E-ISSN: 2981-1686

Journal of Studies in Language Learning and Teaching

www.jsllt.yazd.ac.ir

JSLLT, Vol. 3, No. 2, 19–35.

Paper Type: Original Article

Generative AI-Mediated Task Scaffolding in EFL Reading Comprehension: Cognitive, Self-Regulatory, and Interactional Mechanisms

Ali Hemmati [1,*]

[1] Department of English Language Teaching, Farhangian University, Tehran, Iran.

Received: 02 May, 2026 Revised: 08 July, 2026 Accepted: 08 September, 2026

Abstract

This study examines the integration of generative AI into Task-Based Language Teaching (TBLT) to enhance L2 reading development. Grounded in Interactionist, Sociocultural, and Self-Regulated Learning theories, it proposes a model of AI-mediated task scaffolding, addressing the limited integration of complementary AI scaffolding functions in previous research. In this framework, AI functions as a structured mediator, coordinating interactional feedback, sociocultural support, and metacognitive regulation across pre-, while-, and post-reading phases. Fifty-eight non-English-major student-teachers (A2–B1 CEFR) participated in a quasi-experimental pretest–posttest control-group design. Both groups completed identical texts and tasks, but the experimental group additionally used theoretically aligned ChatGPT, QuillBot, and Rewordify to support lexical access, inferencing, and reflective processing. The reading outcomes were assessed with a validated 40-item comprehension test, complemented by semi-structured interviews on learners' experiences with AI. The quantitative results showed significantly greater gains for the AI-supported group, particularly in overall reading comprehension and strategic engagement with texts. The qualitative findings indicated that the learners reported reduced perceived effort, sustained engagement, and enhanced metacognitive monitoring while using AI-supported reading tasks. By positioning AI as a pedagogically embedded mediator rather than a stand-alone tool, the study provides an integrated account of how interactional modification, scaffolded development, and self-regulated engagement converge in task-based reading. The findings provide empirical support for the principled integration of AI and highlight practical implications for EFL teacher education.

Keywords: AI-mediated scaffolding, Inferential comprehension, L2 reading development, Task-Based Language Teaching.

I | INTRODUCTION

The rapid expansion of artificial intelligence (AI) technologies is reshaping contemporary language education by transforming how instructional support is delivered and experienced. AI systems can provide forms of adaptive assistance that are difficult to sustain consistently in traditional classroom environments, including immediate feedback, iterative clarification, and personalized guidance (Merino-Campos, 2025; Wang & Fan, 2025; Zawacki-Richter et al., 2019). These capabilities have generated considerable interest among researchers and educators seeking ways to enhance learning efficiency and broaden access to individualized support. At the same time, integrating AI into language pedagogy raises important questions about how technological assistance can be aligned with established principles of language learning and instructional design (Huang & Li, 2024; Y. Li et

Corresponding Author: a.hemmati@cfu.ac.ir

Cite this article: Hemmati, A., (2026), Generative AI-Mediated Task Scaffolding in EFL Reading Comprehension: Cognitive, Self-Regulatory, and Interactional Mechanisms, *Journal of Studies in Language Learning and Teaching*, *3*(2), 19-35. doi: 10.22034/JSLLT.2026.24507.1127.

al., 2025). Drawing on sociocultural views of mediated learning (Vygotsky, 1978; Wood et al., 1976), understanding how AI functions in pedagogically meaningful learning environments, therefore, remains an important issue for contemporary language education research.

These questions are particularly relevant in the domain of second language (L2) reading. Reading comprehension involves the coordinated operation of multiple cognitive processes, including lexical access, syntactic parsing, inferential reasoning, and the construction of coherent mental representations of textual meaning (Cain & Oakhill, 2023; Grabe & Stoller, 2020; Liao & Lee, 2024; Ostojić, 2023). Effective comprehension, therefore, requires learners not only to recognize words but also to integrate information across sentences and generate inferences that connect textual ideas. Although digital tools can facilitate access to textual meaning, forms of technological assistance that provide direct translations or simplified outputs may sometimes lead learners toward surface-level understanding while reducing engagement with deeper interpretive processes (Murtisari, 2024; Park, 2024). These concerns highlight the importance of examining how technological tools can support, rather than replace, the cognitive effort involved in academic reading.

One pedagogical approach that offers a useful perspective for examining this issue is Task-Based Language Teaching (TBLT). Grounded in communicative and interactionist perspectives, TBLT conceptualizes language learning as emerging through engagement with meaning-oriented tasks that simulate authentic communicative activity (Ellis, 2009; Long, 2015). Its emphasis on interaction, negotiation of meaning, and scaffolded task completion makes TBLT an appropriate framework for examining AI as a mediational resource (Ellis, 2009; Long, 2015; Y. Li et al., 2025). In such contexts, learners interact with texts, tasks, and mediational resources while constructing meaning. However, classroom realities such as large class sizes, limited instructional time, and delayed feedback often constrain the degree of individualized support that teachers can provide during demanding reading activities.

Recent discussions suggest that AI tools may function as mediational resources that support learners during task engagement by providing prompts, explanations, and opportunities for reflection (Huang & Li, 2024; Y. Li et al., 2025). This role is consistent with interactionist accounts of negotiated meaning (Swain, 1995) and classroom scaffolding perspectives (Gibbons, 2015), encouraging learners to negotiate meaning with texts, monitor their understanding, and regulate cognitive effort during reading.

However, these forms of support are not theoretically equivalent. Depending on how learners engage with AI, the technology may facilitate different learning mechanisms. For example, AI can encourage negotiation of meaning by prompting learners to clarify interpretations and reconsider misunderstandings, provide contingent instructional assistance that helps learners accomplish tasks beyond their independent capability, and stimulate metacognitive monitoring by encouraging learners to evaluate their comprehension and regulate reading strategies (Huang & Li, 2024; Merino-Campos, 2025; Y. Li et al., 2025). Distinguishing these mechanisms is essential for understanding how AI contributes to learning from complementary theoretical perspectives rather than functioning as a single, undifferentiated source of support. Exploring these possibilities requires closer attention to how AI-mediated assistance interacts with established theories of language learning and with pedagogical approaches such as task-based instruction. The following sections, therefore, outline the theoretical perspectives guiding the study and review relevant research on AI-supported language learning and reading comprehension.

II. LITERATURE REVIEW

2.1. Theoretical framework

The conceptual foundation of this study integrates three complementary perspectives, Interactionist Theory, Sociocultural Theory, and Self-Regulated Learning (SRL) Theory, to explain how AI-supported task-based reading may enhance comprehension. Rather than treating technology as an external add-on, AI is positioned as a mediational layer through which interaction, scaffolded development, and metacognitive regulation converge during engagement with complex academic texts. This triadic framework conceptualizes AI not as an autonomous problem-solver, but as a structured intermediary that conditions how learners negotiate meaning, access contingent assistance, and regulate cognitive efforts in task performance.

Although these perspectives collectively explain AI-supported learning, each emphasizes a distinct instructional mechanism. The Interactionist Theory explains how AI facilitates meaning negotiation through iterative dialogue (Long, 2015; Y. Li et al., 2025), the Sociocultural Theory focuses on AI as a source of contingent instructional mediation in learners' developing capabilities (Huang & Li, 2024; Poehner & Infante, 2019), and the SRL Theory explains how AI supports learners' planning, monitoring, and evaluation of their own comprehension processes (Merino-Campos, 2025; Panadero, 2017). Table 1 summarizes how these theoretical mechanisms are operationalized in the present study.

