Journal of Theoretical and Applied Information Technology
15th November 2025. Vol.103. No.21
© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

ASSESSING THE EFFECTIVENESS OF CHATGPT-MEDIATED INSTRUCTION IN DEVELOPING PARAGRAPH WRITING SKILLS

[1]TAJ MOHAMMAD, [*2]ALI ABBAS FALAH ALZUBI, [3]MOHD NAZIM, [4]SOADA IDRIS KHAN

[1] Associate Professor, English Skills Department, Preparatory Year, Najran University, Najran, Saudi Arabia
[*2] Assistant Professor of Applied Linguistics, Department of English, College of Languages and Translation, Najran University, Kingdom of Saudi Arabia (Corresponding author)
[3] Associate Professor, Department of English, College of Languages and Translation, Najran University, Kingdom of Saudi Arabia
[4] Ph.D. in Applied Linguistics

E-mail: [1] tajmohd09@gmail.com, [*2] aliyarmouk2004@gmail.com, [3] nazimspeaking@yahoo.co.in, [4] khan.soada@gmail.com

ABSTRACT

The integration of Artificial Intelligence (AI) tools into education has recently captured the attention of educators and researchers, as they have the potential to improve teaching and learning. Among these tools, ChatGPT has become a key asset for writing instruction. However, its impact on discrete writing skills such as paragraph development is not well studied. This study investigates the effectiveness of ChatGPT-mediated instruction in enhancing the paragraph writing skills of English as a Foreign Language (EFL) students. Using a quasi-experimental design, the research included two groups: an experimental group (N = 30) that received writing instruction with ChatGPT support, and a control group (N = 30) that received conventional instruction. Data were gathered using pre- and post-tests to assess growth in paragraph writing skills and semi-structured interviews to document students' experiences with ChatGPT-supported instruction. The results indicated that the experimental group achieved notable gains in paragraph writing. The findings suggest practical steps for EFL teachers to integrate ChatGPT as an innovative tool to strengthen students' writing competence. Additionally, curriculum designers should create clear guidelines and professional development to facilitate these applications.
Keywords: *ChatGPT-Mediated Instruction, Paragraph Writing Skills, Experimental Research Design.*

1. INTRODUCTION

Writing is an essential skill for academic progression, professional communication, and personal expression [1]. As a reflective process, it fosters critical thinking and creativity, requiring structured thought, clarity of expression, and a strong command of language elements to achieve effective learning outcomes [4]. The integration of Information Technology in writing instruction has recently revolutionized teaching methods by enhancing engagement, feedback, and accessibility [2]. Tools like ChatGPT support personalized learning and foster collaboration [3]. However, the effective use of ChatGPT-mediated instruction requires thoughtful planning to ensure that technology complements rather than replaces core writing skills [4], while also addressing ethical concerns such as plagiarism and data privacy [5]. Therefore, this study seeks to answer the following research question: Does integrating writing pedagogy with IT tools, specifically ChatGPT, create enriched learning environments that empower EFL students to become skilled and confident writers?

Paragraph writing is a challenge for EFL students in Saudi Arabia, particularly at Najran University, with less research on ChatGPT's effectiveness in addressing paragraph-specific skills, such as mechanics, language use, content organization, and phonological interference. Using experimental and qualitative methodologies, this study examines the impact of ChatGPT-mediated instruction on these skills.

In Saudi Arabia, both students and teachers have consistently expressed that paragraph writing poses a significant challenge. Paragraph writing is essential for communicating concepts and information effectively [3]. Coherent paragraphs make ideas easy to understand and communicate

them coherently. Writing well-structured paragraphs effectively draws in and holds the audience's attention, increasing the efficacy of communication [4]. Effective communication relies on tone, topic relevance, clarity, and well-crafted paragraphs [5]. Paragraph organization alone is insufficient if clarity and interest are not guaranteed. Relevance and quality are essential for capturing an audience's attention [6]. If one concentrates on the structure of paragraphs, one may miss other aspects [3]. Message efficacy increases when using an engaging style and matching content to the audience's needs. Message impact and engagement are maximized by combining these components and paragraph structures [7]. The use of a synergistic strategy results in a message that is both compelling and cohesive [6].

Due to the nature of paragraph writing skills, it has always been challenging for EFL learners to master paragraph writing [5]. Teachers have been applying various methods and techniques to enhance students' paragraph writing [32]. Technological integration in the EFL classroom is one of the efforts to address the issues [6]. Consequently, various AI-mediated tools, such as Quillbot, Gemini, Grammarly, and ChatGPT, have been incorporated into the EFL writing classroom [8]. Gemini produced significant improvements in lexical correctness and sentence variety, along with improvements in coherence, syntax, and paragraph organization [7]. Grammarly helped students refine paragraphs by correcting errors in mechanics (spelling & punctuation) and enhancing clarity [5]. Students' proficiency in synonyms, sentence construction, and word choice all improved. The responses are exciting and insightful about using QuillBot to improve in paraphrasing [45]. Like other AI tools, ChatGPT also significantly improved the paragraph-level academic writing skills of tertiary-level ESL students, including grammar, vocabulary, and coherence. It also reduced anxiety and enhanced sentence structure and lexical variety [10]. According to [2, p.3], ChatGPT is considered a "valuable resource for students and professionals" in the field of education. Adiguzel et al. [2] claim that ChatGPT can access a wealth of textual data, including books, papers, and websites, to generate written results at the user's request. Since its inception, studies have confirmed that ChatGPT offers students amazing features, such as language translation, question responding, and draft writing [2], [11]. Students gain greatly from these skills, especially when it comes to academic writing [2], [5]. Students' greater motivation and engagement in academic writing is one of the main outcomes of using ChatGPT [2], [11]. Students can ask questions on ChatGPT, for instance, and modify their answers to suit their requirements and preferences [5]. This capability helps students overcome obstacles such as anxiety and hesitancy in asking questions about their tasks and providing them with a more individualized learning experience [11].

Additionally, ChatGPT's advice and information inspire students to write more persuasively and with greater motivation [5]. Students can use ChatGPT to help them discuss ideas, identify sources, paraphrase phrases, verify grammar mistakes, and suggest topics for their writing papers [9], [4]. These skills were thought to be captivating, which makes EFL students more interested in writing [12]. ChatGPT promoted self-accomplishment and improved writing efficacy. Additionally, they discovered that ChatGPT use was positively correlated with writing self-efficacy, indicating that students felt better about their academic writing assignments the more they utilized ChatGPT [8].

ChatGPT, a relatively new AI tool based on a large language model, can provide meaningful writing samples, adjust the difficulty level of texts to match learners' proficiency levels [8], offer advice on various structural aspects of a text and translate it [13], and facilitate guided writing. ChatGPT, an advanced language generation model, provides real-time assistance with grammar, vocabulary selection, and composition. However, successful integration of these AI-based writing aids into pedagogical frameworks necessitates careful consideration of how to balance ChatGPT's support with the development of learners' intrinsic writing talents [12]. Achieving a delicate balance in which the enhancement of writing skill through AI aid complements the development of learners' intrinsic writing abilities is a vital pedagogical endeavor. The focus is on identifying techniques that successfully leverage ChatGPT's potential to provide invaluable help while overshadowing the critical process of self-initiated and proficient written discourse production [11].

ChatGPT's ability to translate languages, create writing drafts, and answer queries with instant feedback has been shown to improve students' writing productivity and engagement [2], [11]. Furthermore, students said ChatGPT was good for

writing because it helped them paraphrase, make logical solutions, and stay involved [11].

The study is delimited to investigate the impact of ChatGPT-mediated instruction on the paragraph writing skills of EFL learners at Najran University, with an emphasis on four basic domains: mechanics, language use, content organization, and phonological interference. The research's contribution is to address the underexplored field of paragraph writing by integrating experimental evidence with learner comments to provide a full evaluation of ChatGPT's role in writing pedagogy. Beyond contributing to academic discourse, the findings have practical implications for EFL instruction. They suggest that incorporating ChatGPT into classroom practice can improve learner accuracy, organization, and confidence, while also providing teachers with a supplementary tool to support personalized feedback and reduce writing anxiety.

