learning outcomes (LOs) for a university course on AI. Using Bloom's taxonomy as a frame of reference, the researchers prompted GPT-4 by providing specific guidelines for the designed LOs, including a course description, instructional design guidelines, design specifications for the LOs, and examples of well-designed LOs. Analysis of the generated LOs by human reviewers and AI classifiers demonstrated that they were sensible, appropriate for the intended cognitive processes, and consistent with Bloom's taxonomy. This study demonstrates that ChatGPT can assist writing instructors by automating instructional design processes.

Second, L2 writing instructors can use ChatGPT to generate rich instructional resources tailored to specific learning objectives, learning preferences, and pedagogical approaches (Baskara & Mukarto, 2023; Huang & Li, 2023; Hong, 2023; Kohnke et al., 2023; Qadir, 2022; Rahman & Watanobe, 2023). For instance, the model can support writing instruction by generating writing prompts (Baskara & Mukarto, 2023), writing practice activities, and learning tips and suggestions (Yan, 2023). It can also support research writing more specifically by locating and summarizing previous research on a topic (Huang & Li, 2023), assisting students in research design and planning, and offering suggestions for new research directions (Rahman & Watanobe, 2023). Instructors can also use the deep learning capabilities of ChatGPT to develop adaptive learning activities that adjust their pedagogical approaches and complexity levels to students' L2 proficiency levels, learning needs, and preferences (Baidoo-anu & Ansah, 2023). That is, ChatGPT would analyze students' performance data, identify areas of weakness, monitor their progress, and adjust learning design as needed to address the specific learning needs of each student (Rahman & Watanobe, 2023). Also, instructors can use ChatGPT to instill innovative instructional practices, such as using flipped classroom approaches (Lage et al., 2000) by automating L2 writing instruction to be delivered out of class and using class time for intensive L2 writing practice (Rudolph et al., 2023). This way, students could learn foundational constructs outside class and maximize L2 writing practice in class, which should boost teaching and learning effectiveness (Hong, 2023). In addition, ChatGPT can support experiential learning by generating different scenarios and problem-solving activities that could foster collaborative L2 interaction between learners (Rudolph et al., 2023).

A few empirical studies have explored the potential of ChatGPT to automate instruction design (MacNeil et al., 2022; Pardos & Bhandari, 2023). MacNeil et al. (2022) reported on a case study where the Codex and GPT-3 LLMs were used to generate practice programming assignments and automatic explanations (explanation, definitions, hints, and feedback) on learners' code. The researchers were able to use LLMs to automate the processes of designing personalized assignments and offering rich explanations to students, which is typically time-consuming. In another study, Pardos and Bhandari (2023) examined the potential of ChatGPT to offer personalized learning hints to students in an algebra course. The researchers used an online GPT-powered tutoring system (GPT-3.5) to generate automatic learning hints and explored their implications for learning. Using a pre-test–post-test design, they compared the efficacy of ChatGPT-generated and human-generated hints from tutors with 77 college algebra students. The results demonstrated that both groups experienced positive learning gains. Even though these studies did not investigate L2 contexts, their findings suggest that ChatGPT can offer custom-designed and personalized instructional support to L2 writers.

Third, ChatGPT can automate the procedures of generating and implementing assessment and feedback processes (Baidoo-anu & Ansah, 2023; Hong, 2023; Rahman & Watanobe, 2023). ChatGPT can generate assessment tools, including writing prompts, graded reading passages, open-ended or multiple-choice questions, and grading rubrics that align with specific learning objectives (Kohnke et al., 2023; Qadir, 2022). Also, ChatGPT can function as an automated L2 assessment platform that can grade students' work automatically and provide students with instant feedback on their performance (Barrot, 2023), ensuring timely feedback, which is crucial for learners' development and can be challenging for instructors (e.g., Wang et al., 2023). This is especially beneficial in L2 writing practice (Rahman & Watanobe, 2023), where previous research suggests that using Grammarly to offer L2 learners feedback on their writing was an effective intervention for engaging students in writing practice (Koltovskaia, 2020). Automating assessment would also integrate assessment feedback into the learning process. For example, automated essay scoring can be used to identify patterns in students' responses and suggest revisions to students. Intelligent feedback could also support students' autonomy and improve their writing skills through error analysis: recognizing errors, identifying correct patterns, and reformulating their writing (Rudolph et al., 2023).

