www.ijte.net

Investigating the Impact of ChatGPT on Reading Proficiency in EFL Context

Muhammad Mooneeb Ali
Higher Education Department, Government of the Punjab, Pakistan

Ahmed M. Alaa
Badr University in Cairo, Egypt

Wael Alharbi
Yanbu English Language and Preparatory Year Institute, Saudi Arabia

Issa Al Qurashi
Taif University, Saudi Arabia

To cite this article:

Ali, M.M., Alaa, A.M., Alharbi, W., & AlQurashi, I. (2025). Investigating the impact of ChatGPT on reading proficiency in EFL context. *International Journal of Technology in Education (IJTE), 8*(4), 1222-1244. https://doi.org/10.46328/ijte.1563

The International Journal of Technology in Education (IJTE) is a peer-reviewed scholarly online journal. This article may be used for research, teaching, and private study purposes. Authors alone are responsible for the contents of their articles. The journal owns the copyright of the articles. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of the research material. All authors are requested to disclose any actual or potential conflict of interest including any financial, personal or other relationships with other people or organizations regarding the submitted work.

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

International Journal of Technology in Education
2025, Vol. 8, No. 4, 1222-1244 https://doi.org/10.46328/ijte.1563

Investigating the Impact of ChatGPT on Reading Proficiency in EFL Context

Muhammad Mooneeb Ali, Ahmed M. Alaa, Wael Alharbi, Issa AlQurashi

Article Info

Article History

Received:
11 February 2025
Accepted:
7 June 2025

Keywords

Artificial Intelligence
ChatGPT
EFL learners
Reading comprehension
Saudi Arabia

Abstract

Machine and prompt-based Artificial Intelligence (AI) learning has made significant evolution profusely. In education, it has revitalized researchers and educators to scout out subsequent advantages for optimizing learning results. Chiefly, Generative AI has exhibited substantial potential as a tool for language augmentation. This study aims to probe how ChatGPT, a learning model of AI, can be employed as a tool for refining the reading skills of English as a Foreign Language (EFL) students in the Kingdom of Saudi Arabia (KSA). The study sample was one hundred and twenty 10th grade learners of public sector schools in KSA. Ahead of experiment, a preliminary test was taken from the sample, and in accordance with their test scores, two groups of 60 students were formed. One group studied by commonly used classroom methods, while the other learned by exploiting ChatGPT in conjunction with their teacher. The intercession span was three weeks, after that, a final test was administered and the scores were compared. Fifty students were also interviewed to share their insights regarding the use of ChatGPT in their learning. The results displayed a vivid increase in the scores of learners who used ChatGPT. The students also expressed overall positive feedback about learning using ChatGPT. This specifies that ChatGPT is an efficacious and impactful AI learning tool for developing reading proficiency in EFL settings.

Introduction

Technology is a channel for vesting learners and educators (Ali, Alaa, & Shahnaz, 2024). Lately, the coalescing of Artificial Intelligence (AI) has reorganized varied genres in general, and academia in particular. By transforming traditional practices, it has yielded striking advancements. The inception of AI into education has witnessed proliferous progression, thus glinting interest among researchers and practitioners on AI's potential to enhance learning outcomes. Generative AI, in precise, has arose as a promising tool for language enhancement, offering innovative solutions to traditional educational challenges (Ali et al., 2023; Espartinez, 2024; Faisal, 2024).

The swift expansion and application of machine and prompt-based learning, allied with AI, have opened many corridors of novel methodologies in language instruction (Foung, Lin, & Chen, 2024; Huesca et al., 2024). These progressions are also indistinguishable in the context of teaching English as a Foreign Language (EFL), where

efficient and influenceable teaching tools can attract numerous educators and students (Li et al., 2024; Rahimi & Sevilla-Pavón, 2024; Zheng & Steward, 2024).ChatGPT, a sophisticated AI language model, represents one of the forefront technologies in this exploration. Its ability to generate human-like text and engage in interactive dialogue can present unique opportunities for enhancing English language skills (Al Sawi & Alaa, 2024).

The Saudi Arabian English Context

The Kingdom of Saudi Arabia (KSA) does not give English the status of a neutral language, but it is imbued with significant social, political, economic and religious implications. Its exploitation has been increased in recent years, courtesy of the globalization and modernization efforts (Altalhi, 2024). However, this is met with resistance from certain segments of society. This resistance challenges the legitimacy of English and adapts the language to align with local beliefs and traditions. As part of Vision 2030, Saudi Arabia's modernization plan, the goal is to provide students with the English language skills necessary for global competitiveness, while still fostering a strong sense of pride in their Arabic and Islamic heritage (Aldawsari, 2022).

Adhering to the worth and appreciation of English, Saudi Arabia has restructured its higher education programs by incorporating English Medium Instruction (EMI), similar to other nations (Aljehani & Modiano, 2024). The Saudi government appraises English as a global lingua franca. It is sensitized to hoist English proficiency, believing that it is a binding to introduce English as a medium of instruction in the institutes. Besides, the government affirms that teaching a single foreign language is the most effective mode to simplify language acquisition. Despite the efforts placed by the Ministry of Education, implementing this policy has many quandaries (Abdel Latif & Alhamad, 2023). The current Saudi academia direly requires the transformation of syllabus, teaching styles, and learning techniques. With the advent of ChatGPT, an architype shift is vivid amongst academia towards it. So, for the quest for improving English language proficiency among learners, the exploration of AI-based tools is instantiating the novel methods as an auxiliary of conformist teaching systems (Aljabr, 2023; Allehyani & Algamdi, 2023; Alqasham, 2023; Hasanein & Sobaih, 2023; Alammari, 2024; Mugableh; 2024). A climacteric field of foreign language study is the implementation of reading tactics by EFL users.

Reading Skills in the Saudi Arabian Context

Reading is a cognitive process that draws on various knowledge sources, including lexical features and an understanding of current events. Proficiency in reading results from the development of both decoding skills and linguistic comprehension (Indriyani et al., 2024). In Saudi Arabian EFL classrooms, reading is often approached inconsistently, with much of the class time spent on repetition and teachers assuming full responsibility for student reading. As a result, students are unable to gauge their own reading proficiency. Additionally, since EFL students learn the English alphabet and sounds prior to university, English teachers tend to assume their reading skills are already developed. This can make reading a challenging area that needs attention and improvement (Alharbi, 2022).

This study aims to investigate the efficacy of ChatGPT as a learning tool in improving the reading skills of EFL

learners in Saudi Arabia.

Hypothesis

This study posits the following hypothesis:

- ChatGPT has a strong influence on the reading comprehension of Saudi EFL learners.

Research Question

This study attempts to address the following question:

- To what extent does ChatGPT impact the reading comprehension of EFL learners in Saudi Arabia?

Literature Review

OpenAI introduced the first version of ChatGPT (GPT-1) back in 2018, building it on the foundation of large language models (LLMs). This model was updated a year later to be the GPT-2 version, offering a much better performance. However, the true rise of ChatGPT occurred when introducing the GPT-3 series in 2020, which was openly on offer for the public and accentuated the true potential of this tool. This was even emphasized when the model was later refined through supervised training and reinforced learning to the more advanced GPT-3.5 and GPT-4 series (Marr, 2023). ChatGPT excels in a variety of natural language processing (NLP) tasks, such as its outstanding ability to understand and generate natural language and its capacity to provide human-like responses. Add to this its key features, such as understanding different contexts and languages and adapting to them while producing responses that can be fine-tuned on different scales (Ray, 2023).