Even though many studies have shown how AI-driven tools like ChatGPT can improve EFL writing, three major gaps still exist. First, most of the current research focuses on paraphrasing or general academic composition, neglecting paragraph writing, which is a recurring challenge for Saudi EFL learners. Second, there are not many studies that specifically examine the Saudi Arabian setting, and none that assess ChatGPT's effect on students at Najran University's ability to write paragraphs. Finally, while previous research highlights ChatGPT's benefits, a large portion of it relies on perception-based data instead of combining experimental results with learners' helpful criticism.

The data reported in this study adds to and extends earlier research by providing unique empirical evidence on ChatGPT's impact on paragraph writing skills among EFL learners at Najran University, a hitherto unexplored setting. Unlike previous research on general academic writing or paraphrasing [2], [10], this study focuses on paragraph-level skills in four domains: mechanics, language use, content structure, and phonological interference. By integrating quantitative data from a quasi-experimental design with qualitative insights from student interviews, the study provides a more comprehensive evaluation than previous perception-based [8] or solely quantitative studies [24], contributing unique, context-specific findings to the field of AI-mediated EFL writing instruction. The current study addresses these gaps by providing contextual understanding and empirical support for incorporating ChatGPT into EFL instruction to improve writing paragraphs. Therefore, the primary objectives of this study are to investigate any significant differences in performance between the experimental and control groups prior to the intervention, to evaluate the substantial difference in post-test achievement scores between these groups following the intervention, and to document the experiences of EFL students using ChatGPT as a learning aid in the writing classroom.

1.1. Research Questions

1. How effective is the use of ChatGPT-mediated instruction in enhancing EFL students' paragraph writing skills?
2. What are EFL students' experiences of using ChatGPT as a learning aid in paragraph writing?

2. THEORETICAL FRAMEWORK

This study aligns with Computer-Assisted Language Learning (CALL) theory, which serves as a foundation for understanding how technology might help with second and foreign language learning, particularly writing skills. CALL stresses the use of digital resources in language learning to give learners real input, immediate feedback, meaningful engagement, and chances for self-directed learning [14]. In the context of this study, CALL is particularly pertinent to the employment of ChatGPT as an AI-powered writing helper to improve EFL students' writing abilities. ChatGPT is an interactive tool that allows learners to develop, modify, and perfect their writing, which aligns with the CALL principles of increasing linguistic accuracy, fluency, and learner autonomy [15]. ChatGPT allows students to practice actual writing activities that mirror the real world. It reflects CALL's emphasis on contextualized practice and task authenticity, both of which are necessary for skill transfer outside of the classroom [16].

Furthermore, CALL necessitates giving learners influence over the writing process [17], and ChatGPT allows students to plan, compose, edit, and receive feedback at their own pace, increasing their confidence and proficiency in business communication [15]. In this sense, CALL theory not only explains ChatGPT's pedagogical integration but also emphasizes its function in connecting traditional EFL education to technology-enhanced learning environments.

This study reveals how AI tools can be effectively used to encourage precision, clarity, and professionalism in students' paragraph writing skills that are critical for academic achievement.

2.1. Literature Review

Writing is a challenging skill for EFL students. Paragraph writing becomes more difficult as it requires a good command of writing precisely, accurately, coherently, etc. Khuwaileh and Al Shoumali [18] examined Jordanian students' writing in Arabic, their first language, and English, their second language. They discovered that 55% of the students' composed paragraphs in their mother language that were disorganized and lacked proper linking of concepts. Lakkis and Abdel Malak [19] observed similar findings about Jordanian college students' writing arrangement.

On the other hand, Strauch [20] states that when writing in English, authors must begin each paragraph with a topic sentence and conclude each one with a concluding sentence. A general statement made up of a topic and a comment, the topic sentence should be followed by some background information.

To facilitate students, many techniques have been introduced from time to time. Integration of technology and AI-mediated tools is the best example of it. An AI-assisted language learning application was created by [21] to help EFL students overcome cognitive obstacles and enhance their writing abilities. The results showed that the AI-powered language learning tool successfully improved students' writing abilities and decreased the cognitive obstacles they faced when completing writing assignments. Liu et al. [22] investigated the effect of AI on the writing abilities of EFL learners using a quasi-experimental research approach. The results showed significant gains in writing abilities as compared to the traditional curriculum. Effective writing performance was also facilitated by the AI-supported language learning technique, which also improved learners' self-efficacy, self-regulated learning, and decreased cognitive load.

Using AI-mediated tools in the writing classroom has been very rewarding for students. Ng et al. [7] conducted a study that investigated the impact of Gemini on the writing abilities and motivation of EFL learners. Sixty Omani learners were divided into treatment and control groups, each with 30 learners. The treatment group received extra feedback from Gemini, leading to improved performance. The study highlights the importance of Gemini in enhancing EFL learners' motivation and writing abilities. Likewise, Wordcraft, an AI-powered text editor, was used by [23] to assist experienced authors with creative writing activities, specifically brainstorming. The outcomes showed that the AI-powered tool had a beneficial effect on creative writing.

ChatGPT assisting EFL learners in honing their writing abilities, AI has completely transformed language learning and writing teaching. This tool addresses issues like organization and grammatical accuracy by offering interactive learning opportunities and immediate input. Numerous researchers have investigated the usage of ChatGPT in paragraph writing. ChatGPT is one of the AI-mediated tools that has been very popular with students because of its user-friendliness. Yan [24] investigated how ChatGPT, an AI-powered language learning tool, affected the English writing of EFL students. The study looked at how students responded and thought about using the AI-assisted tool to improve their academic writing skills. The findings showed that the AI tool significantly improved students' writing abilities and increased their productivity when completing writing assignments.

According to [25], using ChatGPT in language learners' writing classes helped them complete argumentative writing assignments. They highlighted how difficult it is for teachers or peers to give constructive criticism in interactive writing exercises due to the linguistic and structural complexity of these assignments. This was addressed with the introduction of ChatGPT, which greatly improved students' argumentative writing assignments by offering more effective feedback and comments on language style, organization, and substance.

Mahapatra [10] investigates the efficacy of ChatGPT as a formative feedback tool for ESL undergraduates. A mixed-methods approach that includes exams and focus group discussions revealed a considerable positive influence on academic writing skills, including improvements in grammar, coherence, and organization. Students praised ChatGPT's ability to provide thorough critiques and recommendations. Some, however, noted that it occasionally enforced formulaic patterns, which may have hampered content organization creativity.

Özçelik's [26] study examines the impact of ChatGPT on 11 students' English language writing register knowledge, including formal, neutral, and informal. Using a pre-experimental approach, it was discovered that ChatGPT was

extremely effective in teaching formal register knowledge, assisting students in comprehending ideas relevant to paragraph writing, and providing constructive critique. It was less effective in casual and neutral registers, however, and students questioned its utility in these contexts. According to the study, ChatGPT is a beneficial learning tool, but it underlines the importance of teachers addressing its inadequacies in specific writing circumstances.

Abdi et al. [27] conducted a study in which five English Language Education Department students who had completed a paragraph writing course took part. Through interviews, it was found that ChatGPT taught students how to create paragraphs and provided helpful feedback to help them improve their writing. One issue that would have reduced the tool's utility was the occasional discrepancy between student questions and ChatGPT's responses. According to the study, teachers should play a key role in helping students comprehend ChatGPT's limits. ChatGPT should be used in writing classes as an extra resource to supplement traditional teaching methods.

Song and Song's [28] study evaluated ChatGPT's impact on Chinese EFL students' writing skills and motivation. Quantitative results showed significant improvements in organization, coherence, grammar, and vocabulary. Although the qualitative results showed how much students appreciated ChatGPT's customized input, they also raised concerns about contextual accuracy and potential over-reliance. The study highlights the need for ongoing development of AI tools to support, rather than obstruct, the development of long-term writing skills.

Abdullayeva and Musayeva [29] looked at how ChatGPT affected the writing abilities of EFL students and discovered that it helped by offering writing prompts, real-time feedback, and ideas for revision. An actual experimental investigation examining the impact of AI-assisted language learning on the writing abilities of EFL learners was carried out by [31]. The results showed that students who used the AI-powered tool performed better in their writing than those who did not. Additionally, in AI-supported writing tasks, the AI learners demonstrated high levels of behavioral, cognitive, and emotional involvement.