Several exploratory studies have demonstrated the potential of ChatGPT to automate assessment and grading (Altamimi, 2023; Jia et al., 2022; Kortemeyer, 2023; Kınık & Çetin, 2024). Altamimi (2023) evaluated the performance of ChatGPT (both GPT-3.5 and 4.0 versions) as automatic essay grading and feedback systems across several academic domains. Specifically, the study examined the accuracy, efficacy, and reliability of ChatGPT grading in comparison to human raters. The results indicated that ChatGPT offered an efficient automatic grading platform and that grading accuracy was higher for GPT-4 compared to GPT-3.5. Similarly, Jia et al. (2022) examined the feasibility of a BART-based pre-trained LLM to generate instant feedback on students' report projects. The researcher used a dataset of 484 project reports to train and test the performance of the BART-based feedback system in comparison to human experts according to specific criteria (e.g., readability, helpfulness, and specificity). Analysis of the data revealed that BART could generate feedback on par with expert human feedback. In another study, Kortemeyer (2023) examined the performance of GPT-4 on automatic grading of short-answer questions. Using standardized automatic grading evaluation benchmarks, the researcher evaluated the precision and recall of GPT-4 as a general-purpose automatic grader. The analysis revealed that the performance of ChatGPT-4 was comparable to that of earlier automatic grading models (custom-designed automatic graders) without the need for reference answers, but it underperformed compared to deep learning models that received task-specific pre-training for automatic grading, a limitation that can be rectified with additional training. In another study, Kınık and Çetin (2024) compared the scoring of 20 descriptive essays by student teachers of English and ChatGPT-3.5. They found the human raters were generally more accurate, but they also suggested that ChatGPT had great potential in this area.

Fourth, ChatGPT technology can assist instructors by automating learning management systems and reducing manual administrative interventions (Huang & Li, 2023). It could automatically post activities, grade assignments, monitor students' performance, supervise teacher assistants/tutors, identify learning challenges, and take remedial action (Huang & Li, 2023). Automation of mundane administrative tasks would save instructor's time and help them focus on improving their teaching effectiveness. At least one study has examined the potential of ChatGPT to automate administrative learning management tasks.

Hirunyasiri et al. (2023) examined the potential of GPT-4 to automate the evaluation of tutors' performance and offer them feedback. Using a framework of effective tutoring feedback criteria (e.g., timely delivery, process-focused praise, sincerity, etc.), the researchers investigated the ability of GPT-4 to understand evaluation criteria, analyze tutor-student interactions, accurately evaluate tutors' comments, and offer them feedback on their effectiveness in delivering praise to students. The researchers used 30 synthetic tutor-student dialogues to assess the ability of ChatGPT to offer feedback on tutors' praise of learners' work based on a specific set of criteria. They compared the performance of GPT-4 to that of three experienced and trained human reviewers. Analysis of the data revealed that GPT-4's performance was consistent with that of human reviewers on most criteria, but it underperformed on evaluation criteria that required integrating contextual information.

Implications of ChatGPT for ESL Writing Instruction

Some studies have examined the implications of ChatGPT for ESL writing instruction and demonstrated that it can foster L2 writing practice and development in many ways.

Affective Factors

First, a few studies have examined the effects of ChatGPT use on the affective factors in writing classes and reported that integrating ChatGPT in writing classes can foster student motivation (Fuchs, 2023) and promote writing practice. For example, in an experimental study examining the impact of ChatGPT on student anxiety in writing classrooms with 73 undergraduate English students, Hawanti and Zubaydulloevna (2023) compared the anxiety levels of writing students with and without access to ChatGPT. The study found that anxiety levels were significantly lower for the group with access to ChatGPT and concluded that integrating ChatGPT in writing classes could increase students' optimism and deliver real-time improvements in writing. In another study, Wambsganss et al. (2022) explored the impact of using an AI-powered social comparison nudge on writing students' performance with 71 writing students. The researchers divided the students into two groups; both groups received automatic feedback, but only the experimental group received a social comparison nudge. Analysis of students' essays revealed that the experimental group wrote more convincing and better-developed argumentative essays. The researchers concluded that the use of AI-generated nudges motivated students to write more effective argumentative texts by triggering basic motivational mechanisms. Similarly, Han et al. (2023) explored the use of ChatGPT in English writing courses at a Korean university with over 200 participants. The study did not reach clear conclusions about the usefulness of AI for writing, but it demonstrated that participants were highly satisfied with the experience. In another study, Ali et al. (2023) used a questionnaire to gather data about the implications of ChatGPT for English writing instruction. The results demonstrated that ChatGPT was highly useful and motivational for writing students.

