Canadian Journal of Language and Literature Studies
Vol. 6(1), 2026, pp.89-105

Optimizing Intrinsic Cognitive Load in AI-Supported EFL Vocabulary Learning: Evidence from a Saudi Preparatory-Year Program

Afsaneh Amini Baghbadorani[1]

[1] Ph.D., Umm Al-Qura University, Saudi Arabia
Email: aabaghbadorani@uqu.edu.sa; dr.afsaneh.amini@gmail.com

DOI: 10.53103/cjlls.v6i1.260

Abstract

Vocabulary acquisition is a central component of English as a Foreign Language (EFL) learning, yet it places considerable demands on learners' working memory. Grounded in Cognitive Load Theory (CLT), this study examines whether optimizing intrinsic cognitive load through targeted pedagogical strategies and artificial intelligence (AI) enhances vocabulary learning. Using a quasi-experimental design, Saudi preparatory-year students learned 30 target words from *Evolve 2: Special Edition* under an optimized condition that incorporated AI-assisted tools and cognitively informed instructional strategies (n = 38), while another 30 words were taught through traditional teacher-centered instruction in a non-optimized condition (n = 35). Vocabulary achievement was measured using two parallel 30-item multiple-choice tests, each aligned with the vocabulary taught in its respective condition. Results showed a statistically significant advantage for the optimized condition (M = 96%) over the non-optimized condition (M = 73%), $t(71) = 14.85$, $p < .001$, with a very large effect size (d = 3.48). These findings provide strong evidence for the effectiveness of AI-supported, cognitive load–optimized vocabulary instruction in EFL contexts.

Keywords: Cognitive Load Theory, EFL Vocabulary, Artificial Intelligence, Preparatory Year, Saudi Arabia

Introduction

Vocabulary knowledge forms a central component of second and foreign language proficiency, supporting learners' ability to comprehend texts, express ideas, and participate meaningfully in communicative tasks (Nation, 2013). In many EFL settings, however, students encounter substantial cognitive demands because their exposure to English outside the classroom is limited. The need to learn large numbers of new words within restricted instructional time can place considerable strain on working memory, often resulting in weak retention and reduced engagement.

Cognitive Load Theory (CLT) offers a useful lens for designing instruction that takes into account the constraints of working memory (Sweller, Ayres, & Kalyