

Exploring ChatGPT's Role in Boosting Enhancing English as a Foreign Language (EFL) Reading Skills: A Gender and age Analysis

¹ Mehran Davaribina

² Hossein Siahpoosh

³ Masoomeh Maleki*

Research Paper

IJEAP- 2507-2156

Received: 2025-07-20

Accepted: 2025-11-21

Published: 2025-12-18

Abstract: Artificial intelligence (AI) has emerged as a powerful force reshaping many fields, and education is increasingly benefiting from its innovative potential. This study explores how AI-enhanced reading materials—specifically those supported by ChatGPT—affect the reading comprehension of pre-intermediate EFL learners, compared to traditional text-based approaches. The study involved 62 participants aged 13–16 (28 males and 34 females), selected using an institute-specific placement test aligned with CEFR B1 standards. They were randomly assigned to either a control group (n = 31), which used traditional reading methods, or an experimental group (n = 31), which engaged with ChatGPT-enhanced materials. Both groups completed 15 reading sessions over two months. Reading comprehension was assessed using the TOEFL Junior (2010) pre- and post-tests. Statistical analysis employing ANCOVA indicated a significant improvement in post-test reading comprehension scores for the group that utilized ChatGPT, compared to the traditional instruction group. Paired t-tests revealed significant improvements in reading comprehension from pre-test to post-test within the ChatGPT-assisted group, whereas the traditional group showed no statistically significant gains. ANOVA results indicated no significant interaction between gender and instructional modality, and age did not moderate the impact of instructional methods on reading comprehension outcomes. Comparisons of age and gender distributions revealed no significant differences between groups. The findings suggest that AI-driven learning tools, especially ChatGPT, are more effective in improving reading comprehension skills among pre-intermediate EFL learners than traditional methods, offering increased motivation, personalized learning opportunities, and flexible feedback.

Keywords: Artificial Intelligence, ChatGPT, EFL Learning, Gender, Reading Comprehension

Introduction

The integration of Artificial Intelligence (AI) into educational settings has ushered in transformative advancements that are reshaping how learners interact with content and acquire knowledge. AI's influence extends across various facets of education, with its impact on reading comprehension emerging as a particularly significant development (Chen et al., 2020; Yang et al., 2022). As educational institutions increasingly seek to harness AI's potential to enhance learning outcomes, understanding its effectiveness in specific contexts becomes essential. In the Iranian EFL context, where large class sizes and limited access to interactive resources often hinder personalized

¹ Assistant Professor in TEFL, davaribina@gmail.com; Department of English, Ard. c., Islamic Azad University, Ardabil, Iran.

² Assistant Professor in TEFL, siahpoosh_hossein@yahoo.com; Department of English, Ard. c., Islamic Azad University, Ardabil, Iran.

³ Ph.D. Candidate in TEFL (Corresponding Author), masoomehmaleki8124@gmail.com; Department of English, Ard. C., Islamic Azad University, Ardabil, Iran.

instruction, AI tools like ChatGPT offer a scalable solution to foster individualized learning and engagement.

Reading comprehension is a fundamental component of language learning, especially for intermediate learners transitioning from basic to more complex textual understanding. Traditional reading comprehension methods, while foundational, often lack adaptability, individualized support, and immediate feedback. In many EFL classrooms, students rely heavily on teacher-centered instruction and rote memorization, which limits opportunities for critical engagement and independent reading practice (Heffernan et al., 2014; VanLehn, 2011). These limitations often result in reduced learner motivation and minimal differentiation across proficiency levels. In Iran, these challenges are compounded by resource constraints and traditional pedagogical approaches, underscoring the need for innovative tools to enhance reading comprehension.

ChatGPT, as a leading AI-powered language model, utilizes natural language processing to simulate human-like interaction and provide learners with personalized prompts, comprehension questions, and corrective feedback. It can adjust the complexity of texts and questions based on learner performance, allowing for a more dynamic and responsive learning environment (Huang et al., 2020; Kumar & Rose, 2011). For example, ChatGPT analyzes learners' responses in real-time to generate tailored questions targeting specific skills, such as inference or vocabulary recognition, and provides immediate corrective feedback to address errors in understanding or language use. Such features enable learners to actively engage with texts and receive instant guidance, addressing many shortcomings of static reading materials. To investigate these benefits, this study employs a quasi-experimental design, comparing ChatGPT-assisted reading tasks with traditional textbook-based methods through pre- and post-test assessments of reading comprehension over a six-week intervention period.

Recent research has demonstrated the transformative potential of AI in education. Luckin et al. (2022) emphasize "AI's capacity to revolutionize language education by enhancing personalized learning, highlighting how adaptive systems can tailor instruction to individual learning styles" (p. 87). Similarly, Kasneci et al. (2023) document how tools like ChatGPT provide customized feedback and adapt to learner needs, significantly enhancing language acquisition outcomes (p. 615). This builds on earlier work by Hwang et al. (2020), who demonstrated "how AI-driven platforms create more engaging and interactive learning environments through personalized content delivery" (p. 1042). Despite these advancements, significant research gaps remain, particularly regarding how AI-driven tools like ChatGPT impact reading comprehension across gender and age differences, as these factors may influence learners' engagement with technology in distinct ways. Moreover, there is a notable lack of empirical studies exploring the effectiveness of such tools in the Iranian EFL context, where traditional teaching methods and resource constraints limit personalized instruction. This study addresses these gaps by investigating the impact of ChatGPT-assisted reading instruction on Iranian pre-intermediate EFL learners, with a specific focus on the moderating effects of gender and age.

Similarly, Hamuddin (2018) discuss how AI-driven tools foster more engaging and interactive learning environments, thereby enhancing student motivation and participation. ChatGPT exemplifies these advancements through its user-friendly interface and adaptive capabilities. In addition, recent studies on adaptive learning technologies (e.g., Luckin et al., 2022) highlight the role of AI in creating flexible and accessible learning environments.

Despite these promising developments, the effectiveness of AI-driven tools like ChatGPT in relation to gender and age differences remains underexplored. Existing studies, such as those by Johnson et al. (2021), suggest that gender can influence how students interact with technology and respond to various learning tools. This study aims to address these gaps by evaluating whether ChatGPT impacts reading comprehension differently for male and female learners and considers the role of age in moderating these effects. Despite the growing global interest in AI-assisted learning, research in the Iranian EFL context remains limited. English instruction in Iran often depends on traditional, textbook-centered practices that provide little room for individualized feedback or interactive engagement. Consequently, the use of AI tools such as ChatGPT may offer a promising

alternative for improving reading comprehension. However, few empirical studies have examined the effectiveness of AI-based instruction in Iranian classrooms, particularly concerning learners' age and gender differences. To address these gaps, the present experimental study investigates the impact of ChatGPT-assisted reading instruction compared with traditional methods among Iranian pre-intermediate EFL learners.