Instructional Assistance

A number of studies also suggested that ChatGPT can provide instructional assistance and scaffolding to students' writing practice. For instance, Marzuki et al. (2023) used interview data to explore how four ESL teachers utilized ChatGPT in English writing classes in Indonesia. The study revealed that ChatGPT

improved students' writing output by suggesting effective wording and detecting logical inconsistencies. Similarly, using questionnaire data, Wulandari et al. (2024) explored the benefits of using ChatGPT with 20 English teachers from junior high schools in Indonesia. The study revealed that ChatGPT assisted teachers in enhancing students' writing skills. In a similar vein, Ljujić et al. (2023) compared ChatGPT-written essays about multimedia with highly ranked students' essays. They found that ChatGPT could provide instructional assistance to teachers with regard to grading and instruction, and it could also scaffold students' drafting of writing projects. In another study, Kasneci et al. (2023) conducted a scoping review of the literature on the benefits of AI for writing instruction and reported that it can help instructors teach research and writing skills and train students on mundane writing processes. Similarly, Imran and Almusharraf (2023) reviewed 30 articles examining the use of ChatGPT as a writing assistant and concluded that it can assist learners in generating drafts, brainstorming ideas, and summarizing research articles. In another review, Mhlanga (2023) synthesized the findings of 23 articles on the implications of ChatGPT for education and reported that ChatGPT can enhance students' creativity and analytical thinking by allowing them to explore different writing techniques.

Feedback

In addition to its ability to support L2 writing practice and instruction, several studies have demonstrated that ChatGPT could offer L2 writing students automated feedback on their writing (Bonner et al., 2023). Dai et al. (2023) compared ChatGPT-generated and instructors' feedback on 103 students' reports. The study revealed that ChatGPT provided more detailed, readable, and consistent feedback to students. Similarly, Loem et al. (2023) evaluated the ability of ChatGPT-3 to correct grammatical errors using a variety of prompt designs. The results suggested that ChatGPT was responsive to individual needs, provided the prompts are well designed, and demonstrated that the suggested corrections were accurate and beneficial for learners. In a similar vein, Hawanti and Zubaydulloevna (2023) examined the potential of ChatGPT to support writing instruction in classroom settings. They found that ChatGPT provided instant feedback that enabled students to promptly fix their mistakes and mitigated students' concerns over errors. In another study, Üstünbaş (2024) used semi-structured interviews to explore the ability of ChatGPT to offer corrective feedback to five English writing students at a Turkish university. The study revealed that ChatGPT's feedback was perceived by learners to be useful and accessible, but it lacked the social dimension of human feedback. Also, Davishi et al. (2024) examined the quality of feedback offered by an AI platform, RIPPLE, with 1,625 students across 10 courses using 16,007 peer reviews at an Australian university. The results showed that RIPPLE offered students high-quality feedback and that the quality of their reviews dropped when access to RIPPLE was removed. In addition, Athanassopoulos et al. (2023) examined the impact of ChatGPT-generated feedback on writing quality with a small group of teenage German writing students. Analysis of participants' writing before and after ChatGPT use demonstrated that automated feedback improved participants' vocabulary choice and sentence structure. On the other hand, a study by Yoon et al. (2023) compared the quality of human and GPT-4 feedback on 50 students' essays. They found that GPT-4 feedback was generic and abstract and it failed to identify substantial issues in participants' writings. They suggested that a trained and optimized version of ChatGPT may give more useful and specific feedback.

Discussion

The present review of the available literature demonstrates that ChatGPT has impressive instructional potential that can foster L2 writing pedagogy. One of the essential trends that emerged from the literature review is the potential of ChatGPT integration to boost L2 writers' motivation (Fuchs, 2023). It appears that the availability of ChatGPT to L2 writers as an intelligent tutor can lower their affective filters, which fosters their motivation to experiment with L2 writing and use ChatGPT's feedback to develop their L2 writing skills through trial and error (Wambsganss et al., 2022). A potential explanation for this finding is that the use of machine-generated assistance and feedback removes the social awkwardness that emerges from the face-threatening practice of exposing one's weaknesses and mistakes to others. As a result, students do not experience high levels of anxiety when receiving guidance from ChatGPT (Hawanti & Zubaydulloevna, 2023), which results in a positive learning experience that fosters L2 writers' motivation to experiment with their writing skills (Wambsganss et al., 2022). Incorporating positive affect into the normally face-threatening task of practicing writing in a foreign language should result in extended L2 writing practice and considerable development in students' writing skills (Kasneci et al., 2023; Ljujić et al., 2023; Mhlanga, 2023; Wulandari et al., 2023).

Another potential of ChatGPT to support L2 writing instruction that emerged from the literature review was its ability to provide automated instructional support to L2 writers and scaffold their L2 writing practice (Ljujić et al., 2023; Wulandari et al., 2023). Specifically, the current literature demonstrates that ChatGPT can guide students' practice on standard writing processes, such as brainstorming and outlining ideas, by offering explanations, tips, and examples (Imran & Almusharraf, 2023; Kasneci et al., 2023; Marzuki et al., 2023). This potential appears to (a) offer L2 writers autonomous and individualized learning trajectories tailored to each learner's specific learning needs, (b) provide them with rich opportunities for exploring different writing styles and genres, and (c) promote extended L2 writing practice, all of which are conducive to improving their L2 writing skills and enhancing their creativity and critical thinking (Kasneci et al., 2023; Mhlanga, 2023; Wulandari et al., 2023). Also, the ability of the system to learn from human feedback (Christiano et al., 2017) could allow ChatGPT to adapt and adjust these learning trajectories to learners' needs based on performance data, ensuring the relevance of the learning experience to learners. In addition, using ChatGPT as an instructional assistant would allow writing instructors to delegate instructional design and administration tasks to ChatGPT so they could use most of their time to design and manage active learning opportunities. Freeing instructors' time from mundane administrative responsibilities would give them the time and energy to transform traditional L2 writing classrooms into communities of practices focused on the dissemination of L2 writing literacies through collaborative interaction and apprenticeship (Lave & Wenger, 1991), which was previously infeasible because instructors were hampered by their administrative workloads.

Another transformative potential that ChatGPT brings to L2 writing classrooms, according to the literature, is the ability to automate the processes of offering students on-demand, personalized feedback on their L2 writing output (Bonner et al, 2023; Dai et al., 2023; Davishi et al., 2024). Traditionally, one of the main challenges faced by L2 writers was the scarcity of personalized feedback on their writing output due to limited instructional time and resources. The ability of ChatGPT to offer automatic, instant, and frequent feedback on L2 writers' output would overcome this limitation by expanding the volume and frequency of corrective feedback they can receive on their writing. This should help them capitalize on their personal

strengths and address their individual learning needs (Athanassopoulos et al., 2023) and, as a result, expedite their L2 writing skills development trajectories (Hawanti & Zubaydulloevna, 2023). Also, the automation of feedback delivery would offer L2 writers a safe environment for experimental learning (Üstünbaş, 2024) and motivate them to engage in L2 writing practice more frequently, which should boost their L2 writing skills (Kasneci et al., 2023).

Conclusion

The present study attempted to shed some light on the underexplored potentials and implications of ChatGPT for L2 writing instruction. To this end, it reviewed existing research on the potentials and implications of using ChatGPT on various aspects of writing and L2 instruction. It was hoped that such an account would (a) underline the potentials and implications of ChatGPT for L2 writing instruction, (b) promote the integration of ChatGPT in L2 writing classes, and (c) identify important directions for future research in this new research area. Analysis of the available literature to date demonstrated that ChatGPT has great potential as an instructional assistant in L2 writing classes because it could automate instructional design and assessment processes, allowing writing instructors to focus their energy on creating a community of practice centered around L2 writing skills and literacies. The literature further revealed that ChatGPT could personalize L2 writers' learning trajectories, adjust instructional resources to their individual learning needs, and adapt learning activities to their learning progress based on their performance data. The literature also demonstrated that ChatGPT can overcome instructional problems in traditional writing classes, especially limited access to corrective feedback on students' writing, as ChatGPT can produce regular and frequent feedback, which is crucial for students to develop their writing skills. On the other hand, the review demonstrated that the system is not without limitations and that it needs specialized training on instructional design and assessment processes to perform these tasks more effectively.