In examining the vast amount of research on AI-assisted language learning, many studies have significantly advanced our knowledge of how it affects language acquisition and the field of language learning in general [31], [4], [22]. These studies have offered important insights into a number of language learning topics, including vocabulary growth, grammar correction, and enhancements in language ability. However, when it is read more widely, there is still a big knowledge vacuum about the precise impact of AI on developing paragraph writing of EFL, especially at Najran University, and the present research addresses this knowledge gap.

3. METHODOLOGY

3.1. Research Design

The study utilized a quasi-experimental approach to evaluate the impact of ChatGPT-mediated writing instruction versus traditional instruction. The participants were divided into an experimental group (N=30) and a control group (N=30), which was created by grouping students within matched demographic and pre-test criteria. A pre-test consisting of a standardized 100-word paragraph generated a writing proficiency baseline, which was scored by a validated rubric that assessed the mechanics of language, language use contents, ideas, organization, and phonological interference of the mother tongue. Students in the experimental group participated in writing sessions (100 minutes each) conducted weekly, in which ChatGPT was used to generate outlines and revise drafts. The control group was taught by traditional methods and given the same amount of time and writing instruction (during which they were given manual, teacher-centered instructions on brainstorming, drafting, and revising the text). The intervention was three months long. Using a modified paragraph prompt, a post-test was given to measure any changes in writing performance. Within-group pre- and post-test scores were analyzed with paired t-tests, and for between-group comparisons, an independent t-test with pre-test scores as a covariate. Observations, experiences, and other challenges were explored in semi-structured interviews with 15 students from the experimental group, focusing on their experience with ChatGPT. The interviews were transcribed and thematically analyzed. To control for some extraneous variables (prior English proficiency level, motivation, effort, or attitude, Instructor influence, Familiarity with technology), the two groups were taught by the same instructor who used the same lesson plans and provided the same materials and resources to all students. The study spanned a total of four months: two weeks were

devoted to orientation before pretesting and intervention. Following the orientation, one week was allocated for pre-testing, three months for the intervention, and one week for post-testing and interviews.

3.2. Population and Sample

The participants of the study comprised 90 undergraduate students (aged 18-22) who were enrolled in the Reading and Writing course and were selected using purposeful sampling from PY (boys) in the second semester of the 2024-2025 academic year. Out of this population, a sample of sixty students was drawn. To achieve the objectives of the study, one treatment session was chosen via purposeful sampling. These students are Arabic-speaking learners and aspiring engineers, doctors, and computer science professionals who need to clear PY to continue their studies in their chosen fields. They come to PY after completing upper secondary education and meeting the enrollment prerequisites for Najran University. Their upper secondary education was completed in Najran. They studied English as a foreign language and were at the PY level between elementary and upper intermediate based on a diagnostic test.

The study's ethical approval reference number is 0076-00076-DS. The participants were informed of all the steps of the research prior to giving consent. In case they agreed to, participants were allowed to withdraw or skip any questions at any time. Furthermore, participants were free to ask questions pertaining to the research. They were informed that participation in the study would not result in any benefits, whether direct or indirect. Participants were provided with guarantees that any information collected would remain confidential and would exclusively be used for the purpose of the study. Furthermore, they were provided with the contact information of the research team in case they required further details or clarifications.

3.3. Instruments

The researchers used a test and a semi-structured interview to answer the research questions. Based on the literature analysis and their own teaching experience, the researchers created the test [11], [4]. By comparing the results of the pre- and post-tests, the test sought to evaluate the efficacy of the intervention program. The four main themes of the test were phonological interference, language use, mechanics, and contents and organization. Using the rubric, the students were required to compose a paragraph. "An Interesting Hobby" was the topic assigned to the students. According to the rubrics, they were expected to compose the paragraph that would be used for evaluation. The test took about twenty minutes on average to finish. The same population was given the tests by the researchers in the PY building.

A semi-structured interview was undertaken to gather EFL students' perspectives on the efficacy of ChatGPT-mediated training in enhancing paragraph writing skills. The interviews were conducted at Najran University during the second semester of the 2024-2025 academic year, in a designated lecture room located in the Deanship of Preparatory Year building, Language Skills Unit, which provided a peaceful, professional, and distraction-free environment.

Participants (N = 15) were chosen at random from the experimental group to ensure a balance of proficiency levels. The researcher interviewed each participant individually, while another colleague served as an observer, capturing nonverbal cues such as reluctance, confidence, and enthusiasm. The interviews lasted about 10-15 minutes per student and were performed over a two-week period following the post-test.

The interviews focused on challenging elements of writing, including phonological interference issues, language usage, content, ideas, organization, and mechanics. Based on prior research and the researcher's teaching experience, the semi-structured questions were created to ensure academic integrity [2], [11], [4].

The interview protocol was based on two key categories: experiences and observations.

- **Experiences**: Students were asked to describe how ChatGPT helped them generate, organize, and revise paragraphs, as well as how it impacted their confidence and motivation to write. Questions included the following: "How did ChatGPT change the way you approach paragraph writing?", "Which aspects of your writing improved the most (e.g., grammar, vocabulary, organization)?" or "What challenges did you encounter while using ChatGPT?"
- **Observations**: Students were invited to consider how ChatGPT changed their writing habits, methods, and learning process. These included the following questions: "What differences do you notice in your writing before and after using ChatGPT?", "Did ChatGPT make paragraph writing easier or more difficult for

you?", as well as "How did ChatGPT affect the way you edit and review your work?"

During the interviews, numerous themes appeared. Many students reported positive experiences, stating that ChatGPT assisted them in overcoming issues with language choice, sentence transitions, and logical flow. Others mentioned enhanced autonomy in amending drafts, greater awareness of grammatical accuracy, and better thought structure. However, a few participants highlighted concerns about over-reliance on AI or the tendency to accept its recommendations without critical analysis. Finally, all interview transcripts were examined using Braun and Clarke's [46] thematic content analysis method, which enabled the researcher to find recurring themes, patterns, and discrepancies across student responses.

3.4. Intervention

The intervention assessed whether ChatGPT-enhanced writing instruction improved Preparatory Year (PY) students' paragraph writing compared to traditional instruction. For four months (February-May 2025), the experimental group (N = 30) and control group (N = 30) met twice a week for 200 minutes (100 minutes per class) in the same university language lab. Both groups followed the same curriculum, class objectives, activities, and evaluation rubrics; the only variation was that the experimental group used ChatGPT for 50 minutes per session.

Instruction in both groups was based on the program's prescribed writing textbook units covering paragraph writing. It focuses on

1. Mechanics (spelling, capitalization, punctuation)
2. Language Use (Grammar and syntax, vocabulary, tone, clarity, and conciseness)
3. Content ideas and organizations (relevance, accuracy, clarity, originality, coherence, topic sentence, supporting sentence, concluding sentence)
4. Phonological Interference (spelling influenced by pronunciation, voicing contrast like ‘p’ and ‘b’ sounds, minimal pairs like pin and bin)

Students used OpenAI ChatGPT-4 mostly via institutional logins on lab desktop PCs. When a workstation was unavailable or malfunctioned, students used their personal mobile devices to sign in. To preserve privacy, no personal information was uploaded during the sessions, and students referred to prompts using anonymized task IDs.

A two-hour micro-workshop was led by the instructor and covered topics such as prompt design, safety and ethical issues, troubleshooting, and aligning AI outputs with textbook rubrics. Throughout the class, the teacher practiced with model lessons and assisted students as needed.

The preparatory meetings were designed to provide baselines, standards, and readiness for the use of AI tools. Students first finished a diagnostic and baseline phase that comprised a 40-minute pre-test on writing argumentative paragraphs without AI and a quick needs analysis survey that evaluated their device access, confidence, and past AI experience. Following this, students had a 30- to 40-minute orientation and ethics lesson where they were taught what AI can and cannot do, the significance of avoiding copy-paste submissions, and the necessity of revising outputs in their own words while noting the changes and their justifications. Privacy precautions were highlighted, such as using anonymized task codes and refraining from using names, emails, or personal IDs in prompts. Finally, a 20–25-minute micro-training session included a real explanation of the workflow from prompt development to outline, sample generation, critique, and change.