More specifically, the present study adopted a quasi-experimental, pretest–posttest control group design to compare the effectiveness of two instructional approaches over a six-week period: (1) an experimental group receiving ChatGPT-assisted reading instruction (personalized texts, adaptive comprehension questions, and immediate AI-generated feedback) and (2) a control group using traditional textbook-based reading materials and teacher-led comprehension activities. Sixty Iranian pre-intermediate EFL learners aged 13–16 from two intact classes participated in the study. Reading comprehension was measured using parallel forms of a validated reading test both before and after the intervention.

Recent contributions to this journal have underscored the transformative role of technology-mediated and AI-driven interaction in the Iranian EFL context (Esmaeily & Mahdavi Zafarghandi, 2025; Mirsanjari, 2025). While Esmaeily and Mahdavi Zafarghandi (2025) demonstrated the efficacy of ChatGPT-generated feedback in enhancing teachers' reflective practice, Mirsanjari (2025) highlighted the benefits of dialogic scaffolding within digital environments for improving writing proficiency. Extending this line of research from teacher development and writing instruction to reading comprehension, the present study investigates the direct impact of ChatGPT-assisted reading instruction on pre-intermediate EFL learners, with particular attention to potential moderating effects of gender and age.

Accordingly, the current study seeks to fill this gap by addressing the following research questions (RQs):

Research Question One: Do both Chat-GPT-assisted and traditional methods significantly improve reading task performance over time (from pre-test to post-test) within each group?

Research Question Two: Is there a significant difference in reading comprehension between learners who use AI-assisted materials (Chat-GPT) and those who use traditional methods?

Research Question Three: Is there a significant interaction between the instructional modality (Chat-GPT-assisted vs. traditional) and gender (male vs. female) in influencing reading comprehension?

Research Question Four: Does age moderate the relationship between reading comprehension and the two instructional modalities?

The four research questions were designed separately to avoid double-barreled questions for gender and age and to precisely examine different aspects of the intervention (within-group, between-group, and moderating effects).

Literature Review

The Role of AI in Education

Artificial Intelligence (AI) has emerged as a transformative force in education, offering more than incremental enhancements by fundamentally reshaping instructional delivery and learning experiences. AI's capacity to analyze extensive data and identify learner patterns has enabled the creation of adaptive and personalized learning environments previously unattainable (Luckin et al., 2022; Zawacki-Richter et al., 2019). A prominent example is the development of Intelligent Tutoring Systems (ITS), which simulate individualized tutoring by adapting instruction to learners' real-time needs and progress. Unlike traditional classroom methods—often limited by fixed curricula and large class sizes—ITS can diagnose misconceptions and provide targeted practice and feedback, thereby supporting learner autonomy and mastery (VanLehn, 2011).

Beyond ITS, adaptive learning platforms have gained attention for their dynamic adjustment of instructional content based on learners' performance. Rooted in Vygotsky's concept of the Zone of Proximal Development (ZPD), these platforms maintain instructional content within an optimal challenge range, promoting deeper cognitive engagement and retention (Shute & Rahimi, 2017; Woolf, 2010). Such systems are particularly relevant in language learning contexts, where learner diversity in proficiency and pace requires differentiated instruction to optimize comprehension outcomes.

Natural Language Processing (NLP) tools, another subset of AI, have shown promise in enhancing language and literacy development. Through real-time analysis and generation of human language, NLP tools facilitate authentic interaction and immediate feedback—elements that are essential for language acquisition (Baker & Siemens, 2014). AI-driven conversational agents and chatbots, such as ChatGPT, exemplify this capacity by engaging learners in contextualized dialogue, encouraging reflective thinking and deeper processing of content (Huang et al., 2020). While these tools are praised for their adaptability and responsiveness, it is important to critically acknowledge that their effectiveness can vary based on cultural, linguistic, and contextual factors not fully explored in earlier studies.

AI also offers transformative potential through large-scale data analytics. By continuously monitoring learner engagement and performance, AI tools provide actionable insights for educators to refine instructional strategies and detect learners at risk (Baker, 2016; Luckin, 2018; Siemens, 2013). This data-driven personalization supports a shift from standardized teaching to individualized pathways, aligning with learner-centered pedagogies. Additionally, AI's scalability addresses systemic challenges such as teacher shortages and resource disparities, enabling access to quality instruction across diverse geographic and socioeconomic contexts (Holmes et al., 2019).

AI and Reading Comprehension

Reading comprehension, as a complex cognitive skill, is central to language proficiency. Traditional teaching methods, though foundational, often struggle to accommodate learners' individual differences, leading to varying degrees of engagement and achievement (Chen et al., 2020). AI-powered tools have emerged to bridge this gap by offering adaptive and interactive learning experiences that personalize content to each learner's needs and proficiency level (Huang et al., 2020).

Empirical studies highlight AI's potential to enhance comprehension outcomes through real-time feedback, scaffolded support, and interactive dialogue (Yang et al., 2022). For instance, AI systems can guide learners through texts by clarifying meanings, asking inferential questions, and suggesting strategies for deeper understanding (Nurjaya et al., 2024; Shute & Rahimi, 2017). These adaptive interventions foster critical thinking and metacognitive awareness, which are crucial for processing complex texts (Alan, 2023). However, some studies caution that the effectiveness of such tools depends on learners' prior knowledge, language background, and the cultural appropriateness of the AI-generated content—areas that remain underexplored in diverse EFL contexts.

Comparative Studies: AI vs. Traditional Methods

Research comparing AI-driven instruction to traditional methods often reports positive outcomes in reading comprehension and vocabulary acquisition. Qiao and Zhao (2023) found significant improvement in comprehension among students using AI tools, attributing gains to personalized feedback and real-time adaptation.

Nevertheless, scholars emphasize that these advantages are not universal. Factors such as the quality of AI systems, learners' attitudes toward technology, and classroom integration strategies can moderate outcomes (Hamuddin, 2018). Importantly, while many studies show improved engagement, others note potential challenges, including cognitive overload from interactive features and learners'

preference for human guidance in interpreting nuanced texts. Such mixed findings highlight the need for context-specific research, particularly in underrepresented settings like Iranian EFL classrooms.

Gender and Educational Technology

Demographic factors, including gender, can shape learners' interactions with AI tools. Prior research suggests gender differences in technology acceptance, learning preferences, and perceived ease of use (Venkatesh & Morris, 2000; Wong, 2019). For example, Fathi et al. (2023) found that female learners may exhibit higher anxiety toward unfamiliar technologies, while male learners often demonstrate greater initial enthusiasm for interactive digital tools. Meta-analytic evidence also indicates that males generally hold more favorable attitudes toward technology use, with only minimal reduction in the gender attitudinal gap over recent decades (Cai et al., 2017).

Age and Educational Technology

Age-related factors also moderate the impact of AI on learning. Younger learners may be more receptive to gamified, interactive features, aligning with their developmental preference for visual and experiential learning (Chen et al., 2021; Hwang et al., 2020). Conversely, older adolescents may prefer structured content and explicit instruction (Johnson & Brown, 2023). Cognitive maturity and self-regulation skills further influence how learners engage with AI feedback and adapt to its suggestions. Recognizing these differences supports the development of age-appropriate AI interventions.