Table 1. Alignment of AI functions with theoretical mechanisms

Theoretical perspective	AI function in this study	Primary learning mechanism
Interactionist Theory	AI asks clarification questions, requests elaboration, provides reformulations, and offers confirmation feedback.	Negotiation of meaning through iterative interaction that helps learners' identify and resolve comprehension difficulties (Long, 2015)
Sociocultural Theory	AI provides graduated hints, contextual explanations, vocabulary support, and inferential prompts based on learners' responses.	Contingent instructional mediation that temporarily extends learners' performance in the Zone of Proximal Development (Poehner & Infante, 2019)
Self-Regulated Learning Theory	AI encourages goal setting, comprehension monitoring, strategy adjustment, and reflective evaluation before, during, and after reading.	Metacognitive regulation that strengthens learners' planning, monitoring, and self-evaluation throughout task completion (Panadero, 2017)

From an interactionist standpoint, language development emerges through communicative exchange, particularly when meaning-related breakdowns prompt negotiation, feedback, and modified input (Long, 2015). Extending this view to reading, learners negotiate meaning not only with texts but also with mediational tools. Contemporary syntheses examining generative AI in language pedagogy indicate that AI can operationalize interactionist principles by structuring iterative feedback cycles that prompt clarification, confirmation, and reformulation during task performance (Y. Li et al., 2025). These interactional moves parallel negotiation-of-meaning processes by creating cycles of modified input and output that extend beyond human interlocutor exchange.

In the present study, negotiation of meaning refers to learners' iterative engagement with AI-generated clarification requests, explanatory feedback, and reformulated responses that encourage them to revisit ambiguities and refine interpretations rather than simply receive correct answers, consistent with interactionist accounts of modified interaction and feedback (Long, 2015; Y. Li et al., 2025). When embedded in reading tasks, these interactions transform comprehension from linear decoding into dialogically mediated meaning construction (Etkin et al., 2025).

The Sociocultural Theory conceptualizes learning as a socially mediated process occurring in the Zone of Proximal Development (ZPD), where cognitive growth is supported through the mediational artifacts that extend learners' capabilities while maintaining active engagement (Poehner & Infante, 2019). From this perspective, AI serves as a mediational artifact when its support is adaptive, contingent, and developmentally appropriate, promoting learning rather than merely delivering instructional assistance (Huang & Li, 2024). Effective mediation sustains productive cognitive effort instead of simplifying tasks .

Accordingly, AI-enhanced reading tasks use graduated prompts and adaptive explanations to strengthen inferential reasoning and conceptual integration while preserving learner agency. In the present study, contingent support refers to AI responses that adapt to learners' immediate comprehension needs by providing hints, explanations, or guiding questions rather than complete solutions, consistent with the sociocultural view that instructional assistance should foster independent problem solving through active cognitive engagement (Huang & Li, 2024; Poehner & Infante, 2019).

The SRL Theory emphasizes learners' ability to regulate cognition, motivation, and strategy use throughout learning (Panadero, 2017). Because effective reading requires goal setting, comprehension monitoring, and reflective evaluation, AI can support these regulatory processes through comprehension checks, strategic feedback, and post-reading reflection (Merino-Campos, 2025).

In this study, AI functions were aligned with the pre-, while-, and post-reading stages to support specific cognitive and metacognitive processes. The pre-reading prompts activated prior knowledge, the while-reading scaffolds supported monitoring and inferencing, and the post-reading feedback promoted evaluative reflection (Cain & Oakhill, 2023; Grabe & Stoller, 2020). Unlike the Interactionist Theory, which emphasizes meaning negotiation, and the Sociocultural Theory, which emphasizes contingent mediation, the SRL Theory explains how AI supports learners in monitoring comprehension, evaluating strategy use, and regulating reading processes (Merino-Campos, 2025; Panadero, 2017).

Taken together, the Interactionist, Sociocultural, and SRL theories provide a complementary framework for explaining AI-supported task-based reading. In this framework, AI functions as a pedagogically embedded mediator that supports learners in negotiating meaning, receiving contingent assistance, and regulating comprehension during academically demanding reading tasks. This integration provides the theoretical rationale for examining AI as a tool that enhances cognitive, social, and metacognitive engagement rather than simply improving task performance.

In this study, however, scaffolding refers specifically to the contingent instructional support provided during task completion rather than the full scaffolding model proposed by Wood et al. (1976). Accordingly, the study did not examine the gradual fading of support or the possibility of learner dependence on AI. The findings should, therefore, be interpreted as evidence of AI-mediated contingent assistance during reading tasks rather than a comprehensive evaluation of the complete scaffolding process (Huang & Li, 2024).

2.2. Review of previous studies

Building on the theoretical perspectives outlined above, recent research has increasingly examined the role of artificial intelligence in second language education. Although the field has expanded rapidly, much of the existing work remains technology-driven rather than pedagogy-centered (Kartal & Yeşilyurt, 2024; Zawacki-Richter et al., 2019). Early syntheses highlighted the potential of AI tools but often did not examine skill-specific mechanisms or cognitive processes (Zawacki-Richter et al., 2019). Empirical investigations often target discrete skills such as writing, with research showing improvements in accuracy, organization, and feedback quality through automated tools (Gürbüz, 2024; W. Li et al., 2025; Wei et al., 2023). Bibliometric analyses further indicate that personalization, automated feedback, and learner analytics dominate the field; however, many descriptive studies fail to illuminate how AI directly supports specific learning outcomes, particularly in reading comprehension (Kartal & Yeşilyurt, 2024).

Research has also explored learners' affective and dispositional factors. For instance, AI literacy has been shown to enhance motivation and engagement, though its direct influence on reading performance remains underexplored (Yaşar & Karagücük, 2024). Investigations of generative AI tools (W. Li et al., 2025; Y. Li et al., 2025) suggest improvements in feedback quality and learner engagement, primarily in writing contexts, leaving implications for reading largely untested. Recent studies emphasize that pedagogical integration is crucial: AI-enhanced TBLT can provide adaptive scaffolding and foster deeper interaction with texts (Huang et al., 2024), while research on machine translation highlights that cognitive scaffolds must support effortful processing rather than replace it (Murtisari, 2024; Park, 2024). Meta-analytic and conceptual studies further indicate that AI is most effective when feedback is tailored to learners' proficiency, yet few investigations consider task-based reading or self-regulatory processes (Merino-Campos, 2025; Wang & Fan, 2025).

Contextual research in Iran underscores the need for reading-focused AI support. Iranian EFL learners often encounter challenges in reading comprehension, suggesting the need for instructional scaffolds that support meaning construction beyond literal decoding (Yousefi & Askari, 2024). Although studies show that AI-assisted reading tools can improve comprehension relative to traditional instruction (Huang & Li, 2024; Yousefi & Askari, 2024), the integration of AI in structured, task-based pedagogy remains underexplored. Broader analyses also note that digital literacy, access to technology, and ethical awareness can shape the effectiveness of AI, though these factors are rarely linked directly to measurable outcomes (Semerikov et al., 2024).

At the same time, recent scholarship cautions that AI-supported language learning may be constrained by limitations including inaccurate or biased outputs, learner overreliance, and ethical concerns related to transparency and data privacy (Marino et al., 2024; Selvam & González Vallejo, 2025; Semerikov et al., 2024). Accordingly, AI is increasingly viewed as a pedagogical support that requires critical teacher guidance rather than an autonomous instructional substitute.

