



Original Research

AI meets education: How ChatGPT transforms reading skills in Omani EFL learners

by Behnam Behforouz^{1,2} and Ali Al Ghaithi²

¹University of Technology and Applied Sciences, Shinas, Oman

²Sohar University, Sohar, Oman

The main objective of this study was to measure the impact of ChatGPT on the reading skills of language learners. Therefore, a total of fifty Omani students with intermediate English proficiency were selected and randomly assigned into two groups, one control group and an experimental group, with an equal number of students in each group. Both groups received the traditional face-to-face training to engage with the reading comprehension skills, techniques, and strategies in understanding the texts and finding answers for various types of questions arising from the test, but the experimental group received extra explanation and practice from ChatGPT. To compare the results in both groups, the researcher developed and modified some reading tests. Their reliability and validity were measured and monitored by the experts. After a month of treatment, the findings revealed that both groups initially received higher scores in the reading posttests compared to their pretests, but the experimental group performed significantly better than the control group. Additionally, further analysis of the delayed posttests of reading showed that the control group had no increment in their scores while the experimental group continued its progress, and performance was significantly higher, suggesting improved retention of reading abilities. The results of this study are useful for teachers, students, and educational institutions.

KEYWORDS: ChatGPT, reading skills, EFL learners, AI, artificial intelligence, chatbot

CRediT AUTHOR STATEMENT: Behnam Behforouz: Formal Analysis, Writing – Review & Editing, Project Administration. Ali Al Ghaithi: Methodology, Writing – Original Draft, Resources.

CONFLICT OF INTEREST: The authors declared no conflict of interest.

DATA AVAILABILITY STATEMENT: The data supporting the findings of this study are included within the paper, or it could be shared by the corresponding author upon request.

FUNDING: No funding was reported for this research.

ARTICLE HISTORY: Submitted September 17, 2024 | Revised August 10, 2025 | Accepted September 12, 2025

FOR CITATION: Behforouz, B., & Al Ghaithi, A. (2025). AI meets education: How ChatGPT transforms reading skills in Omani EFL learners. *Training, Language and Culture*, 9(3), 10-21. <https://doi.org/10.22363/2521-442X-2025-9-3-10-21>



This is an open access article distributed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0), which allows its unrestricted use for non-commercial purposes, subject to attribution. The material can be shared/adapted for non-commercial purposes if you give appropriate credit, provide a link to the license, and indicate if changes were made.

1. INTRODUCTION

As Kefalaki et al. (2022) highlight, conventional approaches to language education usually depend on standardised educational programmes and instructional strategies that are unable to meet the diverse needs of learners. Instructors in conventional classrooms may lack the time and tools necessary to modify their lessons to fit every student's demands (Rudolph et al., 2023). With this one-size-fits-all strategy, learners might feel behind or uninterested, which can hinder the growth of their

language skills (Hendriani et al., 2023). In addition, receiving functional feedback on the educational performance in the learning process seems to be one of the challenging criteria (Thi Thuy Oanh & Pham, 2023). Additionally, big class sizes impede instructors from providing personalised feedback, leading to students possibly not receiving timely corrections for their linguistic errors, which reduces their chances for skill improvement (Chiu et al., 2023). In these situations, students may feel frustrated, lose their motivation, and struggle to make progress (Su

& Yang, 2023). Given these constraints, there is an imperative need for novel concepts that may provide tailored support and cultivate dynamic environments for language acquisition. AI-driven technologies like ChatGPT in language teaching enable distinctive educational methodologies for foreign languages by providing tailored assistance and interactive activities (Slamet, 2024).

In the area of language learning and teaching, the use of AI tools attracted some attention. AI might change the directions in language education, providing feedback, and professional development (Kasneci et al., 2023). One such AI tool is ChatGPT, a language model created by OpenAI, which can produce text replies comparable to those of humans (Kim, 2023; Tang, 2023). By realising and questioning user preconceptions and responding automatically to inquiries using hundreds of online resources, ChatGPT sets itself apart from other chatbots and virtual assistants (Hassani & Silva, 2023). By employing artificial intelligence algorithms and large-scale language models, ChatGPT facilitates conversational exchanges with users (Biswas, 2023; Haman & Školník, 2023). A deeper exploration of ChatGPT's potential and preferences by academics guides future educational practices and contributes to the growing body of knowledge about how best to employ AI in language acquisition (Lund & Wang, 2023; Su & Yang, 2023). In this regard, ChatGPT's value as an online language-learning helper becomes clear (Biswas, 2023). Unlike more traditional methods, ChatGPT may be adjusted to fit the requirements of learners by providing them with personalised comments and extra chances for practice (Liebrenz et al., 2023; Qureshi et al., 2023). ChatGPT may assist learners in developing their language abilities by focusing on areas where they need the most help, using deep learning algorithms and massive models of language analysis of student input. Moreover, ChatGPT's interactive features could provide a dynamic and fascinating classroom environment (Rudolph et al., 2023; Shen et al., 2023). Using real-time interactions among peers, ChatGPT lets learners hone their language skills in a natural environment (Hassani & Silva, 2023; Seetharaman, 2023).

Reading is an important ability during educational studies. Some people may think that reading is a basic skill, but it is a complicated process that requires several different skills (Rudolph et al., 2023). Students need to improve their vocabulary and analytical skills to do well on reading comprehension tests. Pido and Mubarakah (2024) show that teachers can help students understand what they read better by using skimming and scanning techniques in reading activities, customising instruction for each student, identifying key points, and encouraging fast reading. Since skimming and scanning techniques improve students' understanding of what they read (Basri et al., 2022), skimming literature may help students understand the author's tone or intonation (Mokalu et al., 2022), help them understand the structure of the text, and facilitate easier reading. On the other hand, scanning may help pupils make the most of their time by making it easier to get information from reading (Gulo, 2020). This strategy became obsolete as most learners

experienced challenges in reading. Because of this, the teacher needs to help the students develop useful reading abilities. To engage in skimming and scanning activities, technology tools could be employed, including AI-assisted instruments, to cultivate an effective teaching environment that enhances students' fundamental reading skills (Pido & Mubarakah, 2024). As a result, AI can be used to improve English lessons for students, particularly in reading. It can be used to make learning more interesting and lead to better grades (Li & Xu, 2020).

In a recent study, Kwon and Lee (2023) investigated the accuracy of ChatGPT in responding to English reading comprehension questions on CSAT and TOEFL iBT. The results showed that ChatGPT could accurately answer nearly 69% of the questions provided. When it came to question typologies, ChatGPT accurately responded to approximately 75% of factual/inferential-type queries and more than 87% of fill-in-the-blank and summary-type ones. Nonetheless, its accuracy rate for vocabulary and grammar tests was much lower than this. However, ChatGPT PLUS, the latest model built on GPT-4, had a proficiency rate of 93%, including vocabulary and grammar. With a comparable objective, Ahn (2023) assessed the effectiveness of ChatGPT on CSAT English reading comprehension test items. In the experiment, ChatGPT achieved a 74% accuracy rate in providing the right responses. The study thus proposes that the performance of ChatGPT could be enhanced by improving training techniques, integrating varied and well-balanced datasets, and using human-AI collaboration. The study also specified testing item categories that ChatGPT could enhance, such as finding the optimal sequence of events in a narrative, detecting pronoun referents, and arranging sentences correctly. Shin (2023) examined the capacity of ChatGPT to create reading comprehension items as an assessment developer. The results indicated that some inquiries necessitate certain criteria for their appropriate design. Consequently, the various types of inquiries encompassed the identification of the contextual significance of underlined phrases, the sequencing of sections within a passage, and the determination of the mode of the text. Through the analysis of measures, the study offered tailored prompts for several categories of reading comprehension questions, as well as suggestions for constructing questions when utilising ChatGPT.

Amidst the present combination of doubt and hope regarding AI, especially ChatGPT, it is crucial to acknowledge that the teaching of a second language (L2) requires research focused on the classroom that may showcase the practicality of this technology for L2 instructors and practitioners (Shin & Lee, 2023). Li et al. (2024) stated that the existing literature on the use of ChatGPT failed to cover the necessities and techniques to use it appropriately within the English reading learning and teaching process. Additionally, Al-Otaibi and Al-Homidhi (2025) believed that there is a lack of sufficient studies and references on the integration of ChatGPT and English language learning in higher education. Therefore, the purpose of this study is to evaluate the effectiveness of ChatGPT in enhancing the reading

comprehension abilities of Omani EFL learners at the college level. This study holds significance in the context of EFL education by exploring the potential of AI technologies, specifically ChatGPT, in providing personalised and adaptable learning experiences. Based on the purpose of the study, the following question will be covered in this study: to what extent does the use of ChatGPT as a facilitator improve the reading comprehension abilities of Omani EFL learners compared to traditional face-to-face instruction?

2. THEORETICAL BACKGROUND

2.1. Reading skills and AI integration

Hidayat (2024) examined the effectiveness of customised reading systems based on artificial intelligence in improving reading comprehension for Indonesian senior high school students. The study included 85 students divided into a treatment group and a control group following the traditional curriculum. Using the AI system, modern algorithms generate personalised reading materials based on the student's comprehension level, interests, and learning style. The findings demonstrated that the AI-based platform significantly enhanced reading comprehension of the experimental group, as evidenced by significantly better results compared to the control group. The strength of the study is a clear design of a quasi-experimental approach that aligns the objectives, instruments, and statistics, and explicitly states that AI-based platforms assist learners in improving their reading skills. However, the study suffers from a short intervention period and a lack of clear statistics to measure longer retention.

Kim (2018) examined the impact of chatbots on the English listening and reading skills of 46 college students. The experimental group interacted with a chatbot named Elbot for 16 weeks, participating in 10 sessions focused on their daily lives. The findings revealed that using chatbots could increase the scores of the experimental group in reading comprehension. One of the strengths of this study is the employment of a controlled testing approach via TOEIC tests, which empirically measure the effects of using chatbots on receptive language skills and support the pedagogical values of chatbot-assisted learning in Korean EFL contexts. However, the weaknesses of the study are the small population size, the short treatment period, and the possible Hawthorne effect, which may limit the generalisability of the study.

Xu et al. (2021) conducted a study to examine learners' performance based on reading comprehension skills in two distinct environments: human-to-human interaction and chatbot-assisted settings. The students in the experimental group received guidance on how to comprehend the narrative from the chatbot. Conversely, students in the control group were given identical instructions by a human instructor. The study's findings indicated that the learners who engaged in directed communication with conversational agents completed story comprehension questions more frequently than those who did not engage in guided conversation. One of the strengths of this study is that the randomised experimental design with the preschool

learners to compare the AI conversational agent and partners' narrations could provide better information on the linguistic outcomes, such as productivity, intelligibility, and lexical diversity. However, the lack of longitudinal monitoring, limited demographic variables such as linguistic background, and the lack of information on whether the conversational agents provide social-emotional support are among the study's weaknesses.

Behforouz and Al Ghaithi (2024) assessed the impact of the design and implementation of artificial intelligence on English language acquisition. An interactive chatbot was developed utilising programming languages to evaluate the effect of instruction delivered through the bot on the reading skills of Omani EFL students at the intermediate competency level. The results of the study indicated that the experimental group, which utilised the interactive chatbot as a learning facilitator, outperformed the control group that received solely traditional in-class teaching. The strength of this study is that it uses validated and reliable tests to provide adequate information on the efficacy of interactive chatbots in the reading comprehension process, but its short intervention period, absence of qualitative data, and the single-institution sampling are among the weaknesses of the study.

Jin (2024) examined the incorporation of artificial intelligence technology into English reading instructions in senior high school environments. The study explored several AI apps emphasising the enhancement of reading comprehension abilities, personalised learning experiences, and adaptive evaluation and feedback mechanisms. The results indicated that the incorporation of AI into English reading instruction improved student engagement and motivation through tailored learning experiences and interactive methods. The strength of the study lies in the conceptual framework, which was provided by integrating AI-driven tools into the learning process, supported by comprehensive teaching cases that demonstrate real classroom implementation. The weaknesses of this study include the shortage of empirical data and control group comparisons.

2.2. ChatGPT and reading skills

Wang and Feng (2024) examined the effect of ChatGPT assistance on reading skills over 4 weeks involving 83 Chinese undergraduate students. The learners were divided into two distinct groups, one relying on conventional paper-based reading while the other utilised ChatGPT for reading support. The findings demonstrated that, compared to the paper book group (63% understanding rate and an average rating of 52/100), the ChatGPT-assisted group performed better in their book evaluations, with an average rating of 76/100 and a reading comprehension rate of almost 86%. According to the results, integrating ChatGPT may significantly improve Chinese students' comprehension of English and their interest in reading English books. The lack of long-term effects could be considered a weakness of this study. Zhang et al. (2025) conducted a study on a novel reading platform powered by ChatGPT, which has rapidly gained popularity as a research assistant among students, owing to its

immediate natural language interaction and question-answering functionalities. Sixty-four undergraduate students were recruited for participation in this quasi-experimental investigation. Substantial disparities were identified between the two groups following the intervention. The characteristics of the ChatGPT-based reading platform resulted in the experimental group students experiencing reduced foreign language reading anxiety and cognitive load compared to the control group. Moreover, they surpassed the control group in critical thinking and academic reading performance. The strength of this study was the implementation of a quasi-experimental study with the combination of quantitative and qualitative data supporting a holistic viewpoint of the impact of AI chatbots on reading performance and motivation among Chinese students. However, the small sample size and the lack of measures for the long-term effect of AI tools could be considered as the weaknesses of this study.

Muman (2025) investigated the effect of ChatGPT-OpenAI and its potential to enhance learners' reading skills. The researcher employed the classroom action research approach to gather and analyse the data among vocational high school students. The findings indicate that ChatGPT-OpenAI serves as a practical teaching resource for enhancing English reading proficiency in vocational schools. This research suggests that teachers should recognise the significance of the 21st century, wherein ChatGPT-OpenAI is a crucial tool that educators must grasp to enhance students' reading skills. The strength of this study is its efficacy in demonstrating the potential of using ChatGPT in improving the reading skills of students by providing structured intervention supported statistically and perceptually. However, the small sample size, lack of control group and some technical barriers could be among the weaknesses of the study.

Amimi and Saragih (2025) examined the impact of ChatGPT utilisation and reading literacy on the learning interest of Business Education students from the 2022 cohort at the Faculty of Economics, State University of Medan, among 63 learners. The analysis results indicated that ChatGPT positively influenced students' learning interest, evidenced by a t-count of 6.447 and a significance level of 0.000. Furthermore, reading literacy exerted a beneficial influence, evidenced by a t-count of 5.471 and a significance level of 0.000. The combined utilisation of ChatGPT and reading literacy accounted for 62.3% of students' enthusiasm in learning. Using robust statistical evidence and offering a data-driven foundation to integrate AI and literacy into digital pedagogy is a notable strength of this study. However, using an ex post facto design, a single-institution scope, and reliance on self-reported surveys are the limitations of the study.

3. MATERIAL AND METHODS

This section is dedicated to the selection procedures of the participants for the study, the steps of designing, piloting, and validating the instruments and tests, and the way SPSS 27.0 and JASP 0.95 were used to analyse the data from the performance of the learners.

3.1. Participants

To conduct this quasi-experimental research study, 50 Omani EFL learners from a higher education institution in Oman were randomly assigned to an experimental group and a control group, with 25 students in each group, comprising both males and females. The English proficiency level of these learners was determined to be at the intermediate level based on the university's placement test. These students were native Arabic speakers, and their ages ranged from 18 to 20 years old. These students were studying in the Foundation Programme, in which they had to study some modules on English, Math, and IT, and upon passing the final assessment, they were eligible to move to their specialisations in higher education departments.

3.2. Instruments

3.2.1. Reading tests

To compare and measure the performance of students in both groups on reading skills, three sets of tests, including pretest, posttest, and delayed posttest, were designed. The tests were aligned with three types of questions, including five multiple-choice questions, five true and false questions, five matching questions, three fill-in-the-blank questions, and two short-answer questions. To ensure the reliability of the tests, a pilot study was conducted before the main round of the study with 25 random Omani EFL learners in the same institution and with the same English proficiency level. Table 1 below shows the results of Cronbach's Alpha for the reliability indexes. Following the reliability, the questions were reviewed by two Omani PhD holders in applied Linguistics to validate the questions.

3.2.2. Reading passages and strategies

The reading passages that were practised during the treatment period were selected from NorthStar3, the Reading and Writing, 3rd Edition. Pearson Education Limited designs these books for GCC (Gulf Cooperation Council) countries. Four reading passages with all the questions were selected from this book. Although there are pre-reading to post-reading strategies that could be covered while working on a text, the following few strategies were covered in this study during the treatment.

Prior Knowledge Elicitation. Ask students to either generate or reflect on what they already know about a subject being read.

Reading Objective. The reader sometimes reads to gain specific information, but at other times, the purpose is to understand the main idea. This allows students to focus their attention and be more active with the text.

Predicting. Engage students in making predictions based on titles, headings, or pictures. Prediction leads to curiosity and active participation.

Skimming. Have students skim through the text quickly to get an overview of how the text is structured and what the text is all about before they proceed to its specifications.

Scanning. Have students do the scanning activity to find information such as the names, dates, and facts. Such an activity will provide an easy way to retrieve important information.

Table 1
The results of the reliability index for all sets of tests

TESTS	NUMBER OF ITEMS	CRONBACH'S ALPHA
Pretest	20	0.856
Posttest	20	0.870
Delayed posttest	20	0.869

Contextual Guessing. Teach students how to infer the meaning of unfamiliar words from their surrounding context rather than immediately reaching for a dictionary.

3.3. Ethical Considerations

Before the study began, ethical approvals were received from the Research Department and the related authorities within the institution. In addition, students were informed and signed the consent form, agreeing that their participation in this study is voluntary and will not affect their regular performance or the continuous assessment. It was also emphasised that the results of their performances will remain confidential and will not be shared publicly.

3.4. Procedure

The present investigation was conducted during the autumn semester of 2024-2025. Participants were informed that their participation in the study was voluntary. Prior to the treatment, and to ensure the homogeneity of students' knowledge of reading skills, a pretest of reading was conducted. After that, students were divided into two groups, including a control group and an experimental group. The treatment period lasted for a month, and according to the curriculum and delivery plan, two sessions, each lasting one hour and forty minutes, focused on reading comprehension activities; therefore, eight reading passages were covered during the treatment. The control group received training and instruction on reading skills, finding answers to various questions, and some reading techniques, such as skimming and scanning, with extra practices within the classroom and through traditional face-to-face teaching techniques. The teacher regularly monitored the students in the class through observation and question-and-answer sessions, and provided extra quizzes. On the other hand, although the experimental group received in-class instructions on reading skills and techniques to cope with different types of questions, the students were instructed to use ChatGPT to practice extra reading comprehension activities, such as creating questions, understanding the general idea, and the main idea. All participants in the experimental group were allocated a ChatGPT account established by the researchers to enable the teacher to oversee the students' progress in ChatGPT, verify adherence to instructions, and ensure the completion of assignments as stipulated by the researchers. Subsequently, the investigator facilitated a one-hour workshop to

instruct the treatment groups on utilising ChatGPT for reading skills improvements and to resolve any potential concerns. Various activities were assigned to the students to use ChatGPT outside of the class setting. ChatGPT could help the students with the opportunity to further their practice to pose follow-up questions, seek clarification on complex vocabulary, or even produce summaries. The instructor motivated learners to participate in meaningful conversations with ChatGPT, which posed comprehension inquiries or stimulated students to anticipate the subsequent events in a narrative. To enhance participation, the teacher instructed students to use ChatGPT to generate their inquiries regarding the reading material and then engage in a quiz amongst themselves in pairs or groups during the upcoming class. These exercises facilitated dynamic learning and provided essential support for learners to enhance their reading abilities, providing both structured instruction and independent practice inside the classroom. As an example (see Appendix), students took a paragraph from the text, made a one-sentence guess about its main idea, and then asked ChatGPT to come up with one literal and one inferential question. Students should answer these questions before looking at any key. Additionally, students asked ChatGPT to find 6 to 8 essential words from the same paragraph and requested a definition, two common collocations, one word-family item, and a contrast for each word. In the next step, students asked ChatGPT to make a six-item cloze exercise from the text, and they tried it without the answer key, and then got feedback on whether it was correct or incorrect. Finally, it should set up a one-minute micro-quiz for the next day. The following week after the end of the treatment, a posttest was conducted to compare the performance of both groups before and after the treatment. A delayed posttest was conducted to measure students' knowledge retention and the efficacy of using ChatGPT three weeks after the treatment period.

4. STUDY RESULTS

Before conducting any test for comparison purposes, it was mandatory to measure the normality of data, which was useful in selecting appropriate parametric or non-parametric tests. Therefore, a Kolmogorov-Smirnov normality test was conducted (Table 2). Table 2 shows that for the pretest, the control had a statistic of 0.229 at 0.002 significance, while for the experimental group, it was 0.222 at 0.002. There is a significant deviance from normality in both groups ($p < 0.05$). The posttest

showed that in the control group, the statistic was 0.178 with a significance of 0.040, whereas in the experimental group, the statistic stood at 0.194 with a significance of 0.016. These results

show both groups continue to be deviant from normality. In the delayed posttest, the control group had a statistic of 0.181, p-value=0.035, while the experimental group had a statistic of

Table 2

The results of the data normality in all sets of the tests

	GROUPS	KOLMOGOROV-SMIRNOV		
		STATISTIC	df	Sig.
Pretest	control	.229	25	.002
	experiment	.222	25	.003
Posttest	control	.178	25	.040
	experiment	.194	25	.016
Delayed posttest	control	.181	25	.035
	experiment	.205	25	.008

0.205, p-value of 0.008, showing another significant deviation from normality. Based on the normality results, the non-parametric test seemed a suitable option to measure the results of tests within the control group.

Table 3 below shows the performance of the control group in all sets of tests. Table 3 revealed that an Asymp. Sig. of <.001 and a Friedman Test Chi-Square of 45.238 with 2 degrees of freedom imply that the difference across pretest, posttest, and

delayed posttest scores within the group has reached significance, hence justifying further post-hoc tests in pointing out the exact locations of these differences. Table 4 provides more details on the performance of the control group and shows the differences in the scores across the pretest and posttest, as well as the posttest and the delayed posttest. The Sig. for the pretest and posttest comparison is less than .001, indicating this difference is significant. That means that the scores significantly differ

Table 3

The Friedman test results for the control group

N	Chi-Square	df	Asymp. Sig.
25	45.238	2	.000

across these two testing times. However, the Sig. for the posttest and delayed posttest is .096, which is not statistically significant at $p > 0.05$. This would mean that the difference in performance between the posttest and delayed posttest was not significant. To continue with the analysis of the data, a Wilcoxon signed-rank test was conducted to measure the performance of students in the experimental group. Table 5 that there are statistically significant differences in the experimental group's performance across all sets of tests. In the posttest versus pretest test, the test obtained an observed Z-value of -4.438 with Sig. of .000. This is thus a significant improvement from the pretest to the posttest. Also, the comparison of delayed posttest versus posttest yields a Z-value of -4.476, with a Sig. of .000 again, indicating continued statistical improvement in the delayed posttest statistics. Thus, the results again show consistent progress in the experimental group over time. Table 6 provides more information about the

comparison of these sets showing that there were significant differences between the pretest and posttest and between the posttest and the delayed posttest of the experimental group. The comparison across pre- and posttest shows a statistically significant increase from pretest to posttest since the Sig. is less than .001 with the Std. of -4.953. Moreover, by comparing the posttest to that of the delayed posttest, Sig. 2-tailed is also less than .001 with a standard of -4.650, showing that the difference between these two-time measures is significant. From this finding, it can be suggested that there was a significant improvement in the experimental group from the pretest to the delayed posttest and that there was a measurable increase at each juncture.

To compare the results of the two groups in all the tests together, a Mann-Whitney U Test was conducted, and the results can be seen in Table 7 that revealed no statistically significant difference between the experimental and control groups at the

Table 4
Marginal homogeneity test

	PRETEST & POSTTEST	POSTTEST & DELAYED POSTTEST
Distinct Values	8	6
Off-Diagonal Cases	25	9
Observed MH Statistic	298.000	129.000
Mean MH Statistic	335.500	131.500
Std. Deviation of MH Statistic	7.921	1.500
Std. MH Statistic	-4.734	-1.667
Asymp. Sig. (2-tailed)	.000	.096

Table 5
The results of the performance of experimental group in all tests

	POSTTEST – PRETEST	DELAYED POSTTEST – POSTTEST
Z	-4.438	-4.476
Asymp. Sig. (2-tailed)	.000	.000

Table 6
The results of the marginal homogeneity test

	PRETEST & POSTTEST	POSTTEST & DELAYED POSTTEST
Distinct Values	10	7
Off-Diagonal Cases	25	25
Observed MH Statistic	293.000	438.000
Mean MH Statistic	365.500	458.000
Std. Deviation of MH Statistic	14.637	4.301
Std. MH Statistic	-4.953	-4.650
Asymp. Sig. (2-tailed)	.000	.000

pretest stage ($p = .478$, CI: -0.410 to 0.208), with a negligible effect size ($r = -0.112$), indicating that both groups began at a comparable level. A statistically significant difference was seen at the posttest stage ($p < .001$, CI: 0.665 to 0.896), with a considerable effect size ($r = 0.810$, CI: 0.988 to 0.997), demonstrating a notable impact of the intervention on the performance of the experimental group.

The difference grew even bigger in the delayed posttest ($p < .001$), and the effect size grew to 0.994 , which means that the experimental group got a significant and long-lasting benefit over time.

5. DISCUSSION

The study objective was to examine the impact of using ChatGPT on reading comprehension. 50 Omani EFL learners were split into two groups, each comprising 25 students, namely the control group and the experimental group. Each group received distinct methodologies of instruction: the control group received conventional face-to-face teaching, while the experimental group received ChatGPT-enhanced learning. Assessment of reading skills performance and retention was conducted by administering three reading tests, including a pretest, a posttest, and a delayed posttest. After one month of teaching and

Table 7
The results of the comparison of both groups in all tests

							95% CI for Effect Size	
	Test	Statistic	df	p	Effect Size	SE Effect Size	Lower	Upper
pretest	Student	0.716	48	.478	0.202	0.284	-0.355	0.757
	Mann-Whitney	347.500		.478	-0.112	0.163	-0.410	0.208
posttest	Student	-6.710	48	< .001	-1.898	0.390	-2.562	-1.220
	Mann-Whitney	59.500		< .001	0.810	0.163	0.665	0.896
delayedposttest	Student	-10.610	48	< .001	-3.001	0.510	-3.809	-2.178
	Mann-Whitney	2.000		< .001	0.994	0.163	0.988	0.997

subsequent testing, the results showed that although both groups showed progress from pretest to posttest, the experimental group performed better than the control group. In addition, the comparison of the delayed posttest between the groups revealed that the control group showed no progress, while the results of the experimental group were significantly positive. The research findings indicated that the use of ChatGPT had a substantial positive impact on reading comprehension and retention after exposure.

The only improvement from the pretest to the posttest was seen in the control group that got face-to-face instruction from a teacher. This improvement can be understood through Vygotsky's (1978) Sociocultural Theory and the concept of the Zone of Proximal Development (ZPD), which emphasises that learners progress most effectively when supported by a more knowledgeable individual through social interaction and scaffolding. This study indicates that the teacher's guided practice, direct interactions, and prompt feedback likely offered scaffolding that facilitated students in attaining enhanced levels of reading comprehension on the posttest. This aligns with Rosenshine's (1987) Principles of Direct Instruction, which stress the necessity of structured, sequential teaching coupled with active oversight and feedback for prompt enhancements. However, without extended practice or adaptive reinforcement, the control group showed no improvement in the delayed posttest.

The experimental group that used ChatGPT to learn outperformed the control group on both the posttest and the delayed posttest. This aligns with Self-Determination Theory, which posits that fulfilling learners' needs for autonomy, competence, and relatedness boosts intrinsic motivation and ongoing engagement (Deci & Ryan, 2000). ChatGPT's personalised feedback and flexible coaching may have bolstered students' perceptions of competence and autonomy, thereby prolonging their drive. Constructivist Learning Theory (Bruner, 1996; Piaget, 1972) asserts that ChatGPT's interaction facilitates learners in actively constructing meaning and integrating new knowledge with pre-existing understanding, rather than only taking

information passively. Cognitive Load Theory asserts that reducing superfluous cognitive load enables learners to allocate more cognitive resources to the learning of essential knowledge (Sweller, 1988). This study indicates that ChatGPT's prompt clarifications and targeted responses may have mitigated unnecessary cognitive burden, hence enhancing learning efficacy. The experimental group's persistent dominance in the delayed posttest exemplifies Spaced Retrieval Practice Theory, since ChatGPT enabled frequent, spaced interactions with reading skills, hence improving long-term recall (Cepeda et al., 2006). The findings of this study align with prior research, including a study by Miller (2019), which demonstrates that AI systems deliver reading content tailored to the reader's comprehension level for individualised learning methods. Liu et al. (2019) conducted a similar study that showed AI can find problems that students have when they read, which allows teachers to give each student personalised help. In a similar study, Salam et al. (2023) showed that the integration of ChatGPT into the educational process significantly impacts students, as seen by improvements in reading proficiency. Moreover, Behforouz and Al Ghaithi (2024) investigated the role of an interactive chatbot as a facilitator in reading, illustrating that chatbots can serve as advantageous tools in education, especially in improving reading proficiency. Sosa Daza et al. (2024) examined ChatGPT-based interventions to enhance English reading proficiency among senior high school students. The main findings showed that 84.7% of teachers were very interested in using ChatGPT for teaching purposes.

6. CONCLUSION

The educational applications of ChatGPT have demonstrated considerable promise, namely in augmenting individualised learning and delivering immediate feedback to students, therefore enabling them to participate in self-directed learning (Kasneci et al., 2023). An illustration of its use in language instruction is its capacity to assist pupils in enhancing their grammar and writing abilities through interactive conversation (Zhai,

'By integrating AI into higher learning and further utilising ChatGPT, the outcomes of educational institutions could be scaled up. It extends the learning experience beyond traditional class hours and allows for more dynamism through the integration of technology in education, which should ultimately translate into better academic performance and put schools at the forefront of modern teaching methodologies'

2023). To this end, this paper's focus was on the integration of ChatGPT in English language classes to improve the reading skills of Omani EFL learners. The comparison of the performance in both groups in posttest and delayed posttest revealed that the experimental group outperformed significantly in all of the tests. This could be attributed to the involvement with ChatGPT outside of the educational setting, which also played a role in the enduring advantages noted in the experimental group.

The results of this study could be helpful for teachers, students, and institutions. For teachers, using ChatGPT will improve pedagogic methods, the ability to provide feedback in a highly individualised form and to support self-directed learning. Students can learn alone and receive errors corrected immediately. Students benefit from increased engagement and more frequent opportunities for interactive practice, resulting in better retention and understanding. By integrating AI into higher learning and further utilising ChatGPT, the outcomes of educational institutions could be scaled up. It extends the learning experience beyond traditional class hours and allows for more dynamism through the integration of technology in education, which should ultimately translate into better academic performance and put schools at the forefront of modern teaching methodologies.

Despite the encouraging results, some constraints of the present study must be recognised. The sample size of 50 participants was relatively small, which may limit the generalisability

to larger populations of EFL learners. Besides, this research was focused on only Omani learners; therefore, cultural or linguistic variables pertinent to them could have intervened in the results. Another limitation was that the intervention period was short, lasting only one month; this might not be enough to gauge the long-term effects or sustainability of improvements in reading comprehension. In addition, this study depended primarily on reading tests for assessment; hence, other aspects of language learning had not been taken into consideration, like writing or speaking. The other limitation of this study could be a lack of qualitative data on the learners' attitudes toward using AI tools within the learning context. Additionally, there was no analysis on the impact of teachers' and students' digital literacy, and local educational norms in this study. Some factors, such as students' exposure to AI tools and their general knowledge about using technology, were not considered in this study, which could affect the findings. Finally, factors such as students' motivation and perceptions toward using technology, that could lead to changes in the results, were not covered in this study.

Therefore, further studies could involve a larger sample of students representing diverse English proficiency levels and a range of institutions, in order to gain a more detailed understanding of AI adoption in Oman. To evaluate the effect of ChatGPT and other AI tools for long-term applicability in the learning process, studies with more treatment time are suggested. Additionally, more studies are suggested to measure other main skills and subskills in language learning, such as writing, speaking, and grammar.

Moreover, future studies should use various types of evaluation instruments besides reading comprehension tests to offer a more complete account of language competence. Further studies based on a qualitative research design could provide a more precise picture of AI-enhanced learning within the Omani EFL learning context. An examination of external elements, such as student motivation, attitude, previous technological experience, and the influence of instructor facilitation, would explain more clearly how these variables impact the efficacy of AI tools in language acquisition.

References

- Ahn, Y. Y. (2023). Performance of ChatGPT 3.5 on CSAT: Its potential as a language learning and assessment tool. *Journal of the Korea English Education Society*, 22(2), 119-145.
- Al-Otaibi, M. K. I., & Al-Homidhi, K. M. (2025). Teaching program based on ChatGPT and its effectiveness in developing reading comprehension skills for English language students at university. *International Journal of Language and Literary Studies*, 7(1), 137-154. <https://doi.org/10.36892/ijlls.v7i1.1972>
- Amimi, F. S., & Saragih, L. S. (2025). The influence of ChatGPT usage and reading literacy on the learning interest of Business Education students of the 2022 Faculty of Economics, State University of Medan. *Economic Journal Economic and Business*, 4(3), 490-497. <https://doi.org/10.56495/ejeb.v4i3.1197>
- Basri, M., Daim, N., & Martora, J. (2022). Teaching reading comprehension in narrative text through skimming and scanning. *Pacific Journal of Education*, 1(1), 27-32. <https://doi.org/10.51135/jukip.v1i1.7>
- Behforouz, B., & Al Ghaithi, A. (2024). Investigating the effect of an interactive educational chatbot on reading comprehension skills. *International Journal of Engineering Pedagogy (ijEP)*, 14(4), 139-154. <https://dx.doi.org/10.3991/ijep.v14i4.48461>
- Biswas, S. S. (2023). Role of Chat GPT in public health. *Annals of Biomedical Engineering*, 51(5), 868-869. <https://dx.doi.org/10.1007/s10439-023-03172-7>
- Bruner, J. (1996). What we have learned about early learning. *European Early Childhood Education Research Journal*, 4(1), 5-16. <https://doi.org/10.1080/13502939685207811>

- Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, 132(3), 354-380. <https://doi.org/10.1037/0033-2909.132.3.354>
- Chiu, T. K. F., Moorhouse, B. L., Chai, C. S., & Ismailov, M. (2023). Teacher support and student motivation to learn with Artificial Intelligence (AI) based chatbot. *Interactive Learning Environments*, 32, 3240-3256. <https://dx.doi.org/10.1080/10494820.2023.2172044>
- Deci, E. L., & Ryan, R. M. (2000). The 'what' and 'why' of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11(4), 227-268. https://doi.org/10.1207/S15327965PLI1104_01
- Gulo, S. N. (2020). The benefits of skimming technique in reading comprehension for the second semester students of Sanata Dharma University. *UC Journal: ELT, Linguistics and Literature Journal*, 1(1), 1-13.
- Haman, M., & Školník, M. (2023). Behind the ChatGPT hype: Are its suggestions contributing to addiction? *Annals of Biomedical Engineering*, 51(6), 1128-1129. <https://dx.doi.org/10.1007/s10439-023-03201-5>
- Hassani, H., & Silva, E. S. (2023). The role of ChatGPT in data science: How AI-assisted conversational interfaces are revolutionizing the field. *Big Data and Cognitive Computing*, 7(2), Article 62. <https://doi.org/10.3390/bdcc7020062>
- Hendriani, S., Yunita, W., & Putra, H. E. (2023). EFL learners' preference of grammar learning model amid covid-19 pandemic: A mixed-methods study. *International Journal of Instruction*, 16(2), 853-870. <https://doi.org/10.29333/iji.2023.16245a>
- Hidayat, M. T. (2024). Effectiveness of AI-based personalized reading platforms in enhancing reading comprehension. *Journal of Learning for Development*, 11(1), 115-125. <https://doi.org/10.56059/jl4d.v11i1.955>
- Jin, M. (2024). Leveraging artificial intelligence for enhanced English reading instruction in senior high school. *Academic Journal of Management and Social Sciences*, 7(3), 29-32. <https://doi.org/10.54097/pekwxsc50>
- Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... & Hüllermeier, E. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, Article 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Kefalaki, M., Diamantidaki, F., & Rudolph, J. (2022). Technology and education. Innovation and hindrances. *Journal of Applied Learning & Teaching*, 5(1), 6-11. <https://dx.doi.org/10.37074/jalt.2022.5.s1.1>
- Kim, N. Y. (2018). A study on chatbots for developing Korean college students' English listening and reading skills. *Journal of Digital Convergence*, 16(8), 19-26. <https://dx.doi.org/10.14400/JDC.2018.16.8.019>
- Kim, S.-G. (2023). Using ChatGPT for language editing in scientific articles. *Maxillofacial Plastic and Reconstructive Surgery*, 45(1), Article 13. <https://doi.org/10.1186/s40902-023-00381-x>
- Kwon, S. K., & Lee, Y. T. (2023). Investigating the performance of generative AI ChatGPT's reading comprehension ability. *Journal of the Korea English Education Society*, 22(2), 147-172.
- Li, M., & Xu, H. (2020). AI-driven language apps and their impact on traditional language learning methods. *Journal of Computer Assisted Learning*, 36(4), 561-574.
- Li, W., Tang, Y., & Wang, S. (2024). Revolutionizing college English reading with ChatGPT: Current predicaments and future directions. *Transactions on Social Science, Education and Humanities Research*, 11, 53-57. <https://dx.doi.org/10.62051/jnmmb52>
- Liebrezn, M., Schleifer, R., Buadze, A., Bhugra, D., & Smith, A. (2023). Generating scholarly content with ChatGPT: Ethical challenges for medical publishing. *The Lancet Digital Health*, 5(3), e105-e106. [https://dx.doi.org/10.1016/S2589-7500\(23\)00019-5](https://dx.doi.org/10.1016/S2589-7500(23)00019-5)
- Liu, Y., Han, F., Li, F., & others. (2019). Inkjet-printed unclonable quantum dot fluorescent (anti-counterfeiting labels with artificial intelligence authentication). *Nature Communications*, 10, Article 2409. <https://dx.doi.org/10.1038/s41467-019-10406-7>
- Lund, B. D., & Wang, T. (2023). Chatting about ChatGPT: How may AI and GPT impact academia and libraries? *Library Hi Tech News*, 40(3), 26-29. <https://dx.doi.org/10.1108/LHTN-01-2023-0009>
- Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, 267, 1-38. <https://doi.org/10.1016/j.artint.2018.07.007>
- Mokalu, P. V. V., Oroh, E. Z., & Tuerah, I. J. C. (2022). Improving students' reading comprehension through skimming and scanning techniques at the tenth grade of SMA. *Kompetensi*, 1(8), 699-709.
- Muman. (2025). Enhancing English reading skills through ChatGPT-OpenAI: A study on vocational high school students. *Jurnal Pemberdayaan Masyarakat*, 4(1), 210-219. <https://doi.org/10.46843/jpm.v4i1.406>
- Piaget, J. (1972). *The psychology of the child*. Basic Books.
- Pido, N. W., & Mubarakah, A. (2024). The effectiveness of using skimming and scanning techniques in improving reading comprehension: A systematic literature review. *VISA: Journal of Vision and Ideas*, 4(2), 995-1001. <https://doi.org/10.47467/visa.v4i2.2885>
- Qureshi, R., Shaughnessy, D., Gill, K. A. R., Robinson, K. A., Li, T., & Agai, E. (2023). Are ChatGPT and large language models 'the answer' to bringing us closer to systematic review automation? *Systematic Reviews*, 12(1), Article 72. <https://doi.org/10.1186/s13643-023-02243-z>
- Rosenshine, B. (1987). Explicit teaching and teacher training. *Journal of Teacher Education*, 38(3), 34-36. <https://dx.doi.org/10.1177/002248718703800308>
- Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching*, 6(1), 342-363.
- Salam, M. Y., Mahayuni, M., Taman, M., & Mudinillah, A. (2023). Using artificial intelligence for education in the Education 5.0 era to improve reading skills. *Arabiyat: Jurnal Pendidikan Bahasa Arab dan Kebahasaaraban*, 10(2), 149-162. <https://doi.org/10.15408/a.v10i2.35487>
- Seetharaman, R. (2023). Revolutionizing medical education: Can ChatGPT boost subjective learning and expression? *Journal of Medical Systems*, 47(1), Article 61. <https://doi.org/10.1007/s10916-023-01957-w>
- Shen, Y., Heacock, L., Elias, J., Hentel, K. D., Reig, B., Shih, G., & Moy, L. (2023). ChatGPT and other large language models are double-edged swords. *Radiology*, 307(2), 1-4.

- Shin, D. (2023). A case study on English test item development training for secondary school teachers using AI tools: Focusing on ChatGPT. *Language Research*, 59(1), 21-42. <https://doi.org/10.30961/lr.2023.59.1.21>
- Shin, D., & Lee, J. H. (2023). Can ChatGPT make reading comprehension testing items on par with human experts? *Language Learning & Technology*, 27(3), 27-40. <https://dx.doi.org/10.64152/10125/73530>
- Slamet, J. (2024). Potential of ChatGPT as a digital language learning assistant: EFL teachers' and students' perceptions. *Discover Artificial Intelligence*, 4(46). <https://dx.doi.org/10.1007/s44163-024-00143-2>
- Sosa Daza, E. P., Macías Angulo, F. del R., & Ramírez Lozada, H. (2024). ChatGPT-based didactic strategies to improve students' English language reading and writing skills. *Revista Científica y Académica*, 4(1), 390-420. <https://dx.doi.org/10.61384/r.c.a.v4i1.108>
- 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://dx.doi.org/10.1177/20965311231168423>
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257-285. [https://doi.org/10.1016/0364-0213\(88\)90023-7](https://doi.org/10.1016/0364-0213(88)90023-7)
- Tang, G. (2023). Academic journals cannot simply require authors to declare that they used ChatGPT. *Irish Journal of Medical Science*, 192(6), 3195-3196.
- Thi Thuy Oanh, T., & Pham, V. P. H. (2023). The effects of online peer feedback on students' writing skills during corona virus pandemic. *International Journal of Instruction*, 16(1), 881-896. <https://doi.org/10.29333/iji.2023.16149a>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Wang, X., & Feng, Y. (2024). An experimental study of ChatGPT-assisted improvement of Chinese college students' English reading skills: A case study of *Dear Life*. In *Proceedings of the 15th International Conference on Education Technology and Computers (ICETC '23)* (pp. 21-26). Association for Computing Machinery. <https://dx.doi.org/10.1145/3629296.3629300>
- Xu, Y., Wang, D., Collins, P., Lee, H., & Warschauer, M. (2021). Same benefits, different communication patterns: Comparing children's reading with a conversational agent vs. a human partner. *Computers & Education*, 161, Article 104059. <https://doi.org/10.1016/j.compedu.2020.104059>
- Zhai, X. (2023). Exploring ChatGPT in language education: Benefits and challenges. *Journal of Educational Technology*, 56(2), 175-189.
- Zhang, Y., Lai, X., Yi, S., Xu, W., Zhao, Y., & Deng, Q. (2025). Does ChatGPT-based reading platform impact foreign language paper reading? Evidence from a quasi-experimental study on Chinese undergraduate students. *Education and Information Technologies*, 30, 9737-9754. <https://doi.org/10.1007/s10639-024-13190-0>

Appendix

Sample conversation between a user and ChatGPT

User: Here is the paragraph: 'Every Saturday, a group of neighbors meets to clean the small river near their homes. They wear gloves and collect plastic bottles, cans, and other trash. The city gives them free bags and shows them how to recycle the materials. After three months, the water looks clearer, and more birds come to the trees. The group also teaches children about keeping the area clean. They post simple tips online, like 'bring your own bottle' and 'use fewer plastic bags.' The neighbors feel proud because the area is safer and more beautiful.'

ChatGPT: Thank you. Please write one sentence that guesses the main idea.

User: Neighbours clean the river every week, and the area becomes cleaner and better.

ChatGPT: Good. Now two questions.

1. Literal: Who gives the group free bags?
2. Inferential: Why do more birds come to the trees?

User: 1. The city gives the free bags. 2. Because the river is cleaner and safer for birds.

ChatGPT: 1. Correct. The text says, 'The city gives them free bags.' 2. Good inference. The text says, 'the water looks clearer,' which supports your idea.

User: Please find 6 priority words and give for each: meaning in this context, two collocations, one word-family item, and one contrast.

ChatGPT: Here are 6 words.

1. **collect** – meaning: pick up and bring together. Collocations: collect trash, collect bottles. Word-family: collection. Contrast: scatter.
2. **recycle** – meaning: process used items so they can be used again. Collocations: recycle materials, recycle plastic. Word-family: recycling. Contrast: waste.
3. **clearer (clean)** – meaning: more clean and easy to see. Collocations: clearer water, clearer view. Word-family: clarity. Contrast: dirty.
4. **proud** – meaning: happy because of good work. Collocations: feel proud, proud of the team. Word-family: pride. Contrast: ashamed.
5. **safer (safe)** – meaning: with less risk of danger. Collocations: safer area, feel safe. Word-family: safety. Contrast: dangerous.
6. **beautiful** – meaning: very nice to see. Collocations: beautiful place, beautiful view. Word-family: beauty. Contrast: ugly.

User: Make a six-item cloze (gap-fill) from the same paragraph. Hide the key.

ChatGPT: Cloze (fill in one word for each blank):

1. Every Saturday, a group of _____ meets to clean the small river.
2. They wear gloves and _____ plastic bottles, cans, and other trash.
3. The city gives them free bags and shows them how to _____ the materials.
4. After three months, the water looks _____, and more birds come.
5. The neighbors feel _____ because the area is safer and more beautiful.
6. They post simple tips online, like 'bring your own _____.'

User: My answers:

1. neighbors. 2. collect. 3. recycle. 4. clearer. 5. proud, safer. 6. bottle.

ABOUT THE AUTHORS

Behnam Behforouz

PhD, Lecturer

English Department, Preparatory Studies Centre, University of Technology and Applied Sciences, Shinas, Oman

Postal address: 324, Shinas, Oman

Email: Behnam.Behforouz@utas.edu.om

ORCID ID: <https://orcid.org/0000-0002-0078-2757>

Ali Al Ghaithi

Lecturer

English Department, Sohar University, Sohar, Oman

Postal address: 311, Sohar, Oman, Al Jamiah Street

Email: Aghaithi@su.edu.om

ORCID ID: <https://orcid.org/0000-0002-9653-508X>