Practical Applications

These findings have a number of important practical applications. First, automating the design of instructional materials will change the role of the L2 writing instructor from that of a lecturer to that of a mentor. The instructor will move from the role of a teacher who imparts knowledge and offers feedback to that of a mentor who supports and monitors the learning process by making effective pedagogical decisions (based on learning analytics from AI) and guiding learners along their learning trajectories (Huang & Li, 2023). Second, given the proliferation of generative AI and the growing dependency of young learners on new technologies in communication and meaning-making activities, it is fair to assume that generative AI skills are going to be key digital literacies that college composition courses will have to integrate in order to equip learners with the necessary skills to use AI ethically and effectively to compose texts. Therefore, it is vital for L2 writing courses to include a new digital literacies component that emphasizes generative AI skills such as prompt design and fair use of AI-generated content (Baskara & Mukarto, 2023). Third, the proliferation of AI-generated content and the increasingly humanlike nature of AI-generated texts necessitates a growing emphasis on critical thinking and research skills to equip learners with the skills to locate and evaluate information in the age of AI-dominated content generation (Qadir, 2022). Fourth, the demand for integrating generative AI literacies in L2 writing courses and the potential the technology can bring to instructional contexts necessitates that L2 educators develop their digital literacy skills so they can

manipulate AI tools and take full advantage of them in their classrooms (Huang & Li, 2023). This would involve acquiring the digital competencies needed to use these tools pedagogically, capitalize on their affordances, and develop a critical awareness of their limitations and risks (Kohnke et al., 2023). Finally, the rapid proliferation of generative AI technologies suggests that educational institutions will have to consider their potentials and limitations (Qadir, 2022) and develop clear guidelines for integrating these tools into their instruction and assessment to prepare students for a new world where AI-powered tools are becoming mainstream communication tools (Kohnke et al., 2023).

Limitations and Future Research Directions

The present study has a number of limitations that point to important future research directions. First, the study offered a systematic review of the literature on ChatGPT's implications for L2 writing instruction, but the empirical investigations to date have been limited and have been primarily descriptive, theoretical, or exploratory in nature, and thus, any conclusions drawn from the literature are tentative. Future research should investigate specific instructional potentials of ChatGPT in depth using rigorous quantitative and qualitative research methods. For instance, future research could explore the accuracy and reliability of ChatGPT's grading of academic essays according to a specific rubric. Second, the present study focused primarily on the positive implications of ChatGPT for instruction and did not cover the potential drawbacks of ChatGPT for L2 writing. Therefore, future research should consider how L2 writers use ChatGPT in real life and the potential drawback of generative AI for L2 writing instruction, especially with regard to violations of academic integrity. Third, the present review demonstrates that ChatGPT can offer promising instructional potential in L2 writing classes based on generic training and general skills that it elicited from general training data. In light of that, it is fair to suggest that these instructional potentials can improve dramatically with specialized training on highly curated datasets exhibiting desired instructional performance (i.e., fine-tuning). Therefore, future research should investigate the training of a specialized GPT version on L2 writing instruction and the impact of specialized training on its performance as an instructional assistant.

References

Ali, J. K. M., Shamsan, M. A. A., Hezam, T. A., & Mohammed, A. A. (2023). Impact of ChatGPT on learning motivation: Teachers and students' voices. *Journal of English Studies in Arabia Felix*, *2*(1), 41–49. https://doi.org/10.56540/jesaf.v2i1.51

Altamimi, A. B. (2023, August). Effectiveness of ChatGPT in essay autograding. In *2023 International Conference on Computing, Electronics & Communications Engineering (iCCECE)* (pp. 102–106). IEEE. https://doi.org/10.1109/iCCECE59400.2023.10238541

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

Ibrahim. K (2023). Using AI-based classifiers to detect AI-generated texts in ESL composition. *Language Testing in Asia*,*13,* 46. https://doi.org/10.1186/s40468-023-00260-2

Baidoo-anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, *7*(1), 52–62. https://doi.org/10.61969/jai.1337500n

Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, *57*, 100745. https://doi.org/10.1016/j.asw.2023.100745