Both groups participated in two weekly classes (each lasting 100 minutes) that included fundamental textbook exercises, and the experimental group also got an extra 50 minutes of ChatGPT-assisted practice that rotated through various paragraph writing components. A 30- to 40-minute model-analysis phase usually started off sessions, which were then followed by 40- to 50-minute guided practice on sentence construction, transitions, and paragraph planning. After that, the experimental group worked independently at lab PCs or mobile devices for 50 minutes while receiving instructor support through a prompted checklist. Meanwhile, the control group completed parallel assignments on paper with teacher comments. The same rubric was used for peer review and editing (30–40 minutes) in both groups. Sessions ended with 15–20-minute exit tickets that included one query and one improvement. Students revised outputs in their own words and documented modifications using ChatGPT for idea development, micro-explanations, alternative phrasing, and paragraph writing. A QuickStart guide, mini-demos for typical problems, and one-on-one coaching to turn ambiguous prompts into smart ones were all included in the support. Topic sentences, supporting sentences, concluding sentences, coherence, argumentative structure,

phrase mechanics, self-editing, and other topics are covered in sample prompts. It is made clear that AI writing must be edited rather than submitted in its original form.

The post-test was given under the identical settings as the pre-test: a 40-minute writing exercise without AI support, assessed using the same rubric. The results revealed a statistically significant improvement in students' writing skills compared to the pre-test. To obtain qualitative insights, a stratified subset of the experimental group took part in 10–15-minute semi-structured interviews. Despite a few small technical issues, students generally had excellent experiences with the intervention, mentioning better writing techniques and confidence. Anonymized IDs were used to keep all AI logs, which were only utilized for the study. Students in the control group spent the same amount of time on teacher-led drills, textbook exercises, and manual drafting/revision as they did on the textbook modules, guided practice, and peer review.

3.5. Validity and Reliability

Achievement Test

The achievement test was constructed, and its final version consisted of four domains: mechanics, language use, content, ideas and organization, and phonological interference.

Test Validity

The test was initially developed and presented to a group of validators and English language specialists from the teaching faculty, totaling 10 validators. Their role was to evaluate the test's suitability, the appropriateness of its questions for their intended purpose, the formulation of questions and domains under it according to test construction standards, the clarity of the linguistic phrasing of test items, and the clarity of test instructions. Based on the validators' feedback, which included deletions, modifications, or additions to items, the final test comprised 4 domains. Here are both versions of the before and after modifications of the validators.

Table 1: Experts' Reviews Of The Test Content.

Stage	Prompt(s)	Domain Assessed
Test before modification	Write a paragraph on an interesting hobby.	Mechanics: Punctuation, capitalization Language Use: grammar, syntax, vocabulary, tone Contents, ideas and organizations: topic sentence, supporting sentence, concluding sentence
Test after modification	Write a paragraph on an interesting hobby. Or How do you like to spend your free time?	Mechanics: Punctuation, spelling, capitalization Language Use: grammar, syntax, vocabulary, tone Contents, ideas and organizations: topic sentence, supporting sentence, concluding sentence, coherence, relevance, clarity, originality, accuracy Phonological Interference: spelling influenced by pronunciation; voicing contrast (e.g., *p* vs. *b*); minimal pairs (e.g., *pin* vs. *bin*)

Potential threats to validity were mitigated by validating group equivalence with pre-tests, employing the same teacher and lesson plans, and assuring equal time-on-task, which reduced internal bias. The single-institution sample limits external validity; therefore, results are interpreted carefully. The four evaluation domains—mechanics, language use, content/organization, and phonological interference—were chosen based on published writing rubrics and expert validation because they mirror the most common challenges Saudi EFL learners have while writing paragraphs. These procedures made certain that the evaluation criteria were both theoretically solid and culturally appropriate.

In addition, the test was administered to a sample outside the study sample, and Pearson correlation coefficients were calculated between each question's score and the total score. Table 2 shows the results.

Table 2: Pearson Correlation Coefficients Between Question Scores And Their Respective Domains And Total Test Scores.

Question	Mechanics	Language Use	Contents, ideas and Organization	Phonological Interference	all1
A1	.819**				.547*
A2	.819**				.663**

A3		.826**			.741**
A4		.841**			.584**
A5			.576**		.619**
A6			.661**		.624**
A7			.661**		.624**
A8			.576**		.619**
A9			.666**		.629**
A10				1.000**	.508*

**. Correlation is significant at the 0.01 level (2-tailed)
*. Correlation is significant at the 0.05 level (2-tailed).

Table 2 shows that the correlation coefficients between questions and their respective domains were statistically significant at the 0.01 level, ranging from (0.576** – 1.00**). The Pearson correlation coefficients between questions and the total test score ranged from (0.508* – 0.741**) and were statistically significant at the 0.05 and 0.01 levels, confirming the test's internal consistency validity. Pearson correlation coefficients were also calculated between the domains and the total score, as shown in Table 3.

Table 3: Pearson Correlation Coefficients Between Domains And Total Test Score.

Domain	Person correlation
Mechanics	.739**
Language Use	.792**
Contents, Ideas and organizations	.973**
Phonological Interference	.508*

Table 3 indicates that the correlation coefficients between domains and the total test score ranged from (0.508* – 0.973**) and were statistically significant at the 0.05 and 0.01 levels, further supporting the test's internal consistency validity.

Test Reliability

The test's reliability was assessed using Cronbach's Alpha, as shown Table 4.

Table 4: Cronbach's Alpha Coefficient.

Domain	Reliability Coefficient
Total Score	0.86

Table 4 shows that the Cronbach's Alpha coefficient reached 0.86, which is a high value, indicating that the test has high reliability.

Test Item Analysis

The test was administered to an exploratory sample of 20 students outside the study sample, and the test questions were analyzed to ensure their validity by calculating difficulty and discrimination coefficients.

Difficulty Coefficients

The difficulty coefficient for objective test items, which yield a score of (zero or one), was measured according to [47] using the following formula:

Difficulty Coefficient = Total Question Scores / (Number of Students * Question Score)

Table 5. shows the difficulty coefficients for the test questions based on the exploratory sample results

Table 5: Difficulty Coefficients For Achievement Test Questions.

Item	Difficulty Coefficient
1	0.65
2	0.65
3	0.65
4	0.55
5	0.60
6	0.65
7	0.65
8	0.70
9	0.50
10	0.65

Table 5 shows that the difficulty coefficients for the test questions ranged between 0.50 – 0.70. According to [47], any item with a difficulty coefficient between 0.20 – 0.80 is considered acceptable and recommended for retention in the test.

The overall average difficulty of the test was calculated to be (0.63), indicating that the test is of moderate difficulty [47].

Discrimination Coefficients

The discrimination coefficients for the objective test items were calculated according to the following formula [47]:

Discrimination Coefficient = (Nu – NI) / N

- **Nu**: Number of students in the upper group who answered the question correctly.
- **NI**: Number of students in the lower group who answered the question correctly.

- **N**: Number of students in one of the two groups.

To calculate the discrimination coefficient for each question, the students were divided into two groups: an upper group consisting of 50% of the students with the highest scores and a lower group consisting of 50% of the students with the lowest scores (10 students in the upper group and 10 in the lower group). According to [47], specialists in measurement have established reference values for judging test items as follows:

- Items with a negative discrimination coefficient are discarded.
- Items with a discrimination coefficient less than 0.20 are recommended for removal.
- Items with a discrimination coefficient of 0.20 or higher are accepted.

Table 6 shows the discrimination coefficients for each test item.

Table 6. Discrimination Coefficients For Achievement Test Questions.

Item	Discrimination Coefficient
1	0.50
2	0.30
3	0.50
4	0.50
5	0.40
6	0.30
7	0.30
8	0.40
9	0.40
10	0.50

Table 6 shows that the discrimination coefficients for the test items ranged between 0.30 – 0.50, and thus all test questions were accepted as they fell within the acceptable discrimination range.

Normality of Distribution

The Kolmogorov-Smirnov test was used to verify the normality of the distribution of the study sample's scores on the test in both the pre-test and post-test, as shown in Table 7.

Table 7: Kolmogorov-Smirnov Test for Normality of Distribution of Study Sample Scores in Pre-Test and Post-Test.

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Group	Statistic	Df	Sig.	Statistic	Df	Sig.
Pre	Control	.119	30	.200*	.968	30	.478
	Experimental	.106	30	.200*	.974	30	.656
Post	Control	.140	30	.137	.945	30	.127
	Experimental	.152	30	.67	.937	30	.112

Table 7 shows that the scores follow a normal distribution, so parametric tests were used.