Theoretical Framework and Research Gap

Grounded in adaptive learning theory and learner-centered pedagogy, this study examines how ChatGPT—a widely accessible conversational AI—can enhance reading comprehension among pre-intermediate Iranian EFL learners. English as a Foreign Language (EFL) learners, particularly in Iran, face unique challenges, such as limited exposure to authentic language input, reliance on teacher-centered instruction, and cultural differences that influence text comprehension. ChatGPT addresses these challenges by providing personalized reading materials, interactive prompts, and immediate feedback tailored to learners' proficiency levels, thereby fostering engagement and autonomy in diverse linguistic and cultural contexts. While prior research confirms AI's potential, few studies have systematically investigated its effectiveness in contexts characterized by linguistic, cultural, and demographic diversity. Furthermore, limited attention has been paid to the moderating roles of gender and age in shaping learning outcomes. Building on their findings, the present study extends this line of research by focusing on a more advanced and interactive tool—ChatGPT—and by examining the moderating roles of gender and age.

Purpose of the Study

This study aims to address these gaps by comparing AI-assisted reading comprehension instruction using ChatGPT with traditional methods, while exploring how gender and age moderate these effects. By situating the investigation in the Iranian EFL context, the research seeks to contribute empirically grounded insights into the opportunities and challenges of integrating AI tools into language learning curricula. Findings are expected to inform educators and policymakers on designing inclusive, effective AI-enhanced learning interventions tailored to diverse learner characteristics.

Methodology

Design of the Study

This study adopted a quasi-experimental pretest–posttest control group design to examine the effects of ChatGPT-assisted reading instruction on EFL learners' reading comprehension. Two intact groups of pre-intermediate Iranian EFL learners were assigned to an experimental group receiving AI-enhanced reading instruction and a control group receiving traditional textbook-based instruction. Both groups completed a standardized reading comprehension pre-test and post-test over a six-week instructional period. The design allowed for comparison of within-group progress over time and between-group differences, while also examining the potential moderating effects of gender and age on learning outcomes.

Participants

The study involved 62 pre-intermediate Iranian EFL learners aged 13–16, recruited from a language institute in Ardabil, Iran. Participants' proficiency levels were assessed during registration using an institute-specific placement test aligned with CEFR B1 standards. Four classes (each with approximately 15–16 students) were randomly assigned to either the experimental (ChatGPT-assisted, $n = 31$) or control (traditional, $n = 31$) group by drawing class names from a container to ensure an unbiased allocation. Pre-test scores on the TOEFL Junior Standard Test confirmed group equivalence (ChatGPT-assisted: $M = 23.13$, $SD = 3.55$; traditional: $M = 23.58$, $SD = 3.69$). While the groups were equal in size, the gender distribution was uneven, with 38 female and 24 male participants—an imbalance acknowledged as a study limitation due to the natural variation in class composition resulting from the cluster randomization approach. This imbalance did not significantly affect the study outcomes, as gender was found to have no moderating effect ($p = 0.488$). The control group received instruction through traditional reading comprehension methods, including reading printed texts, answering comprehension questions, and completing vocabulary exercises without digital support. In contrast, the experimental group engaged with AI-assisted reading activities using ChatGPT, enabling a comparative analysis of the effectiveness of traditional and AI-enhanced instructional approaches on reading comprehension outcomes.

Instruments

Reading Comprehension Assessment: Participants' reading comprehension abilities were measured using the reading section of the TOEFL Junior Standard Test (2010), an internationally recognized and widely validated assessment for English language learners. The test was administered as both a pre-test and a post-test to track learners' progress over the study period. The pre-test established baseline comprehension levels, while the post-test assessed improvements after the intervention. The reliability of the reading section of the TOEFL Junior Standard Test (ETS, 2010) was reported as a Cronbach's alpha of 0.85, indicating high internal consistency (So et al., 2015). The test comprised 11 short passages (100–150 words each) on varied topics (e.g., science, history, daily life) with 36 multiple-choice questions (3–4 per passage), completed in 60 minutes. Scoring was based on CEFR B1 criteria, with a maximum score of 36. The reported reliability ($\alpha = 0.85$) was confirmed by statistical analysis of pre-test scores in the present study. This high reliability, coupled with the test's standardized structure, ensured valid comparisons of pre- and post-test results between the ChatGPT-assisted and traditional groups.

Each test consisted of 11 short reading passages covering various topics, with three to four multiple-choice questions per passage, totaling 36 questions per test. Participants had approximately 60 minutes to complete each test. The use of the TOEFL Junior ensured a standardized, reliable, and consistent measure of reading comprehension across both groups, allowing for valid comparisons between pre- and post-test results. On the pre-test, participants in the ChatGPT-assisted group scored

between 13 and 30 ($M = 23.13$, $SD = 3.55$), while those in the traditional group scored between 17 and 30 ($M = 23.58$, $SD = 3.69$), out of a maximum score of 36. These scores align with the pre-intermediate proficiency level (CEFR B1), confirming the participants' suitability for the study based on their baseline reading comprehension abilities.

AI-Driven Language Learning Tool: The experimental group engaged with *ChatGPT-4*, an AI-powered conversational tool, to support reading comprehension development. Participants accessed ChatGPT through its official web platform on personal devices such as smartphones, tablets, and laptops; no installation was required, ensuring uniform access. Over the course of 15 instructional sessions spanning two months, learners used ChatGPT to enhance their understanding of texts through several strategies: looking up unfamiliar vocabulary, generating summaries of reading passages, and answering comprehension questions.

ChatGPT provided immediate feedback on learners' responses, helping them identify and correct mistakes. It also offered follow-up questions and prompts that encouraged deeper engagement with the material and critical reflection. This adaptive, learner-centered approach aimed to create a more interactive and personalized reading experience tailored to individual learners' needs. Examples of the reading texts used in the intervention are provided in Appendix A.

Traditional Reading Comprehension Materials: The control group received instruction through conventional reading comprehension methods. Materials included texts selected from age- and proficiency-appropriate books and reputable online resources. Each session featured two to three reading passages accompanied by comprehension questions focusing on main ideas, key details, and vocabulary development.

Teacher feedback was provided through oral explanations and class discussions, which aimed to reinforce understanding and clarify misconceptions. This traditional approach relied on established instructional strategies, including teacher guidance and group interaction, to support learners' comprehension and vocabulary retention.

Data Collection Procedure

The study was conducted over a two-month period, during which both the experimental and control groups participated in 15 reading comprehension sessions, each lasting approximately 60 minutes and held twice per week. No pilot study was conducted prior to the main intervention, as the brief training session on using ChatGPT-4 was deemed sufficient to familiarize participants with the tool and ensure the intervention's feasibility.

The control group followed conventional reading comprehension instruction. In each session, learners were provided with reading passages and related tasks, including answering comprehension questions and completing vocabulary exercises. Teacher-led explanations and group discussions were used to reinforce understanding, clarify key ideas, and encourage learner participation. This approach adhered to a traditional teacher-centered instructional framework, emphasizing direct instruction and structured group activities.