Baskara, R., & Mukarto, M. (2023). Exploring the implications of ChatGPT for language learning in higher education. *Indonesian Journal of English Language Teaching and Applied Linguistics*, *7*(2), 343–358. https://files.eric.ed.gov/fulltext/EJ1391490.pdf

Bonner, E., Lege, R., & Frazier, E. (2023). Large language model-based artificial intelligence in the language classroom: Practical ideas for teaching. *Journal of Teaching English with Technology*, *23*(1), 23–41. http://dx.doi.org/10.56297/BKAM1691/WIEO1749

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam P., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., ... & Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, *33*, 1877-1901. https://doi.org/10.48550/arXiv.2005.14165

Christiano, P. F., Leike, J., Brown, T. B., Martic, M., Legg, S., & Amodei, D. (2017). Deep reinforcement learning from human preferences. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Eds.), *Advances in Neural Information Processing Systems: Vol. 30* (pp. 4299–4307). https://proceedings.neurips.cc/paper_files/paper/2017/file/d5e2c0adad503c91f91df240d0cd4e49-Paper.pdf

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, *61*(2), 228–239. https://doi.org/10.1080/14703297.2023.2190148

Dai, W., Lin, J., Jin, H., Li, T., Tsai, Y.-S., Gašević, D., & Chen, G. (2023, July). Can large language models provide feedback to students? A case study on ChatGPT. In *2023 IEEE International Conference on Advanced Learning Technologies (ICALT)* (pp. 323–325). IEEE. https://doi.org/10.1109/ICALT58122.2023.00100

Darvishi, A., Khosravi, H., Abdi, S., Sadiq, S., & Gašević, D. (2022, June). Incorporating training, self-monitoring and AI-assistance to improve peer feedback quality. In *Proceedings of the ninth ACM conference on learning@ scale* (pp. 35-47). https://doi.org/10.1145/3491140.3528265

Francke, E., & Alexander, B. (2019). The potential influence of artificial intelligence on plagiarism a higher education perspective. In Proceedings of the European Conference on the Impact of Artificial Intelligence and Robotics (pp. 131–140). EM Normandie Business School, Oxford. https://doi.org/10.1186/s40468-023-00260-2

Fuchs, K. (2023). Exploring the opportunities and challenges of NLP models in higher education: Is ChatGPT a blessing or a curse? *Frontiers in Education*, *8*, 1–4. https://doi.org/10.3389/feduc.2023.1166682

Gao, C. A., Howard, F. M., Markov, N. S., Dyer, E. C., Ramesh, S., Luo, Y., & Pearson, A. T. (2023). Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers. *NPJ Digital Medicine*, *6*, article 75. https://doi.org/10.1038/s41746-023-00819-6

Han, J., Yoo, H., Myung, J., Kim, M., Lee, T. Y., Ahn, S.-Y., & Oh, A. (2023). Exploring student-ChatGPT dialogue in EFL writing education. In *Thirty-Seventh Annual Conference on Neural Information Processing Systems*. Neural Information Processing Systems Foundation. https://gaied.org/neurips2023/files/19/19_paper.pdf

Haque, M. U., Dharmadasa, I., Sworna, Z. T., Rajapakse, R. N., & Ahmad, H. (2022). *"I think this is the most disruptive technology": Exploring sentiments of ChatGPT early adopters using Twitter data*. arXiv. https://doi.org/10.48550/arXiv.2212.05856

Hawanti, S., & Zubaydulloevna, K. M. (2023). AI chatbot-based learning: Alleviating students' anxiety in English writing classroom. *Bulletin of Social Informatics Theory and Application*, *7*(2), 182–192. https://doi.org/10.31763/businta.v7i2.659

Hirunyasiri, D., Thomas, D. R., Lin, J., Koedinger, K. R., & Aleven, V. (2023). *Comparative analysis of GPT-4 and human graders in evaluating praise given to students in synthetic dialogues*. arXiv. https://doi.org/10.48550/arXiv.2307.02018

Hong, W. C. H. (2023). The impact of ChatGPT on foreign language teaching and learning: Opportunities in education and research. *Journal of Educational Technology and Innovation*, *5*(1). https://jeti.thewsu.org/index.php/cieti/article/view/103

Hu, K. (2023, February 2). *ChatGPT sets record for fastest-growing user base: Analyst note*. Reuters. https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/

Huang, J., & Li, S. (2023). Opportunities and challenges in the application of ChatGPT in foreign language teaching. *International Journal of Education and Social Science Research*, *6*(4), 75–89. http://dx.doi.org/10.37500/IJESSR.2023.6406

Imran, M., & Almusharraf, N. (2023). Analyzing the role of ChatGPT as a writing assistant at higher education level: A systematic review of the literature. *Contemporary Educational Technology*, *15*(4), ep464. https://doi.org/10.30935/cedtech/13605

Jia, Q., Young, M., Xiao, Y., Cui, J., Liu, C., Rashid, P., & Gehringer, E. (2022). Insta-reviewer: A data-driven approach for generating instant feedback on students' project reports. *Proceedings of the 15th International Conference on Educational Data Mining* (pp. 5–16). International Educational Data Mining Society. https://zenodo.org/doi/10.5281/zenodo.6853098

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … & Kasneci, G. (2023). *ChatGPT for good? On opportunities and challenges of large language models for education* [Position paper]. Department of Educational Sciences, Technische Universität München. https://www.edu.sot.tum.de/fileadmin/w00bed/hctl/_my_direct_uploads/ChatGPT_for_Good_.pdf

Khalil, M., & Er, E. (2023). *Will ChatGPT get you caught? Rethinking of plagiarism detection*. arXiv. https://doi.org/10.48550/arXiv.2302.04335

Koltovskaia, S. (2020). Student engagement with automated written corrective feedback (AWCF) provided by Grammarly: A multiple case study. *Assessing Writing*, *44*, 100450.

King, M. R., & ChatGPT. (2023). Editorial: A conversation on artificial intelligence, chatbots, and plagiarism in higher education. *Cellular and Molecular Bioengineering*, *16*, 1–2. https://doi.org/10.1007/s12195-022-00754-8

Kınık, B., & Çetin, H. (2024). Human vs. AI: The use of ChatGPT in writing assessment. In Z. Çetin Köroğlu & A. Çakır (Eds.), *Fostering foreign language teaching and learning environments with contemporary technologies* (pp. 194–215). IGI Global. https://doi.org/10.4018/979-8-3693-0353-5.ch009

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, *54*(2), 537–550. https://doi.org/10.1177/00336882231162868

Kortemeyer, G. (2023). *Performance of the pre-trained large language model GPT-4 on automated short answer grading*. arXiv. https://doi.org/10.48550/arXiv.2309.09338

Lage, M. J., Platt, G. J., & Treglia, M. (2000). Inverting the classroom: A gateway to creating an inclusive learning environment. *Journal of Economic Education*, *31*(1), 30–43. https://doi.org/10.1080/00220480009596759

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press. https://doi.org/10.1017/CBO9780511815355

Ljujić, B., Miljković, J., & Grozdić, V. (2023). ChatGPT and academic writing in higher education. In M. Radulović & M. Trajković, *Towards a more equitable education: From research to change* (pp. 104–109). Institute for Educational Research. https://reff.f.bg.ac.rs/handle/123456789/5759

Loem, M., Kaneko, M., Takase, S., & Okazaki, N. (2023). *Exploring effectiveness of GPT-3 in grammatical error correction: A study on performance and controllability in prompt-based methods*. arXiv. https://doi.org/10.48550/arXiv.2305.18156

Lund, B. D., & Wang, T. (2023). Chatting about ChatGPT: How may AI and GPT impact academia and libraries? *Library Hi Tech News*. https://dx.doi.org/10.2139/ssrn.4333415

MacNeil, S., Tran, A., Mogil, D., Bernstein, S., Ross, E., & Huang, Z. (2022, August). Generating diverse code explanations using the GPT-3 large language model. *Proceedings of the ACM Conference on International Computing Education Research, 2*, 37–39. https://doi.org/10.1145/3501709.3544280

Marzuki, 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), 2236469. https://doi.org/10.1080/2331186X.2023.2236469

Mhlanga, D. (2023). OpenAI in education, the responsible and ethical use of ChatGPT towards lifelong learning. SSRN. http://dx.doi.org/10.2139/ssrn.4354422

OpenAI. (n.d.-b). *Overview*. Retrieved January 5, 2024, from https://openai.com/chatgpt

Pardos, Z. A., & Bhandari, S. (2023). *Learning gain differences between ChatGPT and human tutor generated algebra hints*. arXiv. https://doi.org/10.48550/arXiv.2302.06871

Qadir, J. (2022). *Engineering education in the era of ChatGPT: Promise and pitfalls of generative AI for education*. TechRxiv. https://doi.org/10.36227/techrxiv.21789434.v1

Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). *Improving language understanding by generative pre-training*. Retrieved June 30, 2023, from https://www.mikecaptain.com/resources/pdf/GPT-1.pdf

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. *OpenAI Blog*, *1*(8), 9. https://insightcivic.s3.us-east-1.amazonaws.com/language-models.pdf

Rahman, M. M., & Watanobe, Y. (2023). ChatGPT for education and research: Opportunities, threats, and strategies. *Applied Sciences*, *13*(9), 5783 https://doi.org/10.3390/app13095783

Rakhmonov, I. U., & Kurbonova, R. S. (2023). The pedagogical principles and effectiveness of utilizing ChatGPT for language learning. *Research and Education*, *2*(9), 226–243. https://researchedu.org/index.php/re/article/view/4899

Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching, 6*(1), 342–363. https://doi.org/10.37074/jalt.2023.6.1.9

Shaikh, S., Yayilgan, S. Y., Klimova, B., & Pikhart, M. (2023). Assessing the usability of ChatGPT for formal English language learning. *European Journal of Investigation in Health, Psychology, and Education*, *13*(9), 1937–1960. https://doi.org/10.3390/ejihpe13090140

Solaiman, I., Brundage, M., Clark, J., Askell, A., Herbert-Voss, A., Wu, J., Radford, A., Krueger, G., Kim, J. W., Kreps, S., McCain, M., Newhouse, A., Blazakis, J., McGuffie, K., & Wang, J. (2019). *Release strategies and the social impacts of language models*. arXiv. https://doi.org/10.48550/arXiv.1908.09203

Sridhar, P., Doyle, A., Agarwal, A., Bogart, C., Savelka, J., & Sakr, M. (2023). *Harnessing LLMs in curricular design: Using GPT-4 to support authoring of learning objectives*. arXiv. https://doi.org/10.48550/arXiv.2306.17459

Susnjak, T. (2022). *ChatGPT: The end of online exam integrity?* arXiv. https://doi.org/10.48550/arXiv.2212.09292

Üstünbaş, Ü. (2024). EFL learners' views about the use of artificial intelligence in giving corrective feedback on writing: A case study. In Z. Çetin Köroğlu & A. Çakır (Eds.), *Fostering foreign language teaching and learning environments with contemporary technologies* (pp. 115–133). IGI Global. https://doi.org/10.4018/979-8-3693-0353-5.ch006

Wambsganss, T., Janson, A., & Leimeister, J. M. (2022). Enhancing argumentative writing with automated feedback and social comparison nudging. *Computers & Education*, *191*, 104644. https://doi.org/10.1016/j.compedu.2022.104644

Wang, Y., Derakhshan, A., Pan, Z., & Ghiasvand, F. (2023). Chinese EFL teachers' writing assessment feedback literacy: A scale development and validation study. *Assessing Writing*, *56*, 100726. https://doi.org/10.1016/j.asw.2023.100726

Wulandari, F., Astuti, M. T., & Marhamah, M. (2024). Enhancing writing literacy teachers through AI development. *Jurnal Onoma: Pendidikan, Bahasa, dan Sastra*, *10*(1), 246–256. https://doi.org/10.30605/onoma.v10i1.3175

Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies*, *28*, 13943–13967. https://doi.org/10.1007/s10639-023-11742-4

Yeadon, W., Inyang, O.-O., Mizouri, A., Peach, A., & Testrow, C. (2022). *The death of the short-form physics essay in the coming AI revolution*. arXiv. https://doi.org/10.48550/arXiv.2212.11661

Yoon, S.-Y., Miszoglad, E., & Pierce, L. R. (2023). *Evaluation of ChatGPT feedback on ELL writers' coherence and cohesion*. arXiv. https://doi.org/10.48550/arXiv.2310.06505

Young, J. C., & Shishido, M. (2023). Investigating OpenAI's ChatGPT potentials in generating chatbot's dialogue for English as a foreign language learning. *International Journal of Advanced Computer Science and Applications*, *14*(6). https://doi.org/10.14569/IJACSA.2023.0140607