3.6. Data Analysis

The data were analyzed using the Statistical Package for the Social Sciences (SPSS) to answer the study questions and ensure the accuracy and reliability of the results. Several statistical procedures were employed. First, the Pearson correlation coefficient was applied to verify the internal consistency validity of the study instrument, ensuring that items measuring the same construct were correlated. To further establish reliability, Cronbach's Alpha was calculated, providing a measure of the internal consistency of the instrument and confirming its stability across items. The Kolmogorov-Smirnov test was used to examine the normality of the data distribution, which is a prerequisite for applying parametric tests. For the main analysis, the t-test for independent samples was conducted to determine the significance of the differences between the mean scores of the control and experimental groups in both the pre-test and the post-test, thus assessing the effectiveness of the intervention. Finally, Eta squared was calculated as a measure of effect size, indicating the practical significance and strength of the observed differences between groups.

4. STUDY RESULTS

4.1. Research Question 1: How effective is the use of ChatGPT-mediated instruction in enhancing EFL students' paragraph writing skills?

Table 8 presents the results of the independent samples t-test conducted to compare the pre-test performance of the control and experimental groups across the four domains of writing assessment: mechanics, language use, content/organization, and phonological interference, as well as the overall total score. The results indicate that there were no statistically significant differences at the 0.05 level between the two groups in any of the assessed domains, confirming that the groups were equivalent before the intervention. In mechanics, the control group ($M = 0.20$, $SD = 0.153$) and the experimental group ($M = 0.19$, $SD = 0.142$) showed almost identical performance, $t(58) = 0.219$, $p = .827$. Similarly, for language use, the control group ($M = 0.50$, $SD = 0.394$) and the experimental group

(M = 0.53, SD = 0.414) demonstrated no significant difference, t(58) = -0.320, p = .750, indicating comparable grammatical knowledge at the outset. In terms of content, ideas, and organization, both groups achieved nearly the same mean scores, with the control group at (M = 1.00, SD = 0.881) and the experimental group at (M = 0.93, SD = 0.716), t(58) = 0.322, p = .749. For phonological interference, the mean scores were identical across both groups (M = 0.25, SD = 0.354), t(58) = 0.000, p = 1.000, reflecting equal influence of first-language interference in the pre-test. Finally, when the total scores were compared, the control group (M = 1.95, SD = 1.057) and the experimental group (M = 1.91, SD = 0.957) did not differ significantly, t(58) = 0.160, p = .873. Overall, these findings confirm that the two groups were statistically equivalent in their writing abilities at the beginning of the study, thereby establishing a valid baseline for evaluating the effectiveness of ChatGPT-mediated instruction in the post-test phase.

Table 8: T-Test For Independent Samples To Determine Significance Of Differences Between Mean Scores Of Control And Experimental Groups In Pre-Test.

Domain	Group	N	Mean	Std. Deviation	T	Df	Sig. (2-tailed)
Mechanics	Control	30	.20	.153	.219	58	.827
	Experimental	30	.19	.142			
Language use	Control	30	.50	.394	-.320	58	.750
	Experimental	30	.53	.414			
Contents, ideas and organizations	Control	30	1.00	.881	.322	58	.749
	Experimental	30	.93	.716			
Phonological Interference	Control	30	.25	.354	.000	58	1.000
	Experimental	30	.25	.354			
Total	Control	30	1.95	1.057	.160	58	.873
	Experimental	30	1.91	.957			

Table 9 presents the results of the independent samples t-test comparing the post-test performance of the control and experimental groups across the four writing domains and the overall total score. The analysis shows statistically significant differences at the 0.05 level in all domains, consistently in favor of the experimental group that received ChatGPT-mediated instruction. In the domain of mechanics, the control group (M = 0.73, SD = 0.262) outperformed the experimental group (M = 0.48, SD = 0.217), and the difference was statistically significant, t(58) = 4.024, p = .000, with a large effect size (η^2 = .218). For language use, the control group (M = 1.15, SD = 0.368) scored higher than the experimental group (M = 0.82, SD = 0.404), with t(58) = 3.304, p = .002, and a large effect size (η^2 = .158), reflecting substantial improvement attributable to the treatment. In the domain of content, ideas, and organization, the most pronounced difference was observed: the control group (M = 3.10, SD = 0.960) scored significantly higher than the experimental group (M = 1.47, SD = 0.742), t(58) = 7.375, p = .000, with an exceptionally large effect size (η^2 = .484), indicating that ChatGPT-mediated instruction had a very strong impact on learners' ability to generate, organize, and develop ideas in writing. Similarly, for phonological interference, the control group (M = 0.59, SD = 0.290) outperformed the experimental group (M = 0.31, SD = 0.126), t(58) = 4.911, p = .000, with a large effect size (η^2 = .294), suggesting a marked reduction in first-language interference errors among students exposed to ChatGPT support. Finally, the total mean score of the control group (M = 5.57, SD = 1.273) was significantly higher than that of the experimental group (M = 3.08, SD = 1.049), t(58) = 8.293, p = .000, with a very large effect size (η^2 = .542). Collectively, these results clearly demonstrate that the experimental group showed substantial gains across all domains of writing after the intervention, establishing the effectiveness of ChatGPT-mediated instruction in enhancing EFL learners' paragraph writing performance.

Table 9: T-Test For Independent Samples To Determine Significance Of Differences Between Mean Scores Of Control And Experimental Groups In Post-Test.

Domain	Group	N	Mean	Std. Deviation	t	Df	Sig. (2-tailed)	Eta Squared	Level
Mechanics	Control	30	.73	.262	4.024	58	.000	.218	Large
	Experi	30	.48	.217					

	ment								
Language use	Control	30	1.15	.368	3.304	58	.002	.158	Large
	Experiment	30	.82	.404					
Content, ideas and organizations	Control	30	3.10	.960	7.375	58	.000	.484	Large
	Experiment	30	1.47	.742					
Phonological Interference	Control	30	.59	.290	4.911	58	.000	.294	Large
	Experiment	30	.31	.126					
Total	Control	30	5.57	1.273	8.293	58	.000	.542	Large
	Experiment	30	3.08	1.049					

4.2. Research Question 2: What are EFL students' experiences of using ChatGPT as a learning aid in paragraph writing?

Experiences

The analysis of students' interviews revealed several themes reflecting their experiences of using ChatGPT to enhance paragraph writing.

Students reported that ChatGPT significantly supported them in improving the mechanics of writing. They explained that the tool helped them correct spelling, grammar, and capitalization, which resulted in more accurate and readable paragraphs. For instance, one student stated, *"My spelling was improved through the ChatGPT"* (S6), while another remarked, *"I learnt the spelling of difficult words through ChatGPT"* (S8). These excerpts show that students relied on the tool to reinforce technical aspects of writing.

With regard to language use, students expressed that ChatGPT enriched their writing by providing useful grammatical structures, vocabulary, and expressions. They highlighted that it helped them grasp grammar rules more clearly and apply them properly in their writing. One student observed, *"ChatGPT helps me improve grammar"* (S1), while another added, *"ChatGPT helps me learn grammatical structures"* (S5). Students also emphasized its usefulness for vocabulary development, with one explaining, *"ChatGPT is very rich in providing automated vocabulary that helps learn new words and expressions"* (S15).

Students also emphasized that ChatGPT played an important role in supporting content development, idea generation, and organization. They found that the tool guided them in creating meaningful content, arranging their ideas logically, and structuring paragraphs coherently. For example, a student noted, *"ChatGPT helps me write better content for a paragraph"* (S7), while another shared, *"ChatGPT makes it easy to organize my ideas to produce a good paragraph"* (S9). Similarly, another student highlighted its role in coherence, explaining, *"ChatGPT helped me improve the sequence of sentences correctly"* (S13).

Another important theme that emerged was related to phonological interference. Some students noted that ChatGPT helped them overcome errors caused by similarities in sounds that often affect spelling and word choice. For instance, one student explained, *"AI tool helped me to learn the difference between 'p' and 'b', 't' and 'd', 'k', and 'g', etc."* (S14). Similarly, another student emphasized, *"ChatGPT... helps learn new words and expressions and differentiate among different sounds"* (S15). This indicates that ChatGPT supported learners in minimizing pronunciation-based spelling mistakes.

Observations

Teachers' and students' observations also provided valuable insights. It was noted that ChatGPT increased students' involvement and interest in writing, making the learning process more engaging. As one student put it, *"The AI tool is very fascinating"* (S7). In addition, ChatGPT was seen as highly accessible, with another describing it as *"very user-friendly"* (S8).