Before starting the intervention, the experimental group received a brief training session on effectively using ChatGPT-4 as an AI-driven reading tool. During the study, learners used their personal devices (smartphones, tablets, or laptops) to access ChatGPT through its web platform, ensuring consistent and equitable access without requiring software installation.

In practice, the experimental group engaged in several AI-supported strategies to enhance reading comprehension:

- **Interactive reading:** querying and discussing passages with ChatGPT to deepen understanding;
- **Vocabulary development:** accessing definitions, example sentences, and synonyms suggested by the AI;

- **Comprehension support:** answering AI-generated questions and receiving immediate feedback to correct misunderstandings;
- **Summarization and reflection:** generating summaries and discussing them with peers;
- **Personalization:** receiving recommendations and prompts tailored to learners' proficiency levels and interests;
- **Collaborative learning:** participating in group discussions facilitated by ChatGPT's follow-up questions.

These integrated strategies aimed to create an engaging, adaptive, and learner-centered environment, enabling participants to interact actively with texts and develop critical reading skills. The experimental group followed a learner-centered instructional approach, leveraging ChatGPT's adaptive capabilities to foster personalized and interactive reading experiences.

Data Analysis

To examine the effect of instructional modality (ChatGPT-assisted vs. traditional) on reading comprehension, two complementary statistical analyses were employed. First, paired-sample t-tests were conducted separately for each group (ChatGPT-assisted and traditional) to assess within-group changes between pre-test and post-test scores. This preliminary analysis helped identify potential improvements within each group and provided an initial understanding of the intervention's impact (e.g., the ChatGPT-assisted group improved from a mean of 23.13 to 25.61, $p < 0.001$; the traditional group from 23.58 to 24.13, $p = 0.281$). Subsequently, an analysis of covariance (ANCOVA) was performed to compare post-test scores between the two groups, with pre-test scores as a covariate. ANCOVA enabled a more precise examination of between-group differences by controlling for initial variations in reading comprehension abilities. The combination of these tests, given their distinct purposes (within-group vs. between-group analysis), provides a comprehensive picture of the intervention's effect and is common in similar quasi-experimental studies (Creswell & Creswell, 2018).

Results

Before presenting the main results addressing the research questions (RQs), we first report the preliminary descriptive statistics. These include analyses of the mean, median, standard deviation, range, minimum, and maximum scores for both the pre-test and post-test in the ChatGPT-assisted and traditional groups. Table 1 summarizes these descriptive statistics, offering an overview of participants' reading comprehension performance before and after the intervention across both instructional modalities.

Table 1

Descriptive Statistics for Pre-Test and Post-Test Scores by Group

Group		Mean	Median	Std. Deviation	Range	Minimum	Maximum
Pre-Test Score	Chat-GPT-Assisted	23.13	24	3.55	17	13	30
	Traditional	23.58	24	3.69	13	17	30
Post-Test Score	Chat-GPT-Assisted	25.61	26	3.60	14	16	30
	Traditional	24.13	24	4.45	14	16	30

Pre-Test Scores: The ChatGPT-assisted group had a mean pre-test score of 23.13 (SD = 3.55), whereas the traditional group had a slightly higher mean of 23.58 (SD = 3.69). Both groups demonstrated similar central tendencies, each with a median score of 24.

Post-Test Scores: Following the intervention, the ChatGPT-assisted group achieved a higher mean post-test score of 25.61 (SD = 3.60), compared to the traditional group's mean score of 24.13 (SD = 4.45). While both groups showed improvement, the ChatGPT-assisted group exhibited a greater overall gain in reading comprehension performance.

To address Research Question One, which investigates whether both instructional methods significantly improve reading comprehension within each group over time, paired samples t-tests were conducted separately for the ChatGPT-assisted and traditional groups (see Table 2).

Table 2

Paired T-Test Results for Improvement in Reading Comprehension (Pre-Test to Post-Test)

Group		Mean	N	Std. Deviation	Paired Differences			P-value
					Mean	Std. Deviation	95% CI Lower Upper	
Chat-GPT-Assisted	Pre-Test Score	23.13	31	0.64				
	Post-Test Score	25.61	31	0.65	-2.48	1.80	-3.15	-1.82
Traditional	Pre-Test Score	23.58	31	0.66				
	Post-Test Score	24.13	31	0.80	-0.55	2.78	-1.57	0.47

The paired t-test results indicate that the ChatGPT-assisted group demonstrated a statistically significant improvement in reading comprehension, with mean scores increasing from pre-test (M = 23.13, SD = 0.64) to post-test (M = 25.61, SD = 0.65). The mean difference of -2.48 was significant ($p < 0.001$), suggesting a substantial gain attributable to the AI-assisted intervention.

In contrast, the traditional group showed only a modest, non-significant increase in reading comprehension, from a mean pre-test score of 23.58 (SD = 0.66) to a post-test mean of 24.13 (SD = 0.80), yielding a mean difference of -0.55 ($p = 0.281$). Therefore, Research Question One was partially supported: the ChatGPT-assisted method led to significant within-group improvement, whereas the traditional method did not.

To address Research Question Two, which examines whether there is a significant difference in reading comprehension between the ChatGPT-assisted and traditional groups, an Analysis of Covariance (ANCOVA) was conducted on post-test scores, controlling for pre-test scores as a covariate (see Table 3).

Table 3

Tests of Between-Subjects Effects

Dependent Variable: Post-Test Score

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	693.219 ^a	2	346.610	63.166	.000	.682
Intercept	17.260	1	17.260	3.145	.081	.051
Pre	659.090	1	659.090	120.113	.000	.671
Group	55.564	1	55.564	10.126	.002	.146
Error	323.749	59	5.487			
Total	39368.000	62				
Corrected Total	1016.968	61				

a. R Squared = .682 (Adjusted R Squared = .671)

Table 3 presents the results of the ANCOVA analysis. The following findings were observed:

ChatGPT-Assisted vs. Traditional: The ANCOVA results further support these findings by revealing a statistically significant effect of instructional modality on post-test scores ($F(1, 59) = 10.126, p = .002$). This indicates that the method of instruction had a meaningful impact on learners' reading comprehension outcomes. The model accounted for a substantial proportion of variance in post-test performance ($R^2 = 0.682$), with pre-test scores emerging as a significant covariate ($F(1, 59) = 120.113, p < .001$), highlighting the critical role of initial proficiency levels. Overall, these results suggest that ChatGPT-assisted instruction is more effective in improving reading comprehension compared to traditional methods.

Figure 1

Mean Test Scores by Time (Pre-Test and Post-Test) for AI-Assisted and Traditional Groups with 95% Confidence Intervals

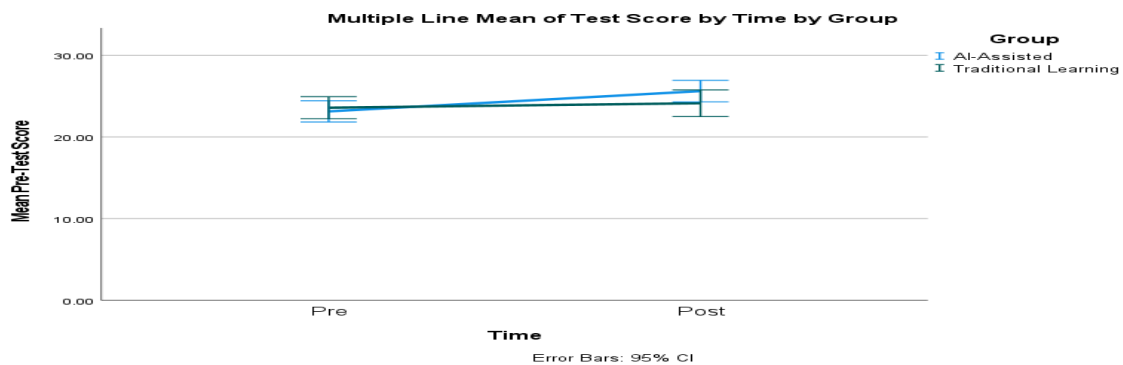


Figure 1 presents the mean pre-test and post-test scores for both the ChatGPT-assisted and traditional groups, accompanied by 95% confidence intervals. The figure illustrates that although both groups demonstrated improvements in reading comprehension over time, the ChatGPT-assisted group exhibited a more pronounced gain. While the overlapping confidence intervals indicate some variability within the data, the overall trend suggests that the AI-assisted approach had a more substantial positive effect on reading comprehension outcomes.

Interaction between Instructional Modality and Gender:

To address Research Question Three, which investigates whether there is a significant interaction between instructional modality (ChatGPT-assisted vs. traditional) and gender (male vs. female) in influencing reading comprehension, a two-way Analysis of Covariance (ANCOVA) was conducted on post-test scores, controlling for pre-test scores as a covariate (see Table 4).

Table 4

ANOVA Results for the Interaction Between Modality and Gender on Reading Comprehension

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	697.104 ^a	4	174.28	31.06	<0.001
Intercept	19.21	1	19.21	3.42	0.069
Pre	610.06	1	610.06	108.71	<0.001
Group	49.30	1	49.30	8.79	0.004
Gender	1.20	1	1.20	0.21	0.645
Group * Gender	2.73	1	2.73	0.49	0.488
Error	319.86	57	5.61		
Total	39368.00	62			
Corrected Total	1016.97	61			

Note: R Squared = .685 (Adjusted R Squared = .663). Dependent Variable: Post-Test Score

The ANOVA results for Research Question 3, which examined whether gender interacts with instructional modality (ChatGPT-assisted vs. traditional) to influence reading comprehension, indicated no significant interaction effect. Specifically, the interaction term between group and gender (Group * Gender) was not significant ($F(1, 57) = 0.49, p = 0.488$), suggesting that gender did not meaningfully moderate the relationship between instructional modality and reading comprehension outcomes (see Table 4).

Moderating Effect of Age on the Relationship between Instructional Modality and Reading Comprehension

To address Research Question Four, which examines whether age moderates the relationship between instructional modality (ChatGPT-assisted vs. traditional) and reading comprehension, a two-way Analysis of Covariance (ANCOVA) was conducted, including the interaction term between group and age (see Table 5).

Table 5

ANOVA Results for the Moderating Effect of Age on the Relationship between Modality and Reading Comprehension

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	695.250 ^a	4	173.81	30.80	0.000
Intercept	2.97	1	2.97	0.53	0.471
Group	649.21	1	649.21	115.02	0.000
Pre	0.17	1	0.17	0.03	0.863
Age	0.11	1	0.11	0.02	0.887
Group * Age	1.23	1	1.23	0.22	0.643
Error	321.72	57	5.64		
Total	39368.00	62			
Corrected Total	1016.97	61			

The results indicate that age did not significantly moderate the effect of instructional modality on reading comprehension, as reflected by a non-significant interaction effect ($F(1, 57) = 0.22, p = 0.643$). Therefore, Research Question Four was not supported, suggesting that the superiority of ChatGPT-assisted instruction over traditional methods was consistent across different age levels in this sample.

Discussion

Regarding the effect of ChatGPT-assisted instruction, the findings of this study highlight the transformative potential of AI-driven tools—particularly ChatGPT—in enhancing the reading comprehension skills of pre-intermediate-level Iranian EFL learners. The significant improvement observed in the experimental group underscores the capacity of AI tools to achieve better educational outcomes compared to traditional methods (Chapelle, 2001; Qiao & Zhao, 2023).

A key factor contributing to the experimental group's enhanced performance is the personalized, interactive, and adaptive nature of AI tools like ChatGPT. Traditional methods often adopt a one-size-fits-all approach, which may not sufficiently address the diverse learning needs of pre-intermediate learners engaging with increasingly complex texts (Riazi & Mosalanejad, 2010). In contrast, AI tools provide real-time feedback, personalized guidance, and adaptive content that aligns

with each learner's proficiency level and pace (Huang et al., 2020; Kukulska-Hulme, 2020). This adaptability allows learners to participate in a more interactive and meaningful learning experience, supported by dialogue, clarification, and scaffolded feedback (Huang et al., 2020; Vandergrift & Goh, 2012). By offering tailored learning pathways, AI tools like ChatGPT help bridge the gaps inherent in traditional instruction and equip learners with the support needed to develop stronger reading comprehension skills (Hamuddin, 2018; Qiao & Zhao, 2023).

The superiority of ChatGPT-assisted instruction in enhancing reading comprehension aligns with recent evidence from this journal demonstrating the efficacy of AI-driven tools in the Iranian EFL context. Specifically, Esmaeily and Mahdavi Zafarghandi (2025)—published in this journal—reported significant improvements in Iranian EFL teachers' reflective practice through the use of ChatGPT and Mote for providing actionable feedback. The present study extends these findings from teacher professional development to direct learner outcomes, showing that conversational AI can similarly transform reading comprehension among pre-intermediate EFL students. Furthermore, Mirsanjari (2025), also published in this journal, highlighted the benefits of dialogic scaffolding in digital learning environments for improving EFL writing proficiency. Taken together, these studies and the current results illustrate that interactive, technology-mediated approaches consistently outperform traditional methods across different language skills in the Iranian EFL context.

These findings align well with broader international research underscoring the effectiveness of AI in education and language learning contexts (Graesser et al., 2005; VanLehn, 2011; Qiao & Zhao, 2023). Moreover, recent research emphasizes the value of integrating multimedia into AI-assisted learning environments. For example, Nurjaya et al. (2024) highlight how features like interactive visuals, audio, and animations can significantly boost student engagement and retention of information. Incorporating these multimedia elements into AI tools offers a multisensory learning experience, making reading comprehension more immersive and effective (Huang et al., 2020; Vandergrift & Goh, 2012). By reinforcing textual content with visual and auditory cues, these tools cater to different learning preferences, creating a more holistic educational approach (Alan, 2023; Kumar & Rose, 2011).

The adaptive capacity of AI tools aligns with contemporary personalized learning theories. Adaptive technologies that respond to learners' evolving needs can enhance educational outcomes by offering timely, individualized support (Kukulska-Hulme, 2020). In this study, the use of ChatGPT for reading comprehension provided such personalized and adaptive support, contributing to the notable gains observed in the experimental group's performance (Huang et al., 2020). By integrating AI-driven, personalized, and multimedia-rich instructional approaches, this study adds further evidence to the expanding literature that highlights the transformative potential of AI in language education (Chapelle, 2001; Hamuddin, 2018).

The interaction between instructional modality (ChatGPT-assisted vs. traditional) and gender in influencing reading comprehension: Findings from this study suggest that AI tools like ChatGPT can help address limitations inherent in traditional educational methods by providing scalable, consistent, and personalized learning opportunities. This is especially significant when considering gender differences in language learning. In educational settings with limited access to highly qualified instructors or advanced resources, AI-powered tools can democratize learning, offering equitable opportunities for all students—regardless of gender—to develop their reading comprehension skills (Chen et al., 2020; Yang et al., 2022).

Interestingly, the study did not find significant gender differences in performance gains among learners using AI tools. This result is notable, given that prior research has often linked gender to differing levels of engagement with technology, with males sometimes reporting higher engagement due to socio-cultural influences (Fathi et al., 2023; Wong, 2019). In contrast, this study shows that both male and female learners benefited similarly from the use of ChatGPT, suggesting that AI tools can effectively support a wide range of learning styles and preferences (Graesser et al., 2005; Kerr, 2019). The adaptive features of AI—such as immediate feedback and individualized support—appear to ensure that students of all genders engage meaningfully with the content (Kumar

& Rose, 2011; Woolf, 2010). This indicates that AI platforms like ChatGPT hold promise as inclusive tools that promote equitable learning outcomes across gender lines.

These findings underscore the potential of AI to foster a more equitable learning experience for both male and female learners. Through features like personalized feedback, adaptive content, and real-time interaction, AI tools empower all students—regardless of gender—to strengthen their reading comprehension skills effectively (Gefen & Straub, 2000; Yang et al., 2022). This reinforces the value of integrating AI into language education, as it creates opportunities for every learner to succeed within a personalized and responsive instructional environment.

In this study, age was examined as a potential moderator but was found to have no statistically significant effect on the effectiveness of instructional modalities (ChatGPT-assisted vs. traditional) ($p = 0.643$). Nevertheless, participants, aged 13–16, represent a developmental stage marked by significant cognitive, emotional, and social changes, which can influence how they interact with and benefit from educational technology. At this stage, learners' cognitive capacities are rapidly evolving, making the adaptability of AI tools particularly valuable for addressing their developmental needs. By offering personalized learning pathways aligned with students' cognitive levels, AI tools can help adolescents gradually construct more sophisticated cognitive frameworks as they mature (Vandergrift & Goh, 2012).

For younger adolescents (13–14 years), greater scaffolding and teacher guidance may be necessary to fully benefit from AI support, whereas older students (15–16 years) might navigate AI-driven systems more independently (Kukulka-Hulme, 2020). Future research could explore how age-specific adaptations of AI tools could further enhance learning outcomes. By aligning AI functionalities with the cognitive and developmental needs of various age groups, educators and developers could maximize learning benefits across the age spectrum (Wong, 2019).

Additionally, examining the intersection of age and gender in AI-assisted learning environments could provide deeper insights into how AI tools can be designed to address learners' diverse needs. Although this study found no significant moderating effects for age or gender ($p = 0.643$ for age; $p = 0.488$ for gender), existing research suggests that gender and developmental differences can shape technology engagement (Kewalramani et al., 2022; Wong, 2019; Woolf, 2010). A nuanced understanding of how cognitive and developmental characteristics vary by age and gender could guide the refinement of AI tools to create more personalized and equitable learning experiences, ultimately supporting reading comprehension among a broader range of students.

This study also makes several theoretical contributions to the fields of educational technology and second language acquisition. First, the findings underscore the significance of integrating adaptive learning technologies into language teaching frameworks. The superior reading comprehension outcomes observed in the experimental group align with personalized learning theories, which emphasize that individualized instruction increases cognitive engagement and retention (Baker & Siemens, 2014; Kukulka-Hulme, 2020).

Furthermore, the results lend support to cognitive load theory, which argues that learning is optimized when instructional materials are matched to learners' cognitive capacities (Paas et al., 2022). By offering real-time feedback and dynamically adjusting task difficulty based on learner responses, AI tools like ChatGPT help reduce extraneous cognitive load, enabling more efficient processing of new information (Vandergrift & Goh, 2012). This demonstrates how AI can help overcome cognitive barriers and promote more effective learning.

Notably, the absence of significant gender differences in the effectiveness of AI tools challenges common assumptions about gender disparities in technology engagement. Previous studies have often reported that male learners engage more readily with technology due to socio-cultural factors (Fathi et al., 2023). However, the present study suggests that AI tools can provide an equitable learning environment in which both male and female students benefit equally from adaptive learning features. This finding contributes to the growing literature on the potential of AI to advance gender equity in education.

Additionally, the moderating effect of age observed in this study supports developmental theories of learning, particularly during adolescence. It underscores that AI tools can accommodate the evolving cognitive abilities of adolescent learners. These findings suggest that further research into age-specific adaptations of AI-based instruction is essential to optimize learning outcomes across different age groups (Hamuddin, 2018). This insight opens new avenues for refining AI tools to better support learners at various developmental stages.

Despite these significant findings, several limitations must be acknowledged. First, the sample size—particularly the imbalance in gender representation, with more female than male participants—may limit the generalizability of the results. This overrepresentation of female learners could affect understanding of how AI tools like ChatGPT impact male students, given that previous studies have suggested gender differences in technology engagement (Cai et al., 2017; Wong, 2019). Future research should aim for a more balanced gender distribution to better assess how AI tools influence both male and female learners.

Another limitation lies in the focus on a narrow age range (13–16 years), which may not capture the full diversity of cognitive and developmental stages present in a broader student population. Including a wider age range in future studies could provide deeper insights into how AI tools perform across different developmental levels, especially among younger and older learners.

Additionally, the study was conducted within a single educational context, which may limit the applicability of the findings to other settings or cultural contexts. Future research could expand to include diverse educational environments, thereby increasing the robustness and cross-cultural generalizability of results. Finally, while this study demonstrated the short-term effectiveness of AI tools in improving reading comprehension, it did not examine the long-term sustainability of these improvements. Future studies should explore whether gains in reading comprehension are maintained over time and whether these gains extend to other academic skills and subjects.

Ultimately, further research is needed to explore the integration of AI tools with diverse teaching methodologies, such as collaborative learning, flipped classrooms, and project-based learning. Investigating how AI can be effectively combined with these pedagogical approaches may lead to even greater improvements in learning outcomes. Understanding the potential synergies between AI technologies and established teaching methods will be essential for optimizing their impact in educational contexts.