How ChatGPT Works

ChatGPT operates using a deep learning model known as a transformer to interpret natural language inputs and produce responses that mimic human conversation. Here is a simplified breakdown of its process:

- *Preprocessing:* The input text is first transformed into numerical representations called tokens, which capture the meaning of each word and allow the transformer model to process the input more effectively.
- *Encoding:* After tokenization, the input passes through encoding layers that use attention mechanisms to highlight the most important parts of the text, allowing the model to focus on key information.
- *Decoding:* The encoded information is then transformed back into natural language through a decoding process, also using attention mechanisms to select the most relevant details for the response.
- *Postprocessing:* Finally, any unnecessary tokens or formatting is removed, presenting the user with a coherent and natural-sounding reply. The transformer architecture enables ChatGPT to generate nuanced and contextually appropriate answers by leveraging its training on a vast dataset from diverse

sources across the internet (Su & Yang, 2023).

Moreover, ChatGPT has shown a deep impact on many fields, including education, where it can be used "to create intelligent tutoring systems capable of providing personalized assistance to students" (Marr, 2023, para. 19). It can serve as an assistant to the educator which empowers them across the educational journey for holding a solid grip over their teaching (Ooi et al., 2023). Therefore, ChatGPT can be easily positioned as a powerful and promising tool in language education.

Reading Skills

Reading is an indispensable language skill that serves learners, exclusively students for the knowledge acquisition. A proficient reader can collect the ideas and information for the newly encounter things (Wibowo et al., 2020). Reading has a critical effect in any second or foreign language learning circumstances. It supports in the recognition of words and vocabulary expansion which ultimately primes to better comprehension. Isaqjon (2022) noted that reading comprehension forms a mind-model where the text is blended with the earlier knowledge to give a new experience of learning. Understanding any text in not just the reading but rather the analysis and comprehension of the context is key. Understanding is also clinked up with personalized experiences and ideas connected in the context and beyond. Some advantages of developing comprehension in the reading skills are the following:

- Allowing learners to think critically
- Argument development
- Decision Making
- Problem solving
- Textual analysis (Isaqjon, 2022).

ChatGPT and Reading Comprehension

The use of ChatGPT in teaching English has been a trending topic in previous research. According to Li et al. (2024), although the use of some form of AI in language learning dates back to the 1960s, the research community has shown particular interest in the topic since the rise of ChatGPT as a powerful open AI tool toward the end of 2022.

Lin and Chen (2024) stated that ChatGPT can affect developing English reading skills in a variety of ways:

1. *Providing Practice Passages:* Providing the user with the accurate response, ChatGPT can give various passages and comprehension questions that can be answered after reading the passages. Also, the learners can check the central idea, details, and other important aspects of the text.
2. *Development of Vocabulary***:** In a reading text, unknown words can be comprehended by asking ChatGPT for the precise definition, meaning, and synonyms and/or antonyms. For better understanding, their use in a sentence can also be enquired. This aids in understanding the new vocabulary in a better way.

3. *Comprehending Context*: Understanding the context can be done by text analysis. Here, ChatGPT assists in presenting the tone, subtext, and then context of the passage. Further, the cultural aspect of a text can also be enquired. This helps to enhance the understanding of the text.
4. *Reading Personalization*: Assessing the interest and level of the user, ChatGPT can suggest some pertinent reading texts for the development of reading stamina. These can be articles, short essays, short stories, etc. In this way, the user can be exposed to various topics and genres and develop concentration and engagement.
5. *Interactive Summaries*: For any text, a summary can be available with ChatGPT. This helps to develop understanding of the major points and the structure of the learning material.
6. *Strategies To Support Reading Analysis and Discussion*: Users can engage in discussions about what they have read, allowing them to practice critical thinking and express their interpretation of the text. ChatGPT can also suggest questions for deeper analysis, enhancing reading engagement (Daza et al., 2024).

Limitations of ChatGPT

Though there are some positives of ChatGPT, there are some limitations too. The usage of ChatGPT in instruction may possess some limitations. Firstly, ChatGPT is a technology which is not tested and developed comprehensively, so errors may take place. Secondly, the quality of the data used may be limited as the model training needs more responses and more validation for producing effective results. It may be perfect in some cases, but not for every prompt. Lastly, ChatGPT may find producing complex tasks problematic. Although it generates human-like responses, it needs more training and its responses require a keen eye for review (Rice et al., 2024).

Theoretical Framework

Su and Yang's (2023) IDEE framework (Figure 1) has been used as the theoretical framework of this research. Their framework shows a structured content for the application of Generative AI in academia. The major components of the theory are as follows:

1. *Identifying desired outcomes*: This involves defining the educational goals, such as improving student engagement or enhancing personalized learning experiences.
2. *Determining automation*: Educators need to assess the appropriate level of automation. ChatGPT can assist with tasks like providing instant feedback or generating instructional content, but human oversight remains essential.
3. *Ensuring ethical considerations*: The ethical implications, such as data privacy, academic integrity, and bias in AI-generated content, must be carefully managed.
4. *Evaluating effectiveness*: Continuous evaluation of AI's impact on educational outcomes is crucial. This includes measuring improvements in learning and student satisfaction, while addressing limitations like untested technology and quality of data.

The framework highlights that while AI offers potential for more personalized and efficient learning, challenges

such as safety concerns, data quality, and ethical risks need to be addressed for successful implementation.

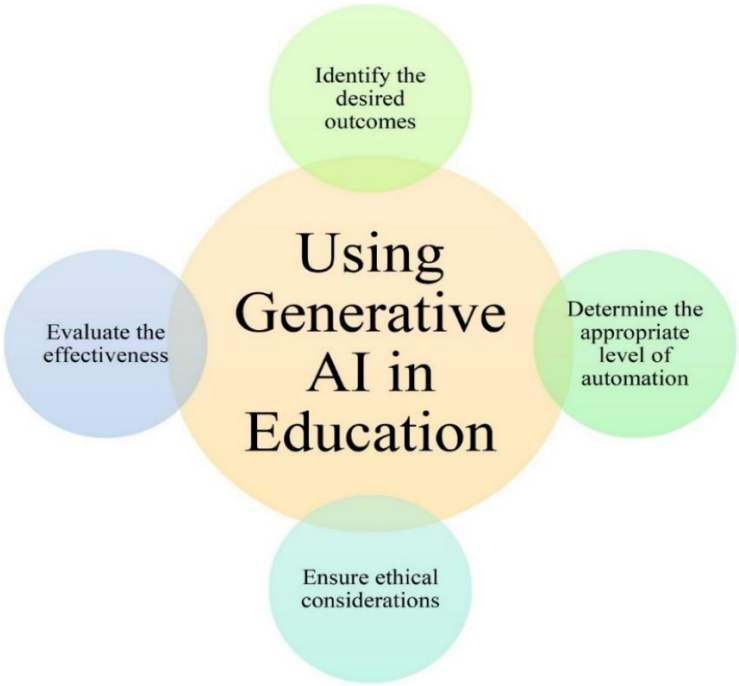

Figure 1. IDEE Framework (Su & Yang, 2023).

This is a latest model that talks about the vitality of ChatGPT into academia. Moreover, it also looks for the various aspects that determine the success of exploring ChatGPT for learning and teaching. The current study can be considered one of the pioneer studies to use this model to evaluate the results extracted from the learners' group.