In terms of its value as an educational instrument, both students and instructors recognized ChatGPT as an effective learning aid that could address many of the challenges students face in paragraph writing. For example, one student stated, *"I think ChatGPT is a good learning tool in writing paragraphs"* (S2), while another emphasized, *"I believe ChatGPT can solve many paragraph writing issues"* (S9).

At the same time, instructors underlined the need for teacher support to ensure that students gain the maximum benefit from the tool. One participant explained, *"I believe ChatGPT can be good if used regularly under teacher support"* (S12).

This reflects the view that ChatGPT should be integrated into learning with proper guidance and continuous practice.

Finally, the use of ChatGPT was found to enhance students' motivation and self-assurance. Teachers observed that the tool allowed learners to face their writing challenges with greater confidence. As one student expressed, *"ChatGPT is good to build confidence and face writing challenges"* (S15). This shows that beyond improving writing skills, the tool also nurtured a sense of empowerment among learners.

Overall, ChatGPT was perceived as a valuable, supportive, and motivating tool that improved the technical, linguistic, and organizational aspects of students' paragraph writing, while also boosting their confidence and engagement in the learning process.

5. DISCUSSION

This study aimed to assess the effectiveness of ChatGPT-mediated instruction in developing paragraph writing skills. The quantitative and qualitative analysis of this study highlight the effectiveness of the ChatGPT-mediated instruction in developing paragraph writing skills, specifically mechanics, language use, contents and organization, and phonological interference. Interviews revealed that the students had no prior formal experience with ChatGPT. This explains why the two groups had equal baseline performance. At the start of the study, the control and experimental groups had similar English proficiency levels and pretest scores, which validates the study. Table 7 contrasts two groups, each consisting of thirty participants, in four domains: phonological interference, content and organization, language use, mechanics, and total. There are not many group-to-group variations in the mean scores and standard deviations. Similar standard deviations and slight mean differences imply that there are no appreciable differences in the performance between the groups in any of the domains.

This baseline finding was triangulated with the qualitative interviews, where students themselves acknowledged that they had not used ChatGPT before the intervention. This confirmed why their pretest results were almost the same and why no initial advantage was observed in the experimental group.

The study's results are consistent with those of another study by [33], who first looked at the descriptive statistics of the participants' general writing, writing proficiency, and writing motivation scores before and after the test in both groups. EFL undergraduate sample (B2 level), experimental vs. control over six weeks (n=33), the mean, median, and mode were all quite near to one another, primarily around the value of 3. The study shows that groups were comparable at baseline (no significant difference in performance) and reports pretest averages.

The study findings are partially consistent with another study by [34], which found that the AI experimental group's post-test scores were significantly higher than the control group's pre-test and post-test scores, as well as the experimental group's pre-test scores. There was no significant difference in pre-test and post-test results for the control or experimental groups. The results might be since the students have almost similar levels of proficiency at the beginning level as the classes start in the semester. The second reason might be owed to the fact that students had no exposure to the AI tool before the intervention, so their performance in the pre-test was almost the same. The researchers could not find any study that could contrast the findings of the study. Triangulating the quantitative and qualitative evidence shows that both strands confirm no baseline difference across groups, thus demonstrating the objectivity of the experiment.

The post-test results showed that the experimental group outperformed the control group in all domains, including mechanics, language use, contents, ideas, and organization, and phonological interference, with large effect sizes indicating that the improvement was both statistically significant and educationally meaningful. The qualitative interviews strongly corroborated these quantitative results: students reported that ChatGPT helped them improve grammar, spelling, vocabulary, content generation, and the ability to arrange and sequence ideas efficiently. They also explained how ChatGPT reduced confusion in phonological contrasts such as "p/b," "t/d," and "k/g", which had previously influenced their spelling errors. This direct testimony explains why the experimental group demonstrated higher post-test scores across all domains. These qualitative findings imply that the post-test improvement was directly related to students' active involvement with ChatGPT as a scaffold for writing practice.

The findings accord strongly with several previous studies. Mahapatra [10] reported that ChatGPT's suggestions for correcting grammatical errors and improving syntax

enhanced the clarity and professionalism of their paragraphs. The findings of the study are in line with another study by [35] that found that ChatGPT significantly improved writing mechanics, including grammar, punctuation, and spelling, through its grammar-checking and proofreading capabilities. Participants demonstrated enhanced accuracy in paragraph writing after using ChatGPT. This is also triangulated with the current qualitative interviews, where learners explicitly mentioned spelling and grammar as the areas where ChatGPT offered the most help. The study emphasizes that ChatGPT's real-time corrections helped students identify and rectify mechanical errors, boosting their confidence in producing polished paragraphs.

The findings of the study support another study by [28] that found that ChatGPT significantly enhanced language use in paragraph writing. Pre- and post-test analyses showed improvements in vocabulary diversity, sentence complexity, and appropriate tone. Students reported that ChatGPT helped them select more precise words and vary sentence structures, leading to more engaging and polished paragraphs. Similarly, in the present study, students' reflections (S15) highlighted that ChatGPT enriched their vocabulary and helped them choose suitable words and expressions, aligning with these post-test improvements.

The findings of the study align with another study by [36] that asserts that the use of ChatGPT showed significant improvements in content quality and organization, with students producing more robust arguments and better-structured paragraphs. The findings of the study support another study BY [25] that found that ChatGPT was particularly effective in improving language use for formal writing, offering suggestions for academic vocabulary and complex sentence structures. Students reported that ChatGPT helped them refine word choice and maintain a consistent formal tone, enhancing the quality of their paragraphs. This triangulates with the qualitative evidence in which students (S7, S9, S13) indicated that ChatGPT guided them in developing content, organizing ideas logically, and sequencing sentences effectively.

The findings of the study align with another study by Li et al. (2024), which found that ChatGPT particularly helped students in identifying and correcting errors in word usage, which could include phonologically influenced errors. The findings of the study align with another study by [34], which found that ChatGPT was helpful for refining ideas and correcting language errors, particularly those related to vocabulary misuse. These quantitative outcomes converge with students' reflections (S14, S15) about overcoming phonological interference and learning to distinguish confusing sound pairs, which improved their word usage. The study highlights ChatGPT's role in enhancing self-efficacy and engagement in writing, which could motivate students to address errors like those caused by phonological interference.

The reason might be owed to the fact that the AI-mediated tool ChatGPT has a lot of features that can assist students in numerous ways to improve their paragraph writing skills. Another reason for ChatGPT to be highly effective is that students' interest in technology and electronic gadgets, especially smartphones, is high.

However, the findings of the study contradict a study by [37] that suggests that over-reliance on ChatGPT for academic tasks, such as generating content and organizing paragraphs, can lead to "memory loss" and weakened cognitive functioning, including critical thinking and idea generation. Continuous dependence on AI tools may foster laziness, reducing students' ability to independently develop and organize ideas in paragraph writing. The findings of the study contradict another study by [38] that found that students using AI tools like ChatGPT reported a loss of pride and ownership in their writing, which could hinder their ability to generate original ideas and organize paragraphs independently. The article emphasizes that writing is a process involving critical thinking, revision, and dialogue, which ChatGPT's instant text generation bypasses. This can lead to underdeveloped literacy skills, particularly in crafting coherent and original content, as students may rely on AI to produce polished paragraphs without engaging in the creative or organizational process. The findings of the study oppose the findings of another study by [39] that found that ChatGPT's efficiency in generating content and organizing text may prevent students from engaging in critical thinking, problem-solving, and imaginative processes essential for writing. Specifically, the study suggests that students relying on ChatGPT for paragraph writing may produce work that appears polished but lacks depth in original thought and structural creativity, as they bypass the iterative process of drafting and revising. The reasons for this contrast may be that some students do not use ChatGPT as an assisting tool but rather want everything to be done by

ChatGPT. In that sense, they become fully dependent on ChatGPT. By contrast, the qualitative evidence from this study showed that when guided properly, students used ChatGPT as a supportive aid rather than a replacement, which explains why their post-test scores indicated genuine skill improvement.

Taken together, the triangulated evidence strongly supports the view that ChatGPT significantly enhances EFL learners' mechanics, grammar, and vocabulary.