The implications of this research for educational practice are substantial. As AI technologies continue to advance, their integration into language curricula offers the potential for highly engaging, personalized, and adaptive learning experiences—particularly for learners at transitional stages of language development. The positive results observed in this study suggest that educators should consider incorporating AI-driven tools as a core component of instructional design, especially in settings where traditional methods may fall short in meeting diverse learner needs (Heffernan et al., 2014). Moreover, the scalability and adaptability of AI technologies present promising solutions for addressing educational inequalities, enabling students from varied backgrounds to access high-quality, individualized instruction (Chen et al., 2020; Yang et al., 2022).

Conclusion and Implications

This study found that ChatGPT-assisted instruction significantly improved reading comprehension among pre-intermediate Iranian EFL learners, compared to the traditional group. This study demonstrates the significant impact of AI-driven tools, such as ChatGPT, on improving the reading comprehension skills of pre-intermediate-level Iranian EFL learners, both male and female. The findings highlight the potential of AI to deliver personalized and adaptive learning experiences that go beyond traditional methods, offering a more engaging and effective approach to language education. As AI technologies continue to develop, it becomes increasingly important for educators and policymakers to explore their integration into educational practice to better meet learners' diverse needs and foster more equitable learning environments.

However, this study was limited to pre-intermediate Iranian EFL learners in a specific context and time frame, with a focus on reading comprehension and access to AI technology. These delimitations and contextual constraints mean that the findings should be generalized to other proficiency levels, skills, or educational settings with caution.

Future research should extend these findings to include a broader range of learner populations and educational contexts. Longitudinal studies, in particular, could provide valuable insights into the long-term effects of AI tools on language acquisition and their capacity to complement or enhance traditional instructional approaches. Specifically, future research could explore why gender and age did not moderate outcomes in this study by examining these variables in diverse EFL contexts or investigate the long-term efficacy of ChatGPT in resource-constrained settings like Iran. This study adds to the growing body of literature advocating for the thoughtful integration of AI into education, offering evidence that such technologies can significantly improve learning outcomes. By leveraging the adaptive capabilities of AI, educators can help create more effective, inclusive, and engaging learning experiences for students around the world.

Acknowledgement

The authors would like to express sincere gratitude to all the participants who took part in this study. Special thanks are extended to the staff and colleagues who provided valuable support and guidance throughout the research process. Their assistance and encouragement were essential in completing this study successfully.

Declaration of Conflicting Interests

The authors declare that there is no conflict of interest regarding the publication of this manuscript.

Funding Details

This research received no specific grant from any funding agency in the public, commercial, or not for-profit sectors.

References

- Alan, S. (2023). Digital storytelling and multimedia tools for teaching language and literature. *Shanlax International Journal of English*, 12(S1), 110–114. <https://doi.org/10.34293/rtdh.v12iS1-Dec.49>
- Baker, R. S. (2016). Baker, R. S. (2016). Stupid tutoring systems, intelligent humans. *International Journal of Artificial Intelligence in Education*, 26(2), 600–614. <https://doi.org/10.1007/s40593-016-0105-0>
- Baker, R., & Siemens, G. (2014). Educational data mining and learning analytics. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (2nd ed., pp. 253–272). Cambridge University Press. <https://doi.org/10.1017/CBO9780511816833.015>
- Cai, Z., Fan, X., & Du, J. (2017). Gender and attitudes toward technology use: A meta-analysis. *Computers & Education*, 105, 1-13. <https://doi.org/10.1016/j.compedu.2016.11.003>
- Chapelle, C. A. (2001). *Computer Applications in Second Language Acquisition: Foundations for Teaching, Testing, and Research*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139524681>
- Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access*, 8, 75264-75278. <https://doi.org/10.1109/ACCESS.2020.2988510>

- Chen, X., Xie, H., & Hwang, G.-J. (2020). A multi-perspective study on artificial intelligence in education: Grants, conferences, journals, software tools, institutions, and countries. *Computers & Education: Artificial Intelligence*, 1, Article 100014. <https://doi.org/10.1016/j.caeai.2020.100014>
- Creswell, J. W., & Creswell, J. D. (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (5th ed.). SAGE Publications.
- Esmaily, J., & Mahdavi Zafarghandi, A. (2025). The Impact of AI-Driven Feedback on Iranian EFL Teachers' Reflective Practice. *Iranian Journal of English for Academic Purposes*, 14(2), 79-94.
- ETS (2012). *TOEFL Junior Comprehensive Test Technical Report*. Educational Testing Service. <https://doi.org/10.1002/j.2333-8504.2012.tb02294.x>
- Fathi, J., Ahmadi, M., & Yeganeh, T. (2023). Gender differences in willingness to communicate and use of language learning strategies in an AI-assisted EFL learning context. *Journal of Language and Education*, 9(2), 92–106. <https://doi.org/10.17323/jle.2023.15596>
- Gefen, D., & Straub, D. W. (2000). Gender Differences in the Perception and Use of E-Learning Technologies. *Information Systems Research*, 11(4), 363–379. <https://doi.org/10.1287/isre.11.4.363.11876>
- Graesser, A. C., McNamara, D. S., & Cai, Z. (2005). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods*, 37(3), 340–356. <https://doi.org/10.3758/BF03192704>
- Hamuddin, S. (2018). The efficacy of AI-driven educational tools in enhancing student learning. *Educational Technology Research and Development*, 66(2), 379-394. <https://doi.org/10.1007/s11423-018-9573-4>
- Heffernan, N. T., & Heffernan, C. L. (2014). The ASSISTments ecosystem: Building a platform that brings scientists and teachers together for minimally invasive research on human learning and teaching. *International Journal of Artificial Intelligence in Education*, 24(4), 470-497. <https://doi.org/10.1007/s40593-014-0024-x>
- Huang, S., Zhou, X., & Zhang, L. (2020). AI in education: Current trends and future directions. *International Journal of Artificial Intelligence in Education*, 30(2), 189-203. <https://doi.org/10.1007/s40593-020-00178-6>
- Hwang, G.-J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of Artificial Intelligence in Education. *Computers and Education: Artificial Intelligence*, 1, 100001. <https://doi.org/10.1016/j.caeai.2020.100001>
- Hwang, G.-J., Wang, S.-C., & Lai, C.-L. (2020). Effects of digital learning on student motivation and achievement: A meta-analysis. *Computers & Education*, 147, 103819. <https://doi.org/10.1016/j.compedu.2020.103819>
- Johnson, K., & Brown, J. (2023). Age-related factors in educational technology use: A review. *Computers in Human Behavior*, 122, 106810. <https://doi.org/10.1016/j.chb.2021.106810>
- Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... & Zierer, K. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Computers and Education: Artificial Intelligence*, 4, 100127. <https://doi.org/10.1016/j.caeai.2023.100127>
- Kewalramani, S., et al. (2022). AI-powered tools for early childhood education: Addressing developmental diversity. *Journal of Educational Technology & Society*, 25(1), 89-104. [https://doi.org/10.30191/ETS.202201_25\(1\).0006](https://doi.org/10.30191/ETS.202201_25(1).0006)

- Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2022). *AI for school teachers*. CRC Press. <https://doi.org/10.1201/9781003162629>
- Luckin, R., George, K., & Cukurova, M. (2022). *AI for school teachers*. Routledge. <https://doi.org/10.1201/9781003193173>
- Luckin, R. (2018). *Machine Learning and Human Intelligence: The Future of Education*. UCL Press. <https://doi.org/10.14324/111.9781787352810>
- Kukulka-Hulme, A. (2020). Intelligent Assistants in Education: Capabilities and Limitations. *Journal of Learning Analytics*, 7(3), 103–116. <https://doi.org/10.18608/jla.2020.73.8>
- Mirsanjari, Z. (2025). Fostering EFL writing proficiency: The impact of dialogic scaffolding in digital learning environments. *Iranian Journal of English for Academic Purposes*, 14(2), 62–78. <http://ijeap.iust.ac.ir/article-1-772-en.html>
- Kumar, R., & Rose, C. (2011). The role of AI in improving reading comprehension: A review. *Computers & Education*, 56(3), 756-763. <https://doi.org/10.1016/j.compedu.2010.10.010>
- Nguyen, A., Ngo, H.N., Hong, Y., Dang, B. and Nguyen, B.P.T. (2023) Ethical Principles for Artificial Intelligence in Education. *Education and Information Technologies*, 28, 4221-4241. <https://doi.org/10.1007/s10639-022-11316-w>
- Nurjaya, Y., Yono, M., Maulana, I., Shofia, M., Maulida, B., Bakri, A., Antonia, J., & Junianty, L. (2024). The impact of multimedia elements on tablets and digital stories in learning process management. *Journal of Education Technology*, 8(1), 45-63. <https://doi.org/10.23887/jet.v8i1.71326>
- Paas, F., van Merriënboer, J. J. G., & Sweller, J. (2022). Cognitive-load theory: Methods to manage working memory load in the learning of complex tasks. *Educational Psychology Review*, 34(1), 1-23. <https://doi.org/10.1007/s10648-021-09523-0>
- Qiao, M., & Zhao, Y. (2023). Personalized learning with AI: Effects on student achievement and engagement. *Journal of Learning Analytics*, 10(1), 60-73. <https://doi.org/10.18608/jla.2023.1001>
- Riazi, A. M., & Mosalanejad, N. (2010). Evaluation of Language Learning and Teaching in Iran: A Mixed-Methods Approach. *TESOL Quarterly*, 44(2), 355–372. <https://doi.org/10.5054/tq.2010.223803>
- Shute, V. J., & Rahimi, S. (2017). Review of computer-based assessment for learning in elementary and secondary education. *Journal of Computer Assisted Learning*, 33 (1), 1-19. <https://doi.org/10.1111/jcal.12172>
- Siemens, G. (2013). Learning analytics: Theoretical perspectives. *Journal of Learning Analytics*, 1(1), 1-17. <https://doi.org/10.18608/jla.2013.1101>
- So, Y., Wolf, M. K., Hauck, M. C., Mollaun, P., Rybinski, P., Tumposky, D., & Wang, L. (2015). TOEFL Junior® Design Framework (Research Report No. RR-15-13). *ETS Research Report Series*, 2015(1), 1–42. <https://doi.org/10.1002/ets2.12058>
- Vandergrift, L., & Goh, C. C. M. (2012). Teaching and Learning Second Language Listening: Metacognition in Action. *Routledge*. <https://doi.org/10.4324/9780203848371>
- VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, 46(4), 197–221. <https://doi.org/10.1080/00461520.2011.611369>
- Venkatesh, V., & Morris, M. G. (2000). Why don't men ever stop to ask for directions? Gender, social influence, and their role in technology acceptance. *MIS Quarterly*, 24(1), 115-139. DOI: 10.2307/3250981

- Wong, G. K. (2019). Gender differences in digital literacy: A meta-analysis. *Computers & Education*, 142, 103636. <https://doi.org/10.1016/j.compedu.2019.103636>
- Woolf, B. P. (2010). Building intelligent interactive tutors: *Student-centered strategies for revolutionizing e-learning*. Morgan Kaufmann/Elsevier. <https://doi.org/10.1016/B978-0-12-373594-2.00001-7>
- Yang, S., Ogata, H., Matsui, T., & Chen, N. S. (2022). Human-centered AI in education: From theory to practice. *Computers & Education*, 189, 104582. <https://doi.org/10.1016/j.compedu.2022.104582>
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – Where are the educators? *International Journal of Educational Technology in Higher Education*, 16, Article 39. <https://doi.org/10.1186/s41239-019-0171-0>

Appendix A: Sample Reading Passage from TOEFL Junior Standard Test (2010)

The following is an official sample from the **Reading Comprehension section** of the TOEFL Junior Standard Test (2010), provided by ETS (Educational Testing Service). This section measures the ability to read and understand academic and non-academic texts. The sample includes one short passage (approximately 150-200 words) followed by 4 multiple-choice questions. In the full test, there are 11 passages with 3-4 questions each, totaling 36 questions, to be completed in about 60 minutes.

Directions: Read the passage below. Then read the questions that follow it and the four possible answers. Choose the best answer for each question.

Sample Passage: The History of Chocolate

Chocolate is one of the world's favorite foods, but few people know its long history. The story of chocolate begins in South America more than 2,000 years ago. The ancient Mayans and Aztecs made a drink from the seeds of the cacao tree. They mixed the seeds with water, spices, and sometimes chili peppers. This bitter drink was used in religious ceremonies and as money.

When Spanish explorers arrived in the 1500s, they brought cacao back to Europe. At first, Europeans found the drink strange, but they soon added sugar and milk to make it sweeter. By the 1800s, chocolate became a popular treat in solid form, thanks to inventions like the chocolate press. Today, chocolate is enjoyed all over the world in bars, cakes, and candies. However, scientists warn that eating too much chocolate can be unhealthy because of its sugar and fat content.

Questions

1. **What is the main topic of the passage?**
 - A. How chocolate is made today
 - B. The history and popularity of chocolate
 - C. The health benefits of cacao seeds
 - D. Chocolate drinks in ancient ceremonies

Best Answer: B

2. **According to the passage, what did the ancient Mayans and Aztecs use chocolate for?**
- A. As a sweet dessert
 - B. In religious ceremonies and as money
 - C. To make solid bars
 - D. To add sugar to European drinks

Best Answer: B

3. **Why did Europeans change the original chocolate drink?**
- A. To make it less bitter
 - B. To use it as medicine
 - C. To grow cacao trees in Europe
 - D. To sell it in South America

Best Answer: A

4. **What does the passage suggest about eating chocolate?**
- A. It is always healthy
 - B. It should be eaten in large amounts
 - C. Too much can be unhealthy
 - D. It has no fat or sugar

Best Answer: C

Note: This sample is adapted from official ETS materials (TOEFL Junior Standard Test, 2010). For the full test structure, refer to the ETS website. Participants in this study completed the entire Reading Comprehension section (36 questions) as pre- and post-tests.

Source: ETS Official TOEFL Junior Preparation Materials (available at: <https://www.ets.org/toefl/junior/prepare/reading-comprehension.html>).