Current Investigations on ChatGPT

In their review articles, Li et al. (2024) and Yang and Li (2024) emphasized that ChatGPT serves various functions in language education, like compiling information, supporting research, acting as an interlocutor or conversation mate, and assisting with writing tasks. It facilitates both independent learning and teacher-led activities by helping in content development, creating educational materials, and conducting assessments. Another review by Faisal (2024) suggested that possible benefits of ChatGPT are creating personalized educational material for different learning styles and needs, providing immediate interaction and feedback, and acting as an academic assistant, all leading to enhancing students' language proficiency. However, despite these possible benefits, ChatGPT should be used with caution due to the potential inaccuracies and biases that may not be suitable for the culture of Saudi Arabia.

In addition to the previously mentioned review articles, we conducted a review of the recent literature regarding the use of ChatGPT in the classroom. We found that some articles studied the perceptions of teachers and students (e.g., Espartinez, 2024; Lee et al., 2024; Stöhr & Malmström, 2024), and others focused on learning and assessing writing (Foung, Lin, & Chen, 2024; Özçelik & Ekşi, 2024), while others discussed the use of ChatGPT to create culture-appropriate English language teaching material (Zheng & Steward, 2024), produce effective lesson plans (Milad & Fayez, 2024), and help English language teachers become more accountable and innovative (Rahimi &

Sevilla-Pavón, 2024). However, none of these studied the effect of using ChatGPT on teaching reading in EFL settings.

ChatGPT and Saudi Arabia

Concerning the Saudi context, a few studies were found related to the topic at hand. Most of these studies (Ahmed, 2023; Ali et al., 2023; Aljabr, 2023; Allehyani & Algamdi, 2023; Alqasham, 2023; Hasanein & Sobaih, 2023; Alammari, 2024) investigated the perceptions of Saudi teachers and students concerning using AI in language education. These studies, despite their different scopes, reached almost a similar result which is that AI tools like ChatGPT can be beneficial to the language classroom and should be used in the Saudi classroom, but they should be used with caution and awareness due to possible ethical challenges, cultural concerns, and limitations. Other studies had a more empirical approach. Alhammad (2024) tested the use of ChatGPT in teaching EFL student's literature appreciation skills. He found the AI tool useful in educating students about several aspects of literature appreciation such as identifying themes and motifs, improving language proficiency, exposure to different cultures, and free interaction. Therefore, he recommended the balanced and healthy use of ChatGPT as a learning aid, while testing its effectiveness in teaching other subject areas in the EFL classroom. Finally, Mugableh (2024) investigated whether ChatGPT would have a significant impact on developing the vocabulary of Saudi EFL students. He found that students who used ChatGPT-generated exercises outperformed those who did not regarding vocabulary size and usage.

While previous research has contributed valuable insights to language learning, none have specifically examined the role of ChatGPT in improving reading proficiency among EFL learners in Saudi Arabia. This study aims to fill that gap by exploring how ChatGPT can benefit local students' reading skills. It will not only provide important findings on the use of AI in education but also serve as a foundational reference for future studies within the Saudi context.

Methodology

Design

This research follows a binary procedure. It is comprised of quantitative and qualitative methods. It has two stages. The first one addresses ChatGPT's effect on the development of reading. To execute this, two tests, one at the start and one at the end of the experiment, are conducted. The analysis is performed by SPSS 25 on participants' scores. Moreover, the second phase is to know about the opinions of those participants who learned through ChatGPT. These learners provide their insights regarding the use of ChatGPT in an interview. The analysis of this interview is done using Claude.ai which is a well-known AI app for thematic analysis.

Population and Sampling

The population of the study was EFL learners in KSA. The sample comprised 120 learners from 12 local government schools. These learners were studying in grade 10. Only male schools were considered as a sample.

Both the schools and the learners were picked by purposive method of sampling. The schools were following English Medium instruction (EMI) and were connected with the Ministry of Education.

Research Instrument

The main instrument was tests which were conducted at two instances to look at the efficacy of ChatGPT on reading abilities in EFL situation. The authenticity of the experiment hinges on the validity and reliability of the instrument. Since the textbook content is curated by experienced educators to cater to learners of varying abilities, selecting test items from this source significantly enhances the validity and reliability of the test. So, it maintained the content validity. The textbook included over 50 comprehension passages. These passages were categorized by difficulty level: easy, normal, and hard. Moreover, the tests items were already made in the book to cater learners with mix ability. This approach ensures that the test is both accurate and consistent in evaluating reading proficiency across different skill levels.

Research Process

To initiate the research data collection is significant. For the very reason, two tests were executed at the beginning and end of the experiment. First test was a placement test and the test takers were placed into two groups (Control and experiment) on the scores they achieved in the test. It was ensured that both groups possess learners with mix abilities to avoid biasness in sampling distribution. At the end, a final test was conducted and results were compared. The time for both tests was 30 minutes. Both the tests had 5 comprehension passages (5) questions each passage. The test sheets had clear instructions, time duration and other details for candidates to avoid confusion.

Learning Process with ChatGPT

The learners in the treatment group learnt using ChatGPT. The protocols of learning were set already. Firstly, the teachers gave pertinent prompts to ChatGPT to self-learn itself for relevant output. Then, the learners were asked to utilize ChatGPT for learning the comprehension passage. The teacher act as a prompt facilitator and a guide in the whole treatment span. Teacher also helped them to clarify if anything became complex for learners. Being a large learning model ChatGPT did not need training to develop its understanding regarding the prompts it received for providing accurate output.

Further ChatGPT helped the learners in the following ways:

- It clarified the difficulties in the comprehension text by giving explanations and also changing the complex sentences into simpler ones. This helped the learners to develop learning coherence.
- It helped in vocabulary acquisition, where learners asked for the synonyms and contextual usage of words. This enhanced their vocabulary comprehension.
- It summarized the main ideas of the text. Moreover, the learners got summary points of the chunks and the whole passage as well. This provided support in retention of paragraph

- Important aspects of the paragraph were highlighted by ChatGPT. This gives pertinent information and connected learners to the major theme (see Appendix 1)

Research Variables

Various variables existed in this study. The independent variable was ChatGPT with learning being the dependent one. As ChatGPT has been trained and retrained universally by experts and till date it is considered the highest-rated conversational AI app, its validity and reliability have already been constructed through numerous trails. The confounding variables were teacher's attitude, level of intellect, and demographics. To address this, only well-experienced instructors were picked up for the experiment process. All of them had been teachers of English for the last 8 years. Also, they were adequate in using ChatGPT. Going on, the participants of this research were from the science group. Therefore, supposedly, they had the same level of intellect. Some of the moderators like age, gender, race, marks in the previous class, English learning level, classroom ambiance, method of instruction, test time and pattern, and digital knowledge were considered.

All participants were Saudi males between 14 and 15. All of them had performed well in their previous class, and they started learning English from grade 5. The classroom ambiance was comfortable and promoted learning.

The instructional mode was English, and the time of both the preliminary and final tests for each group was 45 minutes. Each test contained three comprehension passages with 7 questions each. The test pattern was identical to their comprehension passages in the English prescribed book. Total class time daily was 50 minutes, and the class took place in the morning. Finally, all the experiment group learners had used ChatGPT before and their digital knowledge was excellent.