The findings on EFL students' perceptions and experiences with ChatGPT as a learning tool in the writing classroom show a high convergence of quantitative, qualitative, and literature-based evidence. Quantitative post-test findings showed considerable improvement in paragraph writing, notably in mechanics, grammar, vocabulary use, and content organization, indicating that ChatGPT had a meaningful impact on writing performance. This was triangulated with students' qualitative reflections, where they confirmed that ChatGPT helped them produce relevant ideas, arrange supporting facts, improve spelling and grammar, and link sentences coherently. Students also reported feeling more confident and less anxious when composing paragraphs, albeit a few acknowledged minor technical challenges. These findings are consistent with previous research: Jeyraman et al. [40] and Yang [24] found that ChatGPT improved vocabulary, grammar accuracy, logical sequencing, and writing confidence, while Meniado et al. [41] emphasized its function in brainstorming, refining outlines, and editing drafts. Similarly, Capah et al. [42] discovered that ChatGPT improved idea formation, content organization, and phonological-lexical control. The similarities between studies and current results are most likely due to common EFL student issues in grammar, organization, and vocabulary, which AI solutions directly address.

However, discrepancies emerged: Yoon et al. [43] observed that ChatGPT input was often generic and less useful, while Aksakallı and Daşer [44] cautioned that over-reliance may decrease creativity and critical thinking skills. These disparities may reflect variances in students' prompting abilities, teacher scaffolding, or contextual variables in AI use. The triangulated evidence shows that, when accompanied by teacher direction, ChatGPT considerably enhances EFL learners' writing outcomes and confidence; nevertheless, careful integration is required to avoid dependency and foster critical engagement.

5. CONCLUSION

The current study investigated how well ChatGPT-mediated instruction at Najran University developed the paragraph-writing skills of EFL students. The experimental group showed notable gains in all assessed domains: mechanics, language use, content and organization, and phonological interference—after the intervention, according to the results, even though there were no statistically significant differences in the pre-intervention pre-test scores between the control and experimental groups. These findings support the notion that, when properly overseen and supervised by the instructor, incorporating ChatGPT into writing lessons can improve students' ability to write paragraphs and encourage their participation.

This study is distinctive in that it examines the impact of ChatGPT on paragraph writing skills among EFL learners at Najran University, a previously unexplored context, with a focus on four essential domains: mechanics, language use, content organization, and phonological interference. Unlike previous research that focused on general academic writing or paraphrasing [2], [10], this study's mixed-methods approach, which combines quasi-experimental data with qualitative learner insights, provides a comprehensive evaluation that goes beyond perception-based or solely quantitative studies [8], [24]. Its significance includes expanding Computer-Assisted Language Learning (CALL) by providing empirical evidence for AI-enhanced writing pedagogy as well as practical suggestions for EFL instructors in Saudi Arabia on how to improve student correctness, coherence, and confidence.

Despite these promising findings, a few limitations should be acknowledged. The study employed a small, homogeneous sample of 60 male students from a single university, limiting the findings' generalizability. To improve external validity, future studies should increase the sample size to include both male and female students from diverse colleges and skill levels. The three-month intervention period may not adequately reflect the long-term effects of ChatGPT on independent writing ability. Longitudinal studies that last several semesters or years might be more effective at capturing long-term effects.

Furthermore, the quasi-experimental method controls a wide variety of factors but does not eradicate selection bias. Future research could use randomized controlled trials (RCTs) or matched-group designs to reduce this risk. Furthermore, the study only employed one sort of writing prompt for testing, which may not accurately reflect performance across genres or tasks. Future research should use a variety of tasks, such as essays, tales, and reports, to investigate genre-specific consequences.

The study did not quantify the potential negative repercussions of over-reliance on AI, such as decreased originality or critical thinking skills, in addition to qualitative findings. Future research should include both quantitative and qualitative measurements of creativity, critical thinking, and learner autonomy to assess potential downsides and benefits.

Based on the findings, the study suggests that EFL instructors should use ChatGPT as a guided tool, providing explicit training on AI use, ethical use, and critical evaluation. Assessment rubrics should include both AI-assisted and independent tasks. Institutions should develop policies promoting responsible AI use while addressing plagiarism, data privacy, and academic integrity. Future research should explore ChatGPT's long-term effects on writing proficiency, creativity, and critical thinking, and its genre-specific applications.

This study advances EFL writing pedagogy by investigating ChatGPT's impact on paragraph writing skills among Saudi students at Najran University, a hitherto unexplored environment. Unlike previous research that focused on general academic writing or paraphrasing, this work takes a strong mixed-methods approach to target paragraph-specific skills—mechanics, language use, content organization, and phonological interference. By combining experimental data and vivid learner insights, it provides strong proof of ChatGPT's transformative potential, outperforming studies that rely exclusively on perceptions or quantitative metrics. The findings promote Computer-Assisted Language Learning (CALL) and provide EFL instructors with practical ways for increasing student accuracy, coherence, and confidence in Saudi Arabia.

Despite its single-institution emphasis, which restricts generalizability, this work establishes a solid framework for future research into varied genres and long-term AI integration, resulting in sustainable, ethical advances in writing teaching.

ACKNOWLEDGMENT

The authors are thankful to the Deanship of Graduate Studies and Scientific Research at Najran University for funding this work under the Growth Funding Program grant code (NU/GP/SEHRC/13/182-2).

REFERENCES:

[1] Y. W. Pratama, "The investigation of using Grammarly as online grammar checker in the process of writing," *Journal of English Language Education*, vol. 1, no. 1, 2020, pp. 46–54. https://journal.unsika.ac.id/index.php/IDEAS/article/view/4180

[2] T. Adiguzel, M. H. Kaya, and F. K. Cansu, "Revolutionizing education with AI: Exploring the transformative potential of ChatGPT," *Contemporary Educational Technology*, vol. 15, no. 3, 2023, pp. 1–13. https://doi.org/10.30935/cedtech/13152

[3] T. S. Heathen and D. E. Lin, "A review on the perks of using ChatGPT in education," Preprints, 2024. https://doi.org/10.20944/preprints202406.1060.v1

[4] N. Fitria, C. M. Zahra, M. Usman, and B. Usman, "Exploring strategies to foster cohesion and coherence in EFL writing: A focus on simple view of writing," *Jurnal Sociohumaniora Kodepena (JSK)*, vol. 5, no. 1, 2024, pp. 62–73. https://doi.org/10.54423/jsk.v5i1.180

[5] I. Maakip, O. O. Matanluk, and S. Marshall, "Understanding student perception regarding the use of ChatGPT in their argumentative writing: A qualitative inquiry," *Malaysian Journal of Communication*, vol. 39, no. 4, 2023, pp. 150–165. https://doi.org/10.17576/JKMJC-2023-3904-08

[6] T. Susilowati and A. Fachriza, "Enhancing students' ability in composing a well-developed paragraph through social constructive learning facilitated by peer-assessment," *ELTALL: English Language Teaching, Applied Linguistic and Literature*, vol. 5, no. 1, 2024, pp. 1–13. https://doi.org/10.21154/eltall.v5i1.8738

[7] M. L. Ng, A. Al Ghaithi, and B. Behforouz, "The effect of Gemini on EFL learners' writing skills and motivation," *Journal Name*, vol. X, no. Y, 2025, pp. Z.

[8] A. Bouzar, K. el Idrissi, and T. Ghourdou, "ChatGPT and academic writing self-efficacy: Unveiling correlations and technological dependency among postgraduates," *Arab World English Journal*, vol. 1, 2024, pp. 225–236. https://doi.org/10.24093/awej/ChatGPT.15

[9] S. E. M. Amanda, N. Lubis, and U. Dewi, "Quillbot as an AI-powered English writing assistant: An alternative for students to write English," *Jurnal Pendidikan dan Sastra Inggris*, vol. 3, no. 2, 2023, pp. 188–199. https://doi.org/10.55606/jupensi.v3i2.2026

[10] S. Mahapatra, "Impact of ChatGPT on ESL students' academic writing skills: A mixed methods intervention study," *Smart Learning Environments*, vol. 11, no. 1, 2024, p. 9.