Overall, the literature reveals several persistent limitations. First, relatively little research has examined AI-supported interventions specifically targeting reading comprehension. Second, many studies consider AI tools in isolation rather than integrated in theoretically grounded pedagogical frameworks such as task-based language teaching. Third, the research involving student-teachers, particularly those from non-English academic disciplines, remains limited. Addressing these issues requires investigations that connect technological affordances with established theories of language learning while examining how AI-mediated support may influence both cognitive and metacognitive engagement during reading. Taken together, these limitations highlight the need for empirically grounded research that situates AI-supported reading in theoretically informed task-based pedagogy and examines its implications for learners' comprehension and learning experiences.

III. PURPOSE OF THE STUDY

The present study examines the pedagogical role of generative artificial intelligence when integrated into task-based reading instruction in an English as a Foreign Language (EFL) context. Specifically, it investigates how AI-supported scaffolding can be incorporated into task-based reading activities to support learners' engagement with texts and enhance their reading comprehension. By situating AI tools in the structured phases of task-based instruction, the study considers how such technologies may function as mediational resources that assist learners in addressing lexical challenges, constructing inferences, and reflecting on textual meaning. The study is guided by the following research questions:

1. To what extent does AI-mediated task-based reading instruction improve student-teachers' reading comprehension compared with conventional task-based instruction?
2. How do student-teachers perceive the effects of AI-supported reading tasks on cognitive support for comprehension, engagement, and learner autonomy/self-regulation?

By addressing these questions, the study seeks to advance the understanding of how generative AI can be pedagogically embedded in task-based reading instructions and how such integration may support learners' engagement with academic texts in EFL contexts.

IV. METHODOLOGY

4.1. Research design

The study adopted a quasi-experimental structure featuring baseline and follow-up assessments across treatment and comparison cohorts, within a concurrent mixed-methods framework to examine the impact of AI-integrated instructional practices on learners' reading comprehension. This design enabled a comparison of intact classes while preserving ecological validity (Ellis, 2009; Huang et al., 2024). Because intact classes were used, random assignment was not feasible. Baseline equivalence was established through a placement test confirming comparable English proficiency. The groups were also comparable in age, academic program, and institutional context. Before the intervention, the participants were orally asked about their prior AI familiarity and use, with responses indicating broadly similar AI experience across classes. All the participants confirmed access to smartphones and internet connectivity, which was further verified through their continued participation in digital activities. Although AI experience was assessed informally rather than through a standardized instrument, these checks provided supplementary evidence of comparability in AI experience and digital access. Language-learning motivation was not formally assessed or controlled, so causal interpretations remain cautious.

In line with this design, quantitative learning gains were measured using a standardized 40-item reading comprehension test to address the first research question. Complementary semi-structured interviews explored the learners' perceptions of cognitive support, engagement, and autonomy during AI-mediated reading tasks, providing insight into the processes underlying learning outcomes (Huang & Li, 2024; Zawacki-Richter et al., 2019).

4.2. Institutional context

The study was conducted in the undergraduate teacher-education system of Farhangian University, a public institution responsible for preparing future primary and secondary school teachers in Iran. Admission is based on the national higher education entrance examination, and the students' completion of four-year programs leads to a Bachelor of Arts (B.A.) or Bachelor of Science (B.S.), depending on their field of study.

General English is a compulsory course across all the majors and is typically offered over one 16-week semester with a weekly three-hour session. Rather than training English specialists, the course develops foundational English proficiency, emphasizing reading comprehension and academic vocabulary through general-academic texts shared across disciplines. Because course objectives, instructional hours, and assessment procedures are standardized institutionally, General English follows a consistent structure across campuses. Consequently, the present intervention was implemented in routine curricular practice rather than a specially designed program, enhancing the ecological validity of the findings.

4.3. Participants

A total of 58 male student-teachers aged 19-24 enrolled in General English courses at Farhangian University participated in the study. Because the university provides instruction in gender-segregated classes, only male student-teachers were available for inclusion. The participants represented several non-English disciplines, including social sciences, counseling, and primary education. A placement test confirmed comparable English proficiency (A2-B1, CEFR) across the two intact classes. Before the intervention, the participants were also orally asked about their prior experience with AI tools and confirmed access to smartphones and internet connectivity; digital access was further monitored through ongoing participation in the study activities. Class A (n = 29) served as the control group and received standard TBLT, whereas Class B (n = 29) received AI-enhanced TBLT. Participation was voluntary, and all the students provided written informed consent after being briefed on the research aims, procedures, and ethical safeguards. No participant withdrew; therefore, all the 58 completed the intervention and were included in the analyses.

For the qualitative phase, a purposive subsample of 12 participants from the experimental group completed the semi-structured interviews. The selection ensured variation in proficiency, engagement with AI-supported tasks, and willingness to provide reflective insights, consistent with recommendations for maximizing "information power" (Hennink et al., 2022; Malterud et al., 2016). Based on the recorded AI usage, seven interviewees met the planned engagement benchmark (≥ 500 minutes), whereas five reported lower engagements (< 500 minutes), capturing a range of AI-use experiences. However, the qualitative analysis explored the shared perceptions of AI-supported learning rather than systematically comparing the engagement groups. Because all the reading materials were general-academic rather than discipline-specific, the potential differences associated with the participants' academic backgrounds were minimized.

4.4. Materials and instruments

A 40-item multiple-choice reading comprehension test, aligned with CEFR A2–B1 descriptors and materials commonly used in General English courses, measured the learning gains. The items assessed literal and inferential comprehension, identification of main ideas, and contextual vocabulary interpretation, including recognition of central arguments, supporting information, inferencing, and word meaning from context (Grabe & Stoller, 2020; Ostojić, 2023). Content validity was established through expert review, and pilot testing with a comparable cohort resulted in the removal of six items with low discrimination or inappropriate difficulty. The final instrument demonstrated strong internal consistency (KR-20 = .86, appropriate for dichotomously scored items), consistent with L2 reading research standards (Brown & Abeywickrama, 2019).

Three AI tools including ChatGPT, QuillBot, and Rewordify were integrated into the experimental instruction. ChatGPT provided interactive support through summarization, clarification, and interpretation verification (Baskara & Mukarto, 2023; Etkin et al., 2025). QuillBot supported paraphrasing by encouraging exploration of alternative lexical and syntactic forms (Gürbüz, 2024), whereas Rewordify simplified the complex vocabulary and sentences to reduce the cognitive load for the less proficient readers (Agrawal & Carpuat, 2024; Odo, 2023). These tools functioned as pedagogical scaffolds aligned with specific reading stages and TBLT principles rather than as stand-alone technologies (Ellis, 2009; Y. Li et al., 2025).

Semi-structured interviews with 12 experimental-group participants explored three domains including cognitive support for reading comprehension, engagement and motivation, and learner autonomy and self-regulation. Eight interview questions corresponded to these domains, allowing focused yet flexible exploration. The interviews were conducted in the participants' preferred language (Persian), recorded with consent, and transcribed with minor editing of pauses and repetitions because the analysis focused on intended meaning rather than conversational form. The transcripts were then translated into English, reviewed by an applied linguist to ensure linguistic accuracy and semantic validity, and anonymized (Hashemifardnia & Kooti, 2025; Hennink & Kaiser, 2022; Malterud et al., 2016).

4.5. Procedure

The intervention lasted 12 weeks and comprised twelve 90-minute sessions for both groups, delivered by the same instructor to ensure consistency. The control group followed the standard TBLT cycle of pre-task preparation, task performance, and post-task reflection using authentic texts and collaborative discussion. Thus, AI tools were not incorporated into the planned classroom instruction for the control group. However, the participants' independent use of AI tools outside the intervention was neither restricted nor systematically monitored. The experimental group followed the same sequence, with teacher guidance limited to task instructions and procedural support to maintain comparable instructional conditions. AI support was integrated at each reading stage, including Rewordify for pre-reading lexical and schematic activation, ChatGPT for while-reading comprehension monitoring and inferencing, and QuillBot for post-reading paraphrasing and reflective evaluation. The learners interacted with each tool through structured prompts aligned with the objectives of the corresponding reading stage, while the instructor monitored task completion without providing additional instructional scaffolding.

Beyond the classroom instruction, the experimental-group learners completed approximately eight hours (480 minutes) of supplementary AI-supported reading. AI usage was documented through activity logs and self-reported records. For exploratory analyses, the learners were classified as high-use ($\geq$ 500 minutes) or low-use ($<$ 500 minutes). This benchmark was established a priori as a practical approximation of the planned supplementary practice and distinguished the learners who completed the intended intervention dosage from those with lower engagement, rather than serving as a post hoc statistical threshold.

To enhance methodological transparency and facilitate replication, representative AI-supported tasks were provided. During pre-reading, the learners predicted text content using guiding questions and used Rewordify to simplify the key vocabulary (e.g., "carbon offsetting" and "biodiversity loss") and generate semantic maps linking new vocabulary to prior knowledge. During reading, the learners prompted ChatGPT with questions such as "What evidence in the text supports this conclusion?" and "Explain the relationship between these two paragraphs", which prompted the clarification of causal relationships and referential cohesion. During post-reading, the QuillBot supported paraphrasing of argumentative claims enabled the learners to compare the AI-generated outputs with their own reformulations. This staged integration positioned AI as a scaffold for meaning construction rather than a substitute for cognitive effort.

4.6. Data collection and analysis

The pre- and post-tests were administered under equivalent conditions. The statistical analyses were conducted using SPSS (Version 26). Independent-samples t-tests and ANCOVA, served to control for the pre-test differences and evaluate the instructional effects. Because language-learning motivation was not formally assessed at baseline, it could not be tested for group equivalence or included as a covariate. The effect sizes were estimated using Cohen's d, and the normality was assessed through Shapiro-Wilk tests and visual inspection of Q-Q plots, supporting the conduction of parametric analyses. To examine whether AI engagement was associated with the learning outcomes, the experimental-group participants were classified as high-use ($\geq$ 500 minutes) or low-use ($<$ 500 minutes) based on recorded AI usage. The benchmark, established a priori, approximated the planned supplementary practice and distinguished the learners who completed the intended intervention dosage from those with lower engagement. An independent-samples t-test compared the reading comprehension gain scores between the two groups.

For qualitative analysis, the interview data were analyzed using Braun and Clarke's (2021) six-phase thematic analysis. Open coding identified meaningful units, followed by axial categorization and iterative development of higher-order themes. A second coder independently analyzed one-quarter of the transcripts, yielded an inter-coder agreement of 87% (Cohen's κ = .80), and indicated substantial coding reliability (Cole, 2023). Analytic memoing documented the coding decisions, category development, and theme refinement, while the integration of the qualitative and quantitative findings provided methodological triangulation.

4.7. Rigor and ethical considerations

Methodological rigor was strengthened through expert-reviewed and piloted instruments, standardized instruction, triangulation of quantitative and qualitative data, and partial double-coding of interview transcripts. The documentation of analytic decisions further enhanced dependability and transparency.

The ethical safeguards included voluntary participation, written informed consent, anonymization, secure data storage, and guidance on the responsible use of AI (Marino et al., 2024; Selvam & González Vallejo, 2025). Generative AI was used solely for linguistic editing and stylistic refinement during the manuscript preparation; the study design, data analysis, interpretation, and editorial decisions all remained the authors' responsibility. The study also adhered to established digital research ethics standards to protect the participants' rights, dignity, and well-being throughout the research process. Because the qualitative interviews involved only the experimental-group participants, the findings are interpreted as the describing experiences of AI-supported instruction rather than providing comparative evidence between AI-supported and conventional instruction.

V. RESULTS

5.1. Quantitative findings

The quantitative analyses address the first research question by examining whether AI-mediated task scaffolding produced measurable gains in EFL reading development beyond those observed in traditional instruction. The analytic sequence proceeds in three stages including the confirmation of baseline equivalence, the evaluation of post-intervention differences controlling for initial proficiency, and the examination of whether intensity of AI engagement corresponded with differential gains within the experimental group. AI engagement was estimated using the participants' activity logs and self-reported records collected throughout the intervention. Because there were intact classes, the participants were not randomly assigned at the individual level; however, the two classes were drawn from the same academic cohort and followed the same curriculum before the intervention, thereby minimizing systematic selection bias. Baseline statistical equivalence further supports the comparability of the groups.

5.1.1. Baseline equivalence

An independent-samples t-test was performed to examine whether the two intact classes differed in reading proficiency prior to the intervention (see Table 2). The difference between the control group ($M = 43.1$, $SD = 7.8$) and the experimental group ($M = 42.9$, $SD = 8.0$) was not statistically significant, $t(56) = 0.42$, $p = .67$, $d = 0.03$, 95% CI [-0.49, 0.55].

The near-zero effect size and confidence interval spanning zero indicate practical as well as statistical equivalence at baseline. This comparability strengthens internal validity by reducing the likelihood that subsequent differences reflect pre-existing proficiency disparities.

Table 2. Pre-intervention reading comprehension scores: Independent-samples t-test

Group	N	Mean (SD)	t(56)	p	95% CI (Mean Diff.)	Cohen's d	95% CI for d
Control	29	43.1 (7.8)					
Experimental	29	42.9 (8.0)	0.42	.67	[-3.84, 5.96]	0.03	[-0.49, 0.55]

Note: CI = Confidence Interval. The mean difference is calculated as Experimental-Control.

5.1.2. Assumption checks

Prior to the inferential analyses, the normality assumptions were evaluated using Shapiro-Wilk tests alongside graphical evaluation of Q-Q probability plots. As displayed in Table 3, no statistically significant departures from normality were detected at either testing point ($p > .05$).

In addition, the assumption of homogeneity of regression slopes was examined before conducting ANCOVA. The interaction between the group membership and the pre-test scores was not statistically significant, $F(1, 54) = 0.83$, $p = .37$, indicating that the relationship between the covariate and post-test scores did not differ across the groups. This result supports the appropriateness of ANCOVA for estimating adjusted post-intervention differences. These results collectively support the use of parametric procedures for subsequent analyses.

Table 3. Normality tests for reading comprehension scores

Group	Occasion	Shapiro–Wilk W	df	p
Control	Pre-Test	0.972	29	.623
Control	Post-Test	0.965	29	.451
Experimental	Pre-Test	0.969	29	.557
Experimental	Post-Test	0.976	29	.718

Note: All the p-values exceeded the .05 threshold, demonstrating the absence of statistically meaningful departures from normal distribution.