Results

This section tows out the analyzed data by reflecting the findings acquired by applying t-tests (paired and independent-sample). Tables 1–7 parade the performance of the control and experimental groups in the t-tests. The contrast of the two student groups under examination was viewed at two varied points, i.e., the start and the end. Besides, skewness along with kurtosis was used for data validity.

Comparing the Treatment Circumstances and Initial Assumptions

Before conducting illative analyses to compare treatment circumstances, early assumptions were calculated. This also was comprised of examining the skewness and kurtosis of the constructs. The standard deviation gauges the intensity or extent of variability within and among the samples. Thus, it is a prevalent tool across statistical analyses and holds significant weight across disciplines. It also contributes crucial insights into data variability and distribution.

The following is the key to the terms mentioned in the tables of the analysis:

Mean Value	:	MV
Standard Deviation	:	STDV
Skewness	:	SWS
Kurtosis	:	KTS
Shapiro-Wilk	:	SHW
Control Group	:	CLG
Experimental Group	:	EXPG
Confidence Interval	:	CI
Lower Limit	:	LWRL
Upper Limit	:	UPRL
Cohen's D	:	CD

Table 1 reports the data for the pre- and post-test of two participating groups (CLG and EXPG). Some important analyses like MV, STDV, SKS, and KRT are mentioned. Adding on, the SHW values for the two participating groups are wrote down. The value of SHW justifies the normality of the data. The data is taken as normal if it is between -2 and +2 for SKS (De Luca et al., 2024). The SHW test was also inconsequential. This guaranteed that the two participant groups were distributed normally. Going on, the overall analyses were done by the 5000 bootstrap method which is a popular and authentic way to look for the sample distribution's estimation (Cheung & Cheung, 2024).

Table 1. Statistics Presentation of Pre- and Post-Testing: CLG and EXPG Groups

TC		MV	STDV	SWS	KTS	SHW
CLG	Pre-test	6.66	1.23	0.59	0.59	2.81
	Post-test	9.35	1.05	0.18	0.39	2.90
EXPG	Pre-test	6.70	1.24	-0.07	-0.67	2.93
	Post-test	16.65	0.61	0.30	-0.20	2.03

(no. of participants = 50), including SWS and KTS.

To extract an analysis from the sample data representing both participant groups, an independent-sample t-test was employed through SPSS 25 (Table 2). The values of the mean for CLG and EXPG are 6.66 and 6.68, respectively. This corroborates that a slight difference lies in terms of the performance between the two participant groups. The STDV measuring the range to which the sample data diverge from their particular means was marginally lower for the CLG (SDV = 1.23) compared to the EXPG (SDV = 1.24). This STDV's comparison affirms that a minimal difference lies in the data dispersion around the mean. The homogeneity assumption's variance was inveterate, denoted by an F-value of 0.016 along with a P-value > 0.05, signifying that the variance of the pre-test scores remained consistent across both groups. The independent-sample t-test conducted on the pre-test data further revealed no notable disparities between both groups, reflecting that the groups were comparable in their performance prior to the intervention. However, the t-test statistic (t = -2.81) gave a P-value

of 0.005, which is less than the alpha level of 0.05. This suggests that the two sample groups represent distinct populations. The confidence interval for the t-test, with a lower value of -1.16 and an upper value of -0.21, implies that the true population parameter (such as the mean difference between the groups) likely falls within this range. The absence of zero within this interval indicates that the observed effect is statistically significant at the chosen confidence level. Moreover, the effect size, as measured by Cohen's D, was calculated to be 0.016. According to Cohen's (1988) guidelines, this value suggests a medium magnitude of dissimilarity in the pre-test values of the experimental and control groups. The effect size was further validated by Glass's delta (0.0162) and Hedges' g (0.01626), which also confirmed the magnitude of the effect and the standard deviations of the samples.

Table 2. Independent-Sample t-Test Pre-Test Results (CNG and EXG), N = 100

Variable	EXPG (50)		CLG (50)				95% CI		
	MV	STDV	MV	STDV	t (118)	P	LWRL	UPRL	CD
Pre-test	6.70	1.24	6.66	1.23	-2.81	.005	-1.16	-0.21	0.016

Glass's delta = (6.68 - 6.66)/1.23 = 0.01626.

Hedges' g = (6.68 - 6.66)/1.23 = 0.01626.

The performance of the CLG and the EXPL was assessed through a comparative analysis (Table 3), revealing that the MV of the EXPG (16.65) is significantly higher than the MV of the CLG (9.35). This distinguished variance proposes the high effectiveness of the intervention tool. Observing the standard deviation (STDV = 0.61) of the EXPG indicates a wider dispersion of data points from the mean, presenting the superior performance of the EXPG compared to the CLG (1.05). The homogeneity assumptions' variance was validated, as evidenced by an F-value of 1.291 with a P-value > 0.05, representing that the variance in post-test scores was consistent across both groups.

Table 3. Independent-Sample t-Test Pre-test Results (CNG and EXG), N = 100

Variable	EXPG (50)		CLG (50)				95% CI		
	MV	STDV	MV	STDV	t (118)	P	LWRL	UPRL	CD
Post-Test	16.65	0.61	9.35	1.05	-25.81	.005	-6.21	-4.877	8.50

Glass's delta = (9.35 - 16.65)/0.61 = 11.967213.

Hedges' g = (9.35 - 16.65)/0.858662 = 8.501601.

The t-test statistic (t = -25.81) strongly supports the conclusion that the two groups represent distinctly different populations, with the EXPG exhibiting a significantly higher mean than the CLG. The confidence interval for the t-test, ranging from -6.21 to -4.87, suggests that the true population parameter likely falls within this range. The exclusion of zero within this interval confirms the statistical significance of the observed effect at the chosen confidence level. Moreover, effect size, as measured by CD, was 8.50, indicating a large magnitude of difference between the groups, as per Cohen's (1988) guidelines. This large effect size is further corroborated by the values

of Glass's delta (11.96) and Hedges' g (8.501), which both authenticate the substantial effect size and the standard deviations of the samples.

The analysis of the CNG using a paired-sample t-test dips into their performance across the preliminary and final assessments (Table 4). The MV in the pre-test was 6.66, and 9.35 in the post-test, suggesting that traditional methods had little impact on learning outcomes. Examining the data dispersion, the STDV was 1.23 in the pre-test and 1.05 in the post-test, indicating a narrow spread of data points around the mean in the post-test. This reduced variability suggests a bit of consistency in the post-test. The comparison of pre-test and post-test scores for the CNG is further highlighted by the t-value ($t = -25.81$, $P < .000$), which shows that while the post-test mean score (9.35) was slightly higher than the pre-test mean (6.66), the difference was not substantial. The t-test statistic ($t = -25.81$, with $df = 48$) and the corresponding P-value of 0.000* (which is less than $\alpha = 0.05$) indicate that the CNG's pre-test and post-test scores did not differ much significantly. It suggests that the students' performance, on average, improved, but not markedly so. Moreover, the 95% confidence interval limits (CIL) suggest that the population mean difference (μd) is likely to fall within this interval, confirming the statistical accuracy of the observed changes. Despite some improvement in the CNG's scores, it is crucial to note that this improvement was not substantial. Finally, the effect size, measured by Cohen's D, was 2.35 for the CNG, indicating a lower magnitude of difference between the pre-test and post-test scores, according to Cohen's (1988) classification. This further reinforces the finding that while there was a slight improvement, the impact of the intervention on the CNG was minimal.