[11] H. C. Chui, "ChatGPT as a tool for developing paraphrasing skills among ESL learners," *Journal of Practices in Language Learning and Teaching*, vol. 11, no. 2, 2023, pp. 85–105. https://doi.org/10.24191/cplt.v11i2.21723

[12] M. Perkins, "Academic integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond," *Journal of University Teaching & Learning Practice*, vol. 20, no. 2, 2023. https://doi.org/10.53761/1.20.02.07

[13] M. Imran and N. Almusharraf, "Analyzing the role of ChatGPT as a writing assistant at higher education level: A systematic review of the literature," *Contemporary Educational Technology*, vol. 15, no. 4, 2023, ep464. https://doi.org/10.30935/cedtech/13605

[14] C. A. Chapelle, *Computer applications in second language acquisition: Foundations for teaching, testing and research*. Cambridge University Press, 2001.

[15] B. Zou, L. Huang, and H. Xie, "Artificial intelligence in language education: A review of ChatGPT in EFL writing instruction," *Computer Assisted Language Learning*, 2023, Advance online publication. https://doi.org/10.1080/09588221.2023.2237640

[16] R. Hampel and U. Stickler, "New skills for new classrooms: Training tutors to teach languages online," *Computer Assisted Language Learning*, vol. 18, no. 4, 2005, pp. 311–326. https://doi.org/10.1080/09588220500335455

[17] M. Warschauer, "Computer-assisted language learning: An introduction," in *Multimedia language teaching*, S. Fotos, Ed. Logos International, 1996, pp. 3–20.

[18] A. A. Khuwaileh and A. A. Shoumali, "Writing errors: A study of the writing ability of Arab learners of academic English and Arabic at university," *Language, Culture and Curriculum*, vol. 13, no. 2, 2000, pp. 174–183.

[19] K. Lakkis and M. Abdel Malak, "Understanding the transfer of prepositions: Arabic to English," in *Investigating writing problems among Palestinian students studying English as a foreign language*, K. Mourtaga, Ed. (Unpublished doctoral dissertation), 2000.

[20] A. O. Strauch, *Writers at work: The short composition*. Cambridge University Press, 2005.

[21] J. M. Gayed, M. K. J. Carlon, A. M. Oriola, and J. S. Cross, "Exploring an AI-based writing assistant's impact on English language learners," *Computers and Education: Artificial Intelligence*, vol. 3, 2022, 100055. https://doi.org/10.1016/j.caeai.2022.100055

[22] G. Z. Liu, M. Rahimi, and J. Fathi, "Flipping writing metacognitive strategies and writing skills in an English as a foreign language collaborative writing context: A mixed-methods study," *Journal of Computer Assisted Learning*, vol. 38, 2022, pp. 1730–1751. https://doi.org/10.1111/jcal.12707

[23] D. Ippolito, A. Yuan, A. Coenen, and S. Burnam, "Creative writing with an AI-powered writing assistant: Perspectives from professional writers," *arXiv*, 2022, pp. 1–18. https://doi.org/10.48550/arXiv.2211.05030

[24] D. Yan, "Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation," *Educational Information Technology*, 2023. https://doi.org/10.1007/s10639-023-11742-4

[25] Y. Su, Y. Lin, and C. Lai, "Collaborating with ChatGPT in argumentative writing classrooms," *Assessing Writing*, vol. 57,

2023, 100752. https://doi.org/10.1016/j.asw.2023.100752

[26] N. P. Özçelik and G. Yangın Ekşi, "Cultivating writing skills: The role of ChatGPT as a learning assistant—a case study," *Smart Learning Environments*, vol. 11, no. 1, 2024, p. 10.

[27] M. I. M. Abdi, B. D. B. Ananta, and L. H. Nurbatra, "Exploring the use of ChatGPT in improving students' paragraph composition skills," *Focus on ELT Journal*, vol. 6, no. 4, 2024, pp. 17–36.

[28] C. Song and Y. Song, "Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students," *Frontiers in Psychology*, vol. 14, 2023, 1260843.

[29] M. Abdullayeva and Z. M. Musayeva, "The impact of ChatGPT on students' writing skills: An exploration of AI-assisted writing tools," *Proceedings of the International Conference on Educational Research and Innovation*, vol. 4, 2023, pp. 61–66.

[30] N. Nazari, M. S. Shabbir, and R. Setiawan, "Application of artificial intelligence powered digital writing assistant in higher education: Randomized controlled trial," *Heliyon*, vol. 7, no. 5, 2021.

[31] R. R. Divekar, J. Drozdal, S. Chabot, Y. Zhou, H. Su, Y. Chen, et al., "Foreign language acquisition via artificial intelligence and extended reality: Design and evaluation," *Computer Assisted Language Learning*, vol. 35, 2021, pp. 2332–2360. https://doi.org/10.1080/09588221.2021.1879162

[32] A. Savage and P. Mayer, *Effective academic writing 2*, 2nd ed. Oxford University Press, 2012.

[33] E. Pipia and T. Gurgenishvili, "The impact of ChatGPT on English as a foreign language learners' writing skills—An experimental study at Georgian universities," *ESI Preprints (European Scientific Journal, ESJ)*, vol. 39, 2025, p. 429.

[34] M. F. Teng, ""ChatGPT is the companion, not enemies": EFL learners' perceptions and experiences in using ChatGPT for feedback in writing," *Computers and Education: Artificial Intelligence*, vol. 7, 2024, 100270.

[35] M. A. A. Alkamel and N. A. S. Alwagieh, "Utilizing an adaptable artificial intelligence writing tool (ChatGPT) to enhance academic writing skills among Yemeni university EFL students," *Social Sciences & Humanities Open*, vol. 10, 2024, 101095.

[36] D. Khampusaen, "The impact of ChatGPT on academic writing skills and knowledge: An investigation of its use in argumentative essays," *LEARN Journal: Language Education and Acquisition Research Network*, vol. 18, no. 1, 2025, pp. 963–988.

[37] M. Abbas, F. A. Jam, and T. I. Khan, "Is it harmful or helpful? Examining the causes and consequences of generative AI usage among university students," *International Journal of Educational Technology in Higher Education*, vol. 21, no. 1, 2024, p. 10.

[38] N. S. Baron, "How ChatGPT robs students of motivation to write and think for themselves," *The Conversation*, vol. 19, 2023.

[39] T. Rasul, S. Nair, D. Kalendra, M. Robin, F. de Oliveira Santini, W. J. Ladeira, et al., "The role of ChatGPT in higher education: Benefits, challenges, and future research directions," *Journal of Applied Learning and Teaching*, vol. 6, no. 1, 2023, pp. 41–56.

[40] J. Jeyraman, M. Md Yunus, and N. E. Mohd Said, "ChatGPT for ESL writing: A case study on Year 6 ESL pupils' perceptions," *International Journal of Academic Research in Progressive Education and Development*, 2024. https://doi.org/10.6007/IJARPED/v14i1/24831

[41] J. C. Meniado, D. T. T. Huyen, N. Panyadilokpong, and P. Lertkomolwit, "Using ChatGPT for second language writing: Experiences and perceptions of EFL learners in Thailand and Vietnam," *Computers and Education: Artificial Intelligence*, vol. 7, 2024, 100313.

[42] C. R. Capah, M. Abrar, L. F. Prapassel, and B. S. T. Sukarno, "The role of ChatGPT in EFL writing: Students' perspective," *Journal of English Language and Education*, vol. 10, no. 3, 2025, pp. 703–710.

[43] S. Y. Yoon, E. Miszoglad, and L. R. Pierce, "Evaluation of ChatGPT feedback on ELL writers' coherence and cohesion," *arXiv preprint arXiv:2310.06505*, 2023.

[44] C. Aksakallı and Z. Daşer, "Unlocking EFL learners' insights into ChatGPT use for L2 writing: The impacts of usage frequency and

gender variations," *Current Psychology*, vol. 44, no. 9, 2025, pp. 7957–7977.

[45] T. Mohamad, M. Nazim, A. A. F. Alzubi, and S. I. Khan, "Examining EFL students' motivation level in using QuillBot to improve paraphrasing skills," *World Journal of English Language*, vol. 14, no. 1, 2024.

[46] V. Braun and V. Clarke, "Using thematic analysis in psychology," *Qualitative research in psychology*, vol. 3, no. 2, 2006, pp. 77-101.

[47] R. M. Alareifi, "Analysis of MQs in Summative Exam in English: Difficulty Index, Discrimination Index and Relationship between them," *Journal of Educational and Human Sciences*, no. 20, 2023, pp. 124-135. https://doi.org/10.33193/JEAHS.20.2023.325