5.1.3. Reading comprehension performance

To determine the effect of the intervention, the adjusted post-test outcomes were examined through ANCOVA, incorporating baseline performance as a covariate (Table 4). This analysis revealed a statistically significant primary effect of instructional conditions, $F(1, 55) = 9.47$, $p = .003$, $\eta^2 = .15$, 95% CI [0.04, 0.30].

The effect size suggests that approximately 15% of the variance in the adjusted post-test scores was attributable to instructional conditions, indicating a moderate effect of the intervention within the 12-week instructional period. The confidence interval further suggests that even the lower-bound estimate reflects a non-trivial effect. Table 4 summarizes the descriptive statistics and ANCOVA results for the two instructional conditions.

Table 4. Descriptive statistics and ANCOVA results for reading development

Group	Pre-Test M (SD)	Post-Test M (SD)	Gain M (SD)	Adjusted Post-Test M	F(1,55)	p	η^2	95% CI for η^2
Control	43.1 (7.8)	48.2 (6.9)	5.1 (3.0)	48.6				
Experimental	42.9 (8.0)	54.0 (7.1)	11.1 (3.7)	53.8	9.47	.003	.15	[0.04, 0.30]

Note: Adjusted means reflect the post-intervention scores adjusted for baseline performance. η^2 = Partial eta squared.

The descriptive statistics reveal tha,t while both groups improved over time, the magnitude of gain differed substantially. The experimental group demonstrated a mean gain of 11.1 points (SD = 3.7), whereas the participants in the comparison group improved by an average of 5.1 points (SD = 3.0). Although the total sample size was modest (N = 58), the observed effect size ($\eta^2 = .15$) suggests adequate sensitivity to detect medium effects in a classroom-based intervention. Nevertheless, the precision of the estimates should be interpreted in light of sample size, particularly when considering the confidence interval width.

The greater gain observed in the experimental group provides additional support for the effectiveness of the intervention and aligns with the hypothesis that embedding AI-mediated scaffolding in task-based instruction enhances reading development. The adjusted post-test means indicated an estimated mean difference of 5.2 points in favor of the experimental group (95% CI [1.8, 8.6]).

5.1.4. Relationship between AI use and reading gains

To explore whether the extent of AI engagement was associated with the learning outcomes, the experimental participants were classified according to a pre-established instructional benchmark of AI use. Specifically, the learners who completed the planned supplementary AI practice (≥ 500 minutes) were categorized as high-use, whereas those whose recorded engagement remained below the planned dosage (< 500 minutes) were categorized as low-use. The AI usage time was documented through the participants' activity logs and self-reported records collected throughout the intervention.

The 500-minute benchmark represented a practical approximation of the planned eight-hour (approximately 480-minute) supplementary AI practice and was specified a priori to reflect completion of the intended instructional dosage rather than an arbitrary statistical cutoff. As Table 5 shows, an independent-samples t-test revealed a statistically significant difference in the gain scores, $t(27) = 2.41$, $p = .022$, $d = 0.75$, 95% CI [0.01, 1.49].

Table 5. Reading gains by the level of AI engagement

Usage level	N	Mean gain (SD)	t(27)	p	Cohen's d	95% CI for d
High (≥ 500 min)	14	12.3 (3.5)				
Low (< 500 min)	15	9.8 (3.2)	2.41	.022	0.75	[0.01, 1.49]

Note: Mean difference is calculated as High-use vs Low-use.

The observed effect size indicates a moderate difference between the two engagement groups. The learners who completed the planned supplementary AI engagement demonstrated somewhat greater reading improvement than those whose recorded engagement remained below the intended practice level. Although the confidence interval remains above zero, its considerable width reflects the limited precision of this estimate, likely owing to the small subgroup sample; consequently, the magnitude of the observed difference should be interpreted cautiously. The high-engagement learners achieved a mean gain of 12.3 points (SD = 3.5), compared to 9.8 points (SD = 3.2) among the lower-engagement peers.

Given the reduced statistical power inherent in the subgroup analyses, these results should be regarded as suggestive rather than conclusive. However, the consistency of the direction and magnitude of effects strengthens confidence in the observed usage pattern.

5.2. Qualitative findings

The qualitative analysis addresses the second research question, exploring how the participants in the experimental group interpreted and evaluated their experiences with AI-supported reading. The thematic analysis identified three overarching themes: (1) Cognitive Support for Comprehension, (2) Increased Engagement and Motivation, and (3) Perceived Enhanced Autonomy and Self-Regulation. Table 6 presents these themes alongside their associated sub-themes and illustrates each with two representative participant narratives.

Table 6. Themes, sub-themes, and sample narratives from the learners' interviews

Main theme	Sub-theme	Sample narratives (Two per sub-theme)
Cognitive Support for Comprehension	Clarifying difficult vocabulary and structures	"When the text had unfamiliar words, AI explained them simply, and I didn't get stuck." (Interviewee No. 3) "I used AI to paraphrase difficult sentences, and it made the meaning much clearer." (Interviewee No. 1)
	Identifying main ideas and summarizing texts	"ChatGPT helped me pick out the main ideas instead of getting lost in details." (Interviewee No. 9) "The summaries showed me exactly what I should focus on in long texts." (Interviewee No. 5)
	Reducing cognitive load	"Using AI made reading less tiring because I didn't waste time guessing meanings." (Interviewee No. 8) "I felt more relaxed knowing I could check confusing parts quickly." (Interviewee No. 6)
Increased Engagement and Motivation	More enjoyable and interactive tasks	"AI made reading tasks feel more modern and interesting." (Interviewee No. 7) "It didn't feel like typical homework; it felt interactive." (Interviewee No. 10)
	Immediate feedback and support	"I liked the instant answers. It pushed me to keep going." (Interviewee No. 12) "Whenever I had a question, AI responded fast, so I never lost my rhythm." (Interviewee No. 11)
	Sustained effort beyond class time	"I used the tools at home because they made practice enjoyable." (Interviewee No. 2) "I kept reading after class since I could monitor myself easily." (Interviewee No. 9)
Perceived Enhanced Autonomy and Self-Regulation	Independent comprehension monitoring	"I could check my understanding without asking the teacher." (Interviewee No. 4) "AI helped me see where I was wrong and how to fix it." (Interviewee No. 5)
	Planning and managing strategies	"I planned my reading steps better with AI guidance." (Interviewee No. 12) "The tools helped me decide which parts of the text to work on first." (Interviewee No. 8)
	Increased confidence in self-directed learning	"I felt more confident reading alone." (Interviewee No. 11) "AI made me believe I could handle difficult texts on my own." (Interviewee No. 7)

Of the 12 interviewed participants, seven met the planned supplementary AI engagement benchmark (≥ 500 minutes), whereas five reported lower levels of engagement. Although the participants across both engagement levels described broadly similar experiences with AI-supported reading, the qualitative analysis focused on identifying common themes rather than making formal comparisons between the usage groups because of the small interview sample.

5.2.1. Interpretation of themes

The qualitative findings complement the quantitative results by providing insight into how the participants experienced AI-supported reading during the intervention. Rather than accounting for the observed statistical differences, the themes help contextualize the learners' perceptions of the instructional processes associated with the AI-supported activities.