Table 4. Comparison of Pre-Test and Post-Test Scores in the Control Group Using Paired-Sample t-Test (N = 50)

VR	Pre-test		Post-test				95% CI		
	MV	STDV	MVE	STDV	t (149)	P	LWRL	UPRL	CD
CLG	6.66	1.23	9.35	1.05	-25.81	.000	-2.71	-2.6	2.35

Glass's delta = (9.35 - 6.66)⁄1.23 = 2.186992.

Hedges' g = (9.35 - 6.66)⁄1.143547 = 2.35233.

The values mentioned substantiate the divergence of the CNG in their pre-test denoted by x1, post-test denoted by x2 (Table 5). The minimum value is x1, while the highest is x2. Additionally, x1 / m1 = 6.66 and x2 / m2 = 9.35. This explains the diverse array of MV signified as md = (m2 – m1) = 2.69, whereas MV ratio mr = (m2 / m1) = 0.014 = (0 + 01.4%). The joint ME (m1, m2) is calculated through a center point mc = 8.005, and the variance between them (m1, m2) is calculated by vp = 1.269 relating the unpredictability among the groups. Congruently, v1 (variance) is x1 = 1.074, and v2 (variance) is x2=1.599. Thus, justifying the change in the variance showed by vd (v2 – v1) = 0.485, and also the ratio of variance denoted by vr (v2 / v1) = 1.451. Going ahead, v1, v2 (pooled variance) is calculated by v = 1.3661, telling the variation inside the groups (x1, x2), and vv = 0.166 designates the change of variance of groups (x1, x2). The paired-sample confirmations pointedly have a positive correlation (i.e., r = 0.775) amongst pre- and post-test clusters.

Table 5. Summary Item Statistics

Variables	Mean	Minimum (pre-test)	Maximum (post-test)	Range (x1 and x2)	Ratio	Max./ Min.	Variance
Item Means	mc = 8.005	m1 = 6.66	m2 = 9.35	Md = 2.687	mr = 1.327	vb = 3.394	2
Item Variances	vp = 1.290	v1 = 1.080	v2 = 1.659	vd = 0.579	vr = 1.371	vv = 0.161	2
Inter-Item Correlations	.785	.785	.785	1.000		.000	2

The performance of a single group (EXPG) in the preliminary and final tests is measured by paired-sample t-test (see Table 6). The MV in the preliminary test (6.70) is vividly significant compared to the MV value in the final test (16.65). This elucidates that the intercession instrument has formed substantial effect on students' scores. Here and now, viewing the data disparity, the STDV is 1.24 in the pre- and 0.61 in the post-test group. This verifies that the post-test group reflect wider data points spread around the MV in contrast to the pre-test group, representing high variance in the scores in the post-test. The inequalities amid the pre-test and post-test of the EXPG are mentioned through the t-value (t = -63.75, with df = 47) having a P-value of 0.000* < (α = 0.05). This t-test statistic value designates that both sample groups are from varied populations, thus suggesting a better performance of the participants in the post-test group compared to the pre-test group. Further, the effect size of CD was 10.26, signifying a higher difference, as per Cohen's classification (1988). This increase validates the inclusion of the tool used for the experiment. Overall, the scores reflect that the null hypothesis (Ho: ρ = 0) is to be rejected, suggesting that there is a significant and positive correlation between the two populations.

Table 6. Comparison of Pre-Test and Post-Test Scores in the Experimental Group via Paired-Sample t-Test (N = 50)

VR	Pre-test		Post-test				95% CI		
	MV	STDV	MV	STDV	t (149)	P	LWRL	UPRL	CD
EXPG	6.70	1.24	16.65	0.61	-63.75	.000	-8.91	-7.98	10.26

Glass's delta = (16.65 - 6.68)/1.23 = 8.105691.

Hedges' g = (16.65 - 6.68)/0.970824 = 10.269622.

The figures declared authenticate the aberration of EXPG in their pre-test as x1 and post-test as x2 (see Table 7). The least value is x1, and the highest is x2. Also, x1 or m1 = 6.70 and x2 or m2 = 16.65. This explicates the mean diverseness md = (m2 – m1) = 9.95, where MV ratio mr = (m2 / m1) = 2.481 = (2 + 48.1%). The common ME (m1, m2) is measured on a middle point mc = 11.685, and the variance between them (m1, m2) is premeditated by vp = 3.284 displaying the variance amid the groups. Correspondingly, consider v1 (variance = x1 = 1.112) and v2 (variance = x2 = 3.284). Therefore, authenticate the variance change signified as vd (v2 – v1 = 0.485). The ratio of variance is vr (v2 / v1 = 2.953). Moving forward, v1, v2 (pooled variance) is measured by v = 3.284 showing the change occurring within the groups (x1, x2); and vv = 2.981 labelling the variance difference of the groups (x1, x2). This paired-sample affirmation has a positive correlation (i.e., r = 0.739) between pre- and post-test participant groups.

Table 7. Summary Item Statistics

	Mean	Minimum	Maximum	Range	Max. / Min.	Variance	No. of items
Item Means	mc = 11.675	m1 = 6.70	m2 = 16.65	md = 9.95	mr = 1.291	vb = 29.867	2
Item Variance	vp = 3.284	v1 = 1.112	v2 = 3.284	vd = 2.312	vr = 2.953	vv = 2.981	
Inter-Item Correlations	.739	.739	.739	.000	1.000	.000	2

Discussion

In this research, ChatGPT has shown to significantly enhance learners' reading comprehension skills. The improved performance of the learners is a direct result of the tool's ability to mimic human-like responses through a process involving pre-processing, encoding, decoding, and post-processing of instructions (Su & Yang, 2023). English reading comprehension, especially for second and foreign language learners, can be challenging. However, ChatGPT provided multiple avenues for practice texts tailored to the learners' proficiency levels, which gave the experimental group a distinct advantage. This agrees with Lin and Chen (2024) who confirmed that ChatGPT is a valuable tool for developing English reading skills.

One distinguished feature explored in this research of ChatGPT was its capability to provide practice passages to users. This permitted learners to repetitively practice reading by exploring different texts, thus accustoming them to answering comprehension questions. Additionally, the feedback generated by ChatGPT could be reviewed by teachers for further simplification and clarification which enhanced the learning experience. The tool also played a decisive part in vocabulary building, offering synonyms, contextual antonyms, and simplified sentence meanings. This aligns with the research by Mugableh (2024) who stated that contextual vocabulary is key to understanding texts.

Moreover, ChatGPT exposed learners to cultural contexts, tone, and nuances in reading materials, contributing to a broader understanding. The personalized reading options provided by ChatGPT helped increase the learners' reading stamina by offering varied reading materials. This expanded their concentration and focus. Faisal (2024) also reflected the idea that reading stamina can be improved through exposure to different types of passages, essays, and articles, as noted in various studies. Additionally, ChatGPT's ability to summarize key points and passages facilitated learners to grasp the main ideas and context of the texts. a practice supported by research that highlights the importance of summarization in enhancing comprehension and analytical skills (Daza et al., 2024). This multi-dimensional support from vocabulary expansion to improved focus and comprehension significantly contributed to the improved reading abilities of the experimental group.

Observing it from the IDEE framework by Su and Yang (2023), this study aptly aligns with its core components, particularly concerning the integration of generative AI tools like ChatGPT. A central emphasis of the framework is to classify desired outcomes, which was also the prime goal of the educators and researchers in this study. Through the use of ChatGPT, they aimed at enhancing personalized learning and facilitating improved student outcomes. However, the process did not rely solely on AI; teachers actively monitored student inputs and provided

guidance to ensure that the automation was accurate and appropriate. This aligns with the IDEE framework's emphasis on controlling the degree of automation when generating content and delivering feedback. Throughout the experiment, ethical guidelines were strictly adhered to, ensuring that AI-generated responses were unbiased, further reinforcing another element of the IDEE framework—ethical AI use. The teacher's role in continuous oversight ensured that ChatGPT was effectively meeting the learners' needs. To manage the limitations of the tool, diverse prompts were used to ensure more precise results, which is in line with the IDEE framework's recommendation for the ongoing evaluation of AI's impact to optimize its effectiveness.

Overall, this study ties itself with studies by Espartinez (2024), Lee et al. (2024), and Stöhr and Malmström (2024) who experienced the use of ChatGPT in the EFL classroom and found it to be a productive element. Besides performing a descriptive analysis, a thematic analysis was performed using Claude.ai to look for significant themes present in the test. For this reason, fifty interviews were conducted which was a lengthy procedure. For the interviews, an open-ended questionnaire was structured. It took 20 days for conducting the interviews. These interviews are significantly pivotal as they provide a real opinion of individuals (students) who used AI apps for the experiment process. The questions of the interview can be found in Appendix 1.

Some core conditions of the interview were as follows:

- The participants were informed about the interview prior to recording.
- The recording was done with their consent.
- The interviews were done in a comfortable environment to avoid noise, weather, and other factors.

All the interviews were later transcribed and some themes were picked from Claude.ai such as the following:

- Passionate about usage
- Sufficient information
- Ease of use
- Reliability
- Variety in the learning process
- Beneficial for grammar improvement
- Is it a distraction?

Moving on, for the analysis of the given themes in the text, the following steps were taken:

- Initially, familiarity was developed with the text.
- Then, codes were generated where the key features were labeled as passion and engagement, effectiveness and satisfaction in learning, ease of use, reliability, variety in the process of learning, improvement of grammar, and non-distractive element (see Appendix 2 for the complete codes).
- The next step was to explore these themes in the text. Thus, these codes were then allocated with the broader yet pertinent lines to make a viable theme.
- Along with extracting a theme, each theme was given a name, and after that all the themes were reviewed to finalize the procedure of the data analysis.

Table 8 defines the themes extracted from the Claude.ai analysis.

Table 8. Claude.ai Analysis of Themes

Positive engagement and passion	Includes passion and engagement with ChatGPT
Effectiveness and satisfaction in learning	Covers information sufficiency and accuracy
Ease of use	Relates to ease in using the software
Reliability	Talks about the accuracy of the tool
Variety in the process of learning	Talks about the variation of ChatGPT in explaining the query of the user
Improvement in grammar	Refers to the improvement in grammar for the learners
Non-distractive learning environment	Refers to the focus and relevance of the ChatGPT responses

In the next step, two examples were given for each theme from the text, as presented in Table 9, to understand the themes.

Table 9. Examples of Themes

Positive engagement and passion	Student 12: "Yes, I feel quite energetic and passionate about using ChatGPT." Student 25: "I felt it is attractive and interactive with fun-learning."
Effectiveness and satisfaction in learning	Student 43: "ChatGPT gives sufficient knowledge and information regarding my grammatical questions." Student 31: "The way it explains and exemplifies the question in simple language is easy for me to comprehend."
Ease of use	Student 04: "I don't find any problem in using ChatGPT. The software is very smooth in use and easy in operation." Student 49: "The user can easily operate with simple instructions. There is no difficulty I face in using it."
Reliability	Student 17: "I feel that the responses are accurate and updated." Student 02: "I found it very reliable as it gives me an accurate and correct reply to each of my prompts."
Variety in the process of learning	Student 26: "In my view, ChatGPT is a latest way of learning and understanding." Student 48: "I feel it is hard to ask again for a grammatical question as in class due to shyness and fear."
Improvement in grammar	Student 09: "Indeed, I think ChatGPT is a sensational software for giving variety in the procedure of learning magnificently." Student 27: "It gives multiple replies to satisfy the user. I feel several replies to the same queries are its prolific feature."
Non-distractive learning environment	Student 05: "In my viewpoint, it is an advantage and it does not distract its users. I found it a great learning partner and interactive tech-dependent tool." Student 43: "Only replies to the desired prompts and there is no irrelevant information for any distraction in there."

Figure 2 presents the thematic analysis of the user responses regarding their experience with ChatGPT. The graph shows the prominence of each theme based on the number of references in the text.

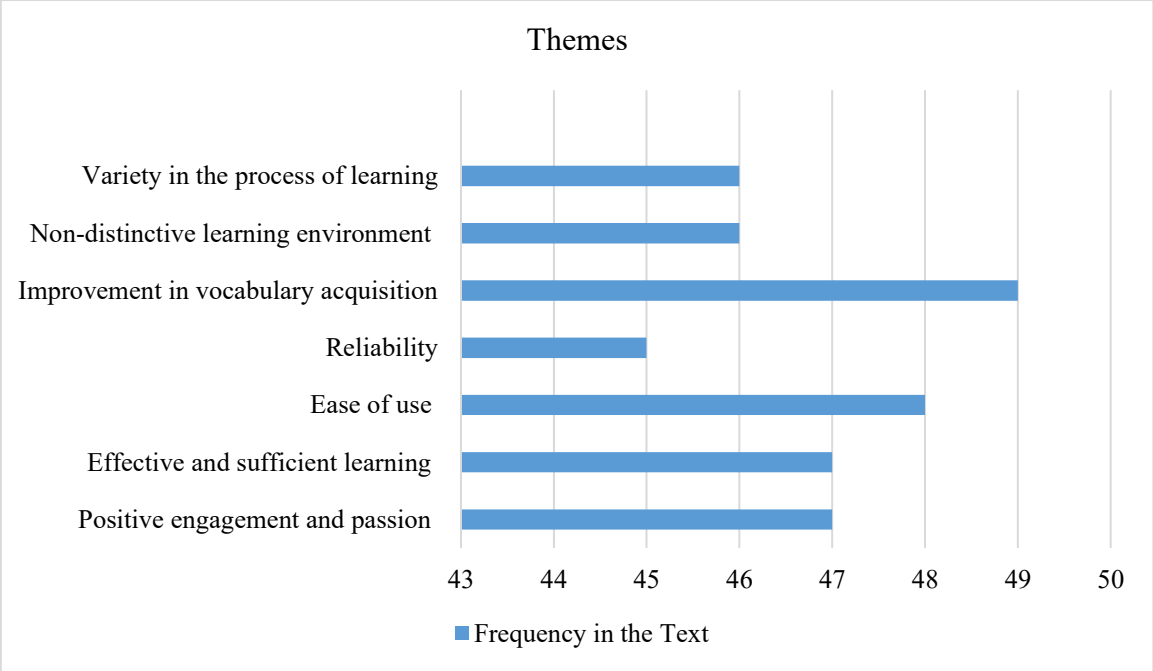

Figure 2. Most Frequently Mentioned Themes in the Text

In the last step, we summarized the findings in a report format and presented the results using visual tools like graphs or tables. We found the following:

- Effective learning and grammar improvement were the most frequently mentioned themes, indicating the significant role ChatGPT plays in enhancing learning and grammar understanding.
- Positive engagement and passion and ease of use and reliability were also prominent, reflecting the user's enthusiasm and the ease of using the tool.
- Non-distractive learning environment was mentioned but with less frequency, emphasizing the focused learning environment that ChatGPT provides.

This visual representation helps to quickly grasp the key themes and their relative importance in the user's feedback. Extracting the themes from any interview gives an inclusive reflection of the ideas shared by the participants of the interviews. The interviews conducted in this research were of high importance as they provide the opinions of those participants who used ChatGPT in the experimentation process. In the context of reading skills, it is essential for any tool to be engaging and stimulating to sustain learners' interest, as reading requires stamina. Students indicated that ChatGPT was a highly engaging tool that fostered their enthusiasm for learning. Additionally, repetitive learning methods can become tedious, but with ChatGPT, the varied responses to the same prompts introduced diversity, effectively satisfying the learners' curiosity. This aligns with the view that variation is a hallmark of ChatGPT (Ooi et al., 2023).

Moreover, many AI tools can be overwhelming due to lengthy instructions or complex operations, which can

frustrate users. However, the participants shared that ChatGPT's simplicity and user-friendliness motivated them to engage with the tool easily. Its reliability in providing accurate responses also increased their trust in the tool, resonating with the opinion that a well-trained AI model can be dependable (Ray, 2023).A crucial component of reading comprehension is understanding grammar. Often, learners struggle with sentence structure and contextual meaning. Participants felt that ChatGPT helped them grasp the context and pragmatic meanings without much difficulty. Finally, all participants agreed that ChatGPT did not hinder their reading comprehension practice. Instead, it offered valuable assistance by providing relevant information in a simplified manner, helping them stay focused and acquire knowledge effectively. This resonates with the conclusions of Li et al. (2024), Lin and Chen (2024), Faisal (2024), and Yang and Li (2024).

Conclusion

ChatGPT has demonstrated significant potential as a transformative tool for language acquisition. To evaluate its effectiveness in enhancing reading skills within Saudi Arabian English as a Foreign Language (EFL) classroom, a controlled experimental design was employed. The study involved two groups: one receiving traditional instruction, and the other integrating ChatGPT interactions with conventional teaching methods. Both cohorts were exposed to identical materials and activities over a three-week period. Upon completion, a reading comprehension assessment was administered, and results were analyzed for comparative purposes. Additionally, a series of interviews was conducted to capture students' perspectives on the utility of ChatGPT in their learning experience.

The findings revealed a marked improvement in reading proficiency for the group utilizing ChatGPT in conjunction with traditional pedagogy. The interview data corroborated these outcomes, with students citing enhanced learning, improved grammatical accuracy, heightened engagement, ease of use, and reliability as key advantages. These results underscore ChatGPT's potential as a powerful AI-driven educational tool capable of significantly augmenting language instruction in EFL contexts. The incorporation of AI into educational practices offers a promising avenue for advancing language proficiency and improving the overall learning experience.

Nevertheless, the study was limited by its focus on a single gender, school grade, and nationality. Future research should explore the application of ChatGPT or other AI technologies across a more diverse population, including different educational levels, university students, and female learners, to facilitate comparative analyses. Moreover, replication of the study in various national and cultural contexts would provide insights into the universal applicability of AI tools in education. As AI continues to gain prominence in educational discourse, it presents a vast array of possibilities for research and instructional innovation.

References

Abdel Latif, M. M., & Alhamad, M. M. (2023). Arabicization or Englishization of higher education in the Arab world? Controversies, policies and realities. *Frontiers in Psychology*, *14*, 1093488. https://doi.org/10.3389/fpsyg.2023.1093488

Ahmed, M. A. (2023). ChatGPT and the EFL classroom: Supplement or substitute in Saudi Arabia's eastern region. *Information Sciences Letters, 12*(7). https://digitalcommons.aaru.edu.jo/isl/vol12/iss7/4

Alammari, A. (2024). Evaluating generative AI integration in Saudi Arabian education: A mixed-methods study. *PeerJ Computer Science, 10*, e1879. https://doi.org/10.7717/peerj-cs.1879

Aldawsari, M. (2022). Deciphering students and teachers' perceptions towards English as a Medium of Instruction: A case study of Prince Sattam Bin Abdulaziz University. *Theory and Practice in Language Studies*, *12*(12), 2687–2694. https://doi.org/10.17507/tpls.1212.27

Alhammad, A. (2024). The impact of ChatGPT in developing Saudi EFL learners' literature appreciation. *World Journal of English Language, 14*(2), 331. http://dx.doi.org/10.5430/wjel.v14n2p331

Alharbi, A. M. (2022). Reading skills among EFL learners in Saudi Arabia: A review of challenges and solutions. *World Journal of Advanced Research and Reviews*, *15*(3), 204–208. https://doi.org/10.30574/wjarr.2022.15.3.0922

Ali, M. M., Alaa, A. M. & Shahnaz, A. (2024). The impact of mobile learning in English language classrooms in Pakistan. *Asian-Pacific Journal of Second and Foreign Language Education, 9*(62). https://doi.org/10.1186/s40862-024-00274-0

Ali, J. K. M., Shamsan, M. A. A., Hezam, T. A., & Mohammed, A. A. Q. (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

Aljabr, F. (2023). Gauging the Saudi EFL learners' level of awareness and attitudes towards the use of ChatGPT. *Information Sciences Letters, 12*(11). https://digitalcommons.aaru.edu.jo/isl/vol12/iss11/20

Aljehani, K., & Modiano, M. (2024). EMI in the Saudi Arabian context. *Discover Education*, *3*(1), 1–13. https://doi.org/10.1007/s44217-024-00244-1

Allehyani, S. H., & Algamdi, M. A. (2023). Digital competences: Early childhood teachers' beliefs and perceptions of ChatGPT application in teaching English as a Second Language (ESL). *International Journal of Learning, Teaching and Educational Research, 22*(11), 343–363. https://doi.org/10.26803/ijlter.22.11.18

Alqasham, F. H. (2023). ChatGPT in the Saudi EFL classroom: A study of learner usage patterns and possibilities in learning optimization. *Migration Letters, 20*(S7), 1251–1263. https://doi.org/10.59670/ml.v20iS7.4828

Al Sawi, I., & Alaa, A. (2024). Navigating the impact: A study of editors' and proofreaders' perceptions of AI tools in editing and proofreading. *Discover Artificial Intelligence, 4*(23). https://doi.org/10.1007/s44163-024-00116-5

Altalhi, N. A. (2024). EFL teachers' and learners' perceptions of code-switching: The role of learners' L2 proficiency levels. *English Language Teaching*, *17*(4), 1–23. https://doi.org/10.5539/elt.v17n4p23

Cheung, S. F., & Cheung, S. H. (2024). manymome: An R package for computing the indirect effects, conditional effects, and conditional indirect effects, standardized or unstandardized, and their bootstrap confidence intervals, in many (though not all) models. *Behavior Research Methods*, *56*(5), 4862–4882. https://doi.org/10.3758/s13428-023-02224-z

Daza, E. P. S., Angulo, F. D. R. M., & Lozada, H. R. (2024). ChatGPT-based didactic strategies to improve students' English language reading and writing skills. *Estudios y Perspectivas Revista Científica y*

Académica, *4*(1), 390–420. https://doi.org/10.61384/r.c.a.v4i1.108

De Luca, D. L., Ridolfi, E., Russo, F., Moccia, B., & Napolitano, F. (2024). Climate change effects on rainfall extreme value distribution: the role of skewness. *Journal of Hydrology*, *634*, 130958. https://doi.org/10.1016/j.jhydrol.2024.130958

Espartinez, S. A. (2024). Exploring student and teacher perceptions of ChatGPT use in higher education: A Q-Methodology study. *Computers and Education: Artificial Intelligence, 7*(2024), 100264. https://doi.org/10.1016/j.caeai.2024.100264

Faisal, E. (2024). Unlock the potential for Saudi Arabian higher education: A systematic review of the benefits of ChatGPT. *Frontiers in Education, 9*(2024), 1325601. https://doi.org/10.3389/feduc.2024.1325601

Foung, D., Lin, L., & Chen, J. (2024). Reinventing assessments with ChatGPT and other online tools: Opportunities for GenAI-empowered assessment practices. *Computers and Education: Artificial Intelligence, 6*(2024), 100250. https://doi.org/10.1016/j.caeai.2024.100250

Hasanein, A. M., & Sobaih, A. E. E. (2023). Drivers and consequences of ChatGPT use in higher education: Key stakeholder perspectives. *European Journal of Investigation in Health, Psychology and Education, 13*(11), 2599–2614. https://doi.org/10.3390/ejihpe13110181

Huesca, G., et al. (2024). Effectiveness of using ChatGPT as a tool to strengthen benefits of the flipped learning strategy. *Education Sciences*, *14*(6), 660. https://doi.org/10.3390/educsci14060660

Indriyani, V., Fendi, H., Jasmienti, J., & Rahmayani, R. D. (2024). Comparing Problem-Based Learning (PBL) and active learning in augmenting students' proficiency in reading comprehension. *Jurnal Pendidikan Progresif*, *14*(1), 74–88. http://dx.doi.org/10.23960/jpp.v14.i1.202406

Isaqjon, T. (2022). Strategies and techniques for improving EFL learners' reading skills. *"Involta" Innovation Scientific Journal, 1*(11), 94–99. https://doi.org/10.5281/zenodo.7277769

Lee, D., et al. (2024). The impact of Generative AI on higher education learning and teaching: A study of educators' perspectives. *Computers and Education: Artificial Intelligence, 6*(2024), 100221. https://doi.org/10.1016/j.caeai.2024.100221

Li, B., Lowell, V. L., Wang, C., & Li, X. (2024). A systematic review of the first year of publications on ChatGPT and language education: Examining research on ChatGPT's use in language learning and teaching. *Computers and Education: Artificial Intelligence, 7*(2024), 100266. https://doi.org/10.1016/j.caeai.2024.100266

Lin, Z., & Chen, H. (2024). Investigating the capability of ChatGPT for generating multiple-choice reading comprehension items. *System*, *123*, 103344. https://doi.org/10.1016/j.system.2024.103344

Marr, B. (2023, May 19). A short history of ChatGPT: How we got to where we are today. *Forbes*. https://www.forbes.com/sites/bernardmarr/2023/05/19/a-short-history-of-chatgpt-how-we-got-to-where-we-are-today/

Milad, M., & Fayez, F. (2024). Your way AI way: Let's meet half-way: Paradigm shift in TEFL. *World Journal of English Language, 14*(5), 468. http://dx.doi.org/10.5430/wjel.v14n5p468

Mugableh, A. I. (2024). The impact of ChatGPT on the development of vocabulary knowledge of Saudi EFL students. *Arab World English Journal, 2024*, 265–281. https://dx.doi.org/10.24093/awej/ChatGPT.18

Ooi, K. B., Tan, et al. (2023). The potential of generative artificial intelligence across disciplines: Perspectives and future directions. *Journal of Computer Information Systems*, 1–32.

https://doi.org/10.1080/08874417.2023.2261010

Özçelik, N. P., & Ekşi, G. Y. (2024). Cultivating writing skills: The role of ChatGPT as a learning assistant—A case study. *Smart Learning Environments, 11*(2024). https://doi.org/10.1186/s40561-024-00296-8

Rahimi, A. R., & Sevilla-Pavón, A. (2024). The role of ChatGPT readiness in shaping language teachers' language teaching innovation and meeting accountability: A bisymmetric approach. *Computers and Education: Artificial Intelligence, 7*(2024), 100258. https://doi.org/10.1016/j.caeai.2024.100258

Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*, *3*, 121–154. https://doi.org/10.1016/j.iotcps.2023.04.003

Rice, S., Crouse, S. R., Winter, S. R., & Rice, C. (2024). The advantages and limitations of using ChatGPT to enhance technological research. *Technology in Society*, *76*, 102426. https://doi.org/10.1016/j.techsoc.2023.102426

Stöhr, C., Ou, A. W., & Malmström, H. (2024). Perceptions and usage of AI chatbots among students in higher education across genders, academic levels and fields of study. *Computers and Education: Artificial Intelligence, 7*(2024), 100259. https://doi.org/10.1016/j.caeai.2024.100259

Su, J., & Yang, W. (2023). Unlocking the power of ChatGPT: A framework for applying Generative AI in education. *ECNU Review of Education, 6*(3), 355–366. https://doi.org/10.1177/20965311231168423

Wibowo, E. W., et al. (2020). An analysis of English teachers' strategies in teaching reading comprehension. *Journal of Applied Linguistics and Literacy*, *4*(1), 20–27. https://jurnal.unigal.ac.id/jall/article/view/3153

Yang, L., & Li, R. (2024). ChatGPT for L2 learning: Current status and implications. *System, 124*(2024), 103351. https://doi.org/10.1016/j.system.2024.103351

Zheng, Y., & Stewart, N. (2024). Improving EFL students' cultural awareness: Reframing moral dilemmatic stories with ChatGPT. *Computers and Education: Artificial Intelligence, 6*(2024), 100223. https://doi.org/10.1016/j.caeai.2024.100223

Author Information

Muhammad Mooneeb Ali Khan
https://orcid.org/0000-0002-0356-320X
Higher Education Department
Government of the Punjab, Lahore
Pakistan

Ahmed M. Alaa
https://orcid.org/0000-0003-4900-3635
Badr University in Cairo, Cairo
Egypt
Contact e-mail: *Ahmed.Alaaelden@buc.edu.eg*

Wael Alharbi
https://orcid.org/0000-0001-7713-5294
Yanbu English Language and Preparatory Year Institute, Yanbu
Saudi Arabia

Issa AlQurashi
https://orcid.org/0009-0008-8610-5327
Department of Foreign Languages
Taif University, Taif
Saudi Arabia

Appendix 1. The Interview Questions

- Do you feel passionate about using ChatGPT?
- Do you think it provides sufficient information to address your query?
- Is it hard to use and rely upon?
- Do you think ChatGPT provides variety in the learning process?
- Do you think ChatGPT is beneficial for improving grammar?
- Do you think ChatGPT is a distraction?

Appendix 2. The Complete Codes

Passion and Engagement: Energetic, passionate, attractive, interactive, fun-learning.

Information Sufficiency: Sufficient knowledge, easy to comprehend, clear explanations, quick feedback.

Ease of Use: Smooth operation, easy to use, accurate, updated.

Learning Enhancement: Eases learning, clarifies concepts, interesting, personalized learning.

Grammar Improvement: Improved grammar, clear explanations, overcoming shyness, effective learning.

Non-Distractive: No distractions, focused learning, relevant information.