Under the theme of Cognitive Support for Comprehension, the participants consistently reported that AI tools helped them simplify unfamiliar vocabulary, clarify complex sentences, and identify main ideas. The learners described moving from fragmented word-level decoding toward a more coherent understanding of entire passages. Several indicated that, when lexical or syntactic difficulty was reduced through AI assistance, they were able to concentrate more fully on meaning construction. These reports align with the stronger post-test performance observed in the experimental group and suggest that the participants perceived AI support as facilitating more efficient allocation of attention during reading tasks.

The theme of Increased Engagement and Motivation further contextualizes the quantitative gains. The participants frequently characterized AI-supported tasks as interactive and responsive, noting that the tools encouraged sustained effort and curiosity. Some learners reported extending their reading practice beyond required classroom tasks, indicating a shift in task involvement. Notably, the high-use participants demonstrated significantly greater improvement ($d = 0.75$), which mirrors qualitative accounts of more frequent tool interaction and deeper task engagement. This convergence suggests that the participants perceived consistent engagement with AI-supported activities as contributing to their learning during the intervention.

The third theme, Perceived Enhanced Autonomy and Self-Regulation, captures learners' growing strategic awareness. The participants described using AI tools to check interpretations, identify misunderstandings, and adjust reading strategies. Rather than relying passively on generated answers, many learners reported using AI feedback to refine their own reasoning. These accounts parallel the observed performance gains within the experimental group and indicate that the participants perceived strategic interaction with AI as supporting more independent comprehension development.

Taken together, the themes suggest that the participants perceived AI-mediated scaffolding as supporting reading development through certain interconnected processes including facilitating comprehension at points of difficulty, sustaining engagement with tasks, and encouraging more strategic reading behaviors. The alignment between these perceptions and the quantitative outcomes strengthens the internal coherence of the findings. However, because the qualitative data were collected exclusively from the experimental group, the themes should be interpreted as accounts of the learners' experiences with AI-supported instruction rather than as evidence of differences between AI-supported and conventional instruction.

VI. DISCUSSION

The results suggest that AI-supported task-based language teaching (TBLT) was associated with improved reading comprehension among non-English-major student-teachers. Beyond the improved test scores, the participants' accounts revealed a qualitative shift in engagement with texts, suggesting that AI reshaped the cognitive dynamics of task performance rather than simply accelerating answer retrieval. Consistent with interactionist perspectives, the learners in the AI-supported condition engaged in more active meaning-making, indicating that AI tools simulated interactive feedback cycles typically associated with peer- or teacher-mediated exchanges (Huang & Li, 2024; Long, 2015). Vocabulary clarification, sentence reformulation, and guided negotiation of meaning enabled the learners to iteratively reprocess textual information, supporting deeper comprehension beyond surface-level decoding. These findings align with Huang and Li's (2024) socio-cognitive account, which positions AI as a mediational scaffold that supports cognitive efforts while preserving productive struggles, consistent with the Sociocultural Theory (Poehner

& Infante, 2019). Rather than displacing cognitive responsibility, AI redistributed interactional work in ways that maintained learner agency while providing contingent support.

These qualitative findings further suggest that AI-mediated tasks operationalized interactionist mechanisms by structuring iterative cycles of clarification, reformulation, and confirmation during reading. Rather than functioning as an information retrieval tool, AI simulated dialogic feedback conditions that prompted the learners to revisit misunderstandings and refine interpretations, closely reflecting the negotiation processes central to the interactionist accounts of language development (Etkin et al., 2025; Huang & Li, 2024; Long, 2015).

These findings also highlight AI's role in fostering self-regulated learning (SRL). The participants perceived greater awareness of comprehension strategies, goal setting, and progress monitoring, suggesting that AI-supported tasks may enhance metacognitive engagement (Merino-Campos, 2025; Panadero, 2017). Rather than simply providing answers or simplified content, AI encouraged the learners to reflect on their understanding, adjust strategies, and consolidate learning, consistent with previous research on AI-supported autonomy and strategic engagement in second language learning (Huang et al., 2024; W. Li et al., 2025). Integrating AI across the pre-, while-, and post-reading phases supported the learners in anticipating lexical and schematic demands, monitoring comprehension, and critically reflecting on textual meaning (Cain & Oakhill, 2023; Grabe & Stoller, 2020). Together, these findings suggest that AI-mediated scaffolding supports both immediate comprehension and the strategic regulation underlying longer-term reading development.

In parallel, the quantitative analyses revealed statistically significant differences between the AI-supported and traditional TBLT groups, with the former demonstrating greater gains in reading comprehension. However, because the use of AI by the control-group participants outside the planned classroom instruction was not restricted or systematically monitored, potential treatment contamination cannot be ruled out. Any independent exposure to AI could have influenced the observed between-group differences and should, therefore, be considered when interpreting the findings. The exploratory subgroup analysis further suggested that the learners who completed the planned supplementary AI practice achieved larger gains, suggesting that active engagement is important for maximizing AI's pedagogical benefits. This finding aligns with the studies showing that personalization and adaptive feedback may enhance learning outcomes when learners engage consistently (Wang & Fan, 2025; Y. Li et al., 2025). The present study extends this evidence by situating these benefits in structured, reading-specific TBLT tasks. Nevertheless, the subgroup findings should be interpreted cautiously because categorizing AI usage reduces naturally continuous variation in learner engagement.

These results should be interpreted in light of prior research reporting mixed outcomes. For example, earlier studies on AI-based writing or translation tools have sometimes showed modest or negligible effects (Murtisari, 2024; Wei et al., 2023). Such discrepancies likely reflect differences in task type, intervention duration, or cognitive demands. Writing-focused interventions may not sufficiently engage inferencing and meaning negotiation processes critical to reading comprehension, and short-term interventions may not allow learners to internalize strategies or establish meaningful interactions with AI scaffolds. In contrast, the twelve-week duration of this study, combined with structured TBLT tasks, may have enabled gradual skill consolidation and fuller use of AI support. These contrasts underscore the importance of task design, theoretical alignment, and cognitive scaffolding when implementing AI in language learning (Ellis, 2009; Huang et al., 2024). Notably, much prior evidence derives from writing-focused studies; thus, the present reading-focused findings extend, rather than replicate, earlier AI research.

From a theoretical perspective, these findings advance the understanding of AI-mediated scaffolding as a dynamic regulatory layer in task performance, redistributing interactional work traditionally performed by teachers or peers without reducing learner agency. Therefore, AI functions as a pedagogically embedded mediator that enhances the interactional and cognitive affordances of reading tasks. The results support a multidimensional framework in which interactionist, sociocultural, and self-regulated learning perspectives converge, enabling AI to mediate comprehension, promote strategic processing, and strengthen metacognitive awareness. They also address a gap in previous research, which has focused largely on writing or isolated language skills (W. Li et al., 2025; Wei et al., 2023), by demonstrating the feasibility of AI-integrated task-based reading for non-English-major student-teachers, a population rarely examined in earlier studies (Semerikov et al., 2024; Yousefi & Askari, 2024). Accordingly, the study offers an integrated account of how AI affordances, task structure, and learner regulation interact to support reading development. Unlike earlier CALL/TELL approaches that primarily provided digital resources or automated feedback, the present framework conceptualizes generative AI as an adaptive mediator integrating interactional feedback, contingent scaffolding, and self-regulatory support in a unified pedagogical model.

Taken together, these patterns suggest that AI-supported TBLT does not merely supplement existing instructional routines but recalibrates the cognitive ecology of reading tasks. By embedding contingent prompts in the task cycle, the intervention reshaped learners' allocation of attention, evaluation of comprehension breakdowns, and persistence through ambiguity. This shift from reactive problem-solving to proactive monitoring reflects a qualitative transformation in strategic engagement rather than a simple performance gain. Recent evidence suggests that GPT-based tools support iterative attention and comprehension monitoring (Etkin et al., 2025), while AI scaffolds embedded in task cycles enhance strategic engagement and reflection (Huang et al., 2024). Adaptive AI feedback can further recalibrate metacognitive monitoring by providing timely cues to evaluate understanding and adjust strategies in real time (Merino-Campos, 2025). Collectively, these findings highlight that AI's instructional value lies less in automation and more in orchestrating structured, feedback-rich interaction in pedagogically principled tasks.

Practical insights regarding engagement and usage also emerge. Although the ≥500-minute threshold facilitated subgroup comparisons, it represented a practical approximation of the planned supplementary practice rather than an optimal benchmark. Future studies should, therefore, examine AI engagement as a continuous predictor to better understand dose-response relationships in AI-supported language learning (Huang et al., 2024; Merino-Campos, 2025).

The learners' affective and motivational responses further underscored the broader impact of AI scaffolds. The participants reported reduced perceived cognitive demands and increased confidence. These findings are consistent with the Cognitive Load Theory (Sweller et al., 2019), while remaining actively engaged in negotiating meaning. The affective benefits align with evidence that AI literacy, self-regulation, and strategic engagement jointly influence learning outcomes (Y. Li et al., 2025; Yaşar & Karagücük, 2024). By integrating cognitive, metacognitive, and affective dimensions, the present study extends the research that has often examined these processes separately (Cain & Oakhill, 2023; Huang & Li, 2024). Because the qualitative data were collected only from the experimental group, these findings should be interpreted as participants' perceptions rather than comparative evidence. They nevertheless suggest that reading development in AI-mediated environments reflects the interaction of scaffolding, strategy use, and learner perceptions rather than technological efficiency alone. Despite these benefits, AI-supported scaffolding may also introduce challenges. Excessive reliance on AI-generated prompts could reduce learners' independent inferencing if support is not carefully calibrated. The findings, therefore, reinforce the importance of using AI as contingent instructional mediation rather than as a substitute for learners' cognitive effort.

In conclusion, while the study demonstrates positive outcomes, contextual and methodological considerations must be acknowledged. Task design, learner heterogeneity, and digital literacy influence the efficacy of AI-supported reading, emphasizing that technology is not inherently transformative. AI functions as a mediator of strategically structured learning experiences, reinforcing the principle that pedagogical alignment remains central to educational impact (Ellis, 2009; Poehner & Infante, 2019). Recognizing these contingencies provides a balanced perspective, highlighting that any observed gain results from principled pedagogical integration rather than technological novelty alone.

VII. CONCLUSION AND IMPLICATIONS

This study provides evidence that strategically embedding artificial intelligence (AI) in task-based language teaching (TBLT) can enhance the reading comprehension of non-English-major student-teachers. The learners receiving the AI-supported TBLT intervention achieved greater gains than those receiving conventional TBLT under the study conditions, although uncontrolled AI exposure outside the intervention cannot be excluded. The qualitative findings suggested that AI promoted vocabulary clarification, sentence reformulation, immediate feedback, and meaning negotiation. The participants also reported greater autonomy and metacognitive control, indicating that thoughtfully integrated AI can foster strategic engagement (Huang & Li, 2024; Merino-Campos, 2025).

Taken together, the findings support a theoretically integrated account in which AI-mediated support is associated with interactional meaning negotiation, sociocultural mediation, and self-regulated learning. Rather than replacing instruction, AI functioned as a pedagogically embedded mediator that supported deeper cognitive and strategic engagement during the reading tasks.

These findings have important implications for tertiary education. AI can reinforce TBLT by providing personalized support, particularly in large or diverse classrooms where individualized feedback is difficult to sustain. In practice, AI-supported pre-reading vocabulary activation, while-reading comprehension monitoring, and post-reading reflective paraphrasing can be systematically incorporated into TBLT syllabi and formative assessment activities. Teacher education programs should, therefore, prepare future educators to use AI critically, ethically, and in ways that align with sound pedagogy, emphasizing AI literacy (Marino, Liska, & Kessler, 2024; W. Li et al., 2025). Educators should also recognize that AI-generated responses may occasionally contain inaccuracies or bias; therefore, AI outputs require critical evaluation rather than unquestioned acceptance. Institutional investment in transparent implementation and professional development will remain essential to maintaining academic integrity while leveraging AI effectively.

Several limitations should be acknowledged. The quasi-experimental design, modest sample size, twelve-week intervention, and inclusion of only male student-teachers from gender-segregated classes limit the generalizability of the results. The individual differences in digital literacy and motivation were not measured and may have influenced engagement (Semerikov et al., 2024; Yaşar & Karagücük, 2024). Moreover, although AI tools were not incorporated into the control group's planned instruction, the participants' independent AI use outside the intervention was not systematically monitored; thus, potential treatment contamination cannot be excluded. In addition, the qualitative findings reflect only the AI group's perspectives; no delayed post-test examined long-term retention, and the exploratory ≥ 500-minute usage comparison should be interpreted cautiously. Future research should include more diverse learner populations, monitor or control AI exposure across study conditions, model AI engagement as a continuous variable, investigate long-term retention, and examine the gradual fading of AI support.

Overall, this study offers a theoretically grounded and potentially scalable model for integrating AI into TBLT to support cognitive, social, and self-regulatory dimensions of academic reading.

ACKNOWLEDGMENTS

The author extends sincere gratitude to the student-teachers who generously contributed their time to participate in this study.

USE OF GENERATIVE AI

The author reports limited use of AI-supported applications (e.g., Grammarly and ChatGPT) exclusively for linguistic editing and stylistic refinement.

AUTHORS' BIOGRAPHIES

Ali Hemmati is an Assistant Professor of English Language at Farhangian University, Tehran, Iran. His Primary research interest focuses on discourse analysis and English language teaching.

REFERENCES

Agrawal, S., & Carpuat, M. (2024). Do text simplification systems preserve meaning? A human evaluation via reading comprehension. Transactions of the Association for Computational Linguistics, 12, 432–448. https://doi.org/10.1162/tacl_a_00653

Baskara, R., & Mukarto, M. (2023). Exploring the implications of ChatGPT for language learning in higher education. Indonesian Journal of English Language Teaching and Applied Linguistics, 7(2), 343–358. https://doi.org/10.21093/ijeltal.v7i2.1387

Braun, V., & Clarke, V. (2021). One size fits all? What counts as quality practice in (reflexive) thematic analysis? Qualitative Research in Psychology, 18(3), 328–352. https://doi.org/10.1080/14780887.2020.1769238

Braun, V., & Clarke, V. (2023). Toward good practice in thematic analysis: Avoiding common problems and be(com)ing a knowing researcher. International Journal of Transgender Health, 24(1), 1–6. https://doi.org/10.1080/26895269.2022.2129597

Brown, H. D., & Abeywickrama, P. (2019). Language assessment: Principles and classroom practices (3rd ed.). Pearson.

Cain, K., & Oakhill, J. (2023). Children's text comprehension from theory & research to support & intervention. Pedagogical Linguistics, 4(1), 1-25. https://doi.org/10.1075/pl.21015.oak

Cole, R. (2023). Inter-rater reliability methods in qualitative case study research. Sociological Methods & Research, 53(4), 1944–1975. https://doi.org/10.1177/00491241231156971

Ellis, R. (2009). Task-based language teaching: Sorting out the misunderstandings. International Journal of Applied Linguistics, 19(3), 221–246. https://doi.org/10.1111/j.1473-4192.2009.00231.x

Etkin, H. K., Etkin, K. J., Carter, R. J., & Rolle, C. E. (2025). Differential effects of GPT-based tools on comprehension of standardized passages. Frontiers in Education, 10, Article 1506752. https://doi.org/10.3389/feduc.2025.1506752

Grabe, W., & Stoller, F. L. (2020). Teaching and researching reading (3rd ed.). Routledge. https://www.routledge.com/Teaching-and-Researching-Reading/Grabe Stoller/p/book/9781138847941

Gürbüz, N. (2024). The impact of QuillBot as an automated writing evaluation tool on EFL learners. Journal of Education in Science, Mathematics and Art (JESMA), 9, Article 101. https://doi.org/10.51383/jesma.2024.101

Hashemifardnia, A., & Kooti, M. (2025). AI‑Mediated Language Learning and EFL Learners' Self‑Confidence, Self‑Regulation, Well‑Being, and L2 Motivation: A Mixed Method Study. English Education Journal, 16(2), 109‑124. https://doi.org/10.24815/eej.v16i2.45696

Hennink, M., & Kaiser, B. N. (2022). Sample sizes for saturation in qualitative research: A systematic review of empirical tests. Social Science & Medicine, 292, Article 114523. https://doi.org/10.1016/j.socscimed.2021.114523

Huang, J., & Li, M. (2024). Artificial intelligence as scaffolding in second language reading: A sociocognitive perspective. Computer Assisted Language Learning, 37(5–6), 1121–1144. https://doi.org/10.1080/09588221.2023.2234567

Huang, X., Wu, Y., & Dou, A. (2024). AI-enhanced task-based language teaching: Fostering personalized college English learning. Frontiers in Educational Research, 7(10), 204-209. https://doi.org/10.25236/FER.2024.071033

Kartal, G., & Yeşilyurt, Y. E. (2024). A bibliometric analysis of artificial intelligence in L2 teaching and applied linguistics between 1995 and 2022. ReCALL, 36(3), 359–375. https://doi.org/10.1017/S0958344024000077

Li, W., Lu, N., & Zhuang, J. (2025). A review of empirical studies on second language writing feedback in the background of artificial intelligence (2014–2024). Education, Science, Technology, Innovation and Life, 7(1), 44–59. https://doi.org/ 10.23977/trance.2025.070107

Li, Y., Shadiev, R., & Chiu, T. K. F. (2025). Applying generative artificial intelligence to task-based language teaching and learning: A systematic review and meta-analysis. TechTrends, 70, 150-163. https://doi.org/10.1007/s11528-025-01140-7

Liao, R. J. T., & Lee, K. (2024). The interplay between metacognitive knowledge, L2 language proficiency, and question formats in predicting L2 reading test scores. Education Sciences, 14(4), 370. https://doi.org/10.3390/educsci14040370

Long, M. H. (2015). Second language acquisition and task-based language teaching. Wiley-Blackwell.

Malterud, K., Siersma, V. D., & Guassora, A. D. (2016). Sample size in qualitative interview studies: Guided by information power. Qualitative Health Research, 26(13), 1753 1760. https://doi.org/10.1177/1049732315617444

Marino, F., Liska, D., & Kessler, M. (2024). Ethical considerations for research involving computer-assisted language learning, social media, and online environments. In P. I. De Costa, A. Rabie-Ahmed, & C. Cinaglia (Eds.). Ethical issues in Applied Linguistics scholarship (pp. 72–86). John Benjamins. https://doi.org/10.1075/rmal.7

Merino-Campos, C. (2025). The impact of artificial intelligence on personalized learning in higher education: A systematic review. Trends in Higher Education, 4(2), 1-15, Article 17. https://doi.org/10.3390/higheredu4020017

Moslemi Nezhad Arani, S. (2024). Navigating the Future of Language Learning: A Conceptual Review of AI's Role in Personalized Learning. Computer‑Assisted Language Learning Electronic Journal, 25(3). 1-22. https://callej.org/index.php/journal/article/view/78

Murtisari, E. T. (2024). Self-directed use of machine translation among language learners. Foreign Language Annals, 57(4), 1094-1114. https://doi.org/10.1111/flan.12768

Odo, D. M. (2023). The effect of automatic text simplification on L2 readers' text comprehension. Applied Linguistics, 44(6), 1030–1046. https://doi.org/10.1093/applin/amac057

Ostojić, A. B. (2023). Reading comprehension processes: A review based on theoretical models and research methodology. Hrvatska Revija Za Rehabilitacijska Istrazivanja, 59(1), 122–143. https://doi.org/10.31299/hrri.59.1.8

Panadero, E. (2017). A Review of Self-Regulated Learning: Six Models and Four Directions for Research. Frontiers in Psychology, 8, 422. https://doi.org/10.3389/fpsyg.2017.00422

Park, G. J. (2024). Exploring the influence of machine translation on L2 reading comprehension. English Teaching, 79(4), 33-54. https://files.eric.ed.gov/fulltext/EJ1459825.pdf

Poehner, M., & Infante, P. (2019). Mediated development and the internalization of psychological tools in second language education. Learning, Culture and Social Interaction, 22, 100322. https://doi.org/10.1016/jcsi.2019.100322

Selvam, M., & González Vallejo, R. (2025). Ethical and privacy considerations in AI-driven language learning. LatIA, 3, 328–342. https://doi.org/10.62486/latia2025328

Semerikov, S. O., Striuk, A. M., & Shalatska, H. M. (2024). AI‑Assisted Language Education: Critical Review. Educational Dimension, 4, 1-7. https://doi.org/10.31812/ed.623

Sweller, J., Ayres, P., & Kalyuga, S. (2019). Cognitive load theory (2nd ed.). Springer.

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Harvard University Press.

Wang, J., & Fan, W. (2025). The effect of ChatGPT on students' learning performance: Insights from a meta-analysis. Humanities and Social Sciences Communications, 12(1). Article 621. https://doi.org/10.1038/s41599-025-04787-y

Wei, P., Wang, X., & Dong, H. (2023). The impact of automated writing evaluation on second language writing skills of Chinese EFL learners: A randomized controlled trial. Frontiers in Psychology, 14, 1249991. https://doi.org/10.3389/fpsyg.2023.1249991

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. Journal of Child Psychology and Psychiatry, 17(2), 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

Yaşar, H., & Karagücük, V. (2024). Exploring the relationship between artificial intelligence literacy and English language learning motivation. International Journal of Languages' Education and Teaching, 12(4), 107–124. https://doi.org/10.71084/ijlet.1561914

Yousefi, M., & Askari, M. I. (2024). Exploring the Effectiveness of Artificial Intelligence (AI) on Reading Comprehension among Iranian EFL Learners. Journal of New Trends in English Language Learning (JNTELL), 3(5), 1-13. https://doi.org/10.71528/2024.202406221123351

Zawacki‑Richter, O., Marin, V., Bond, M., & Gouvernuer, 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/s41239-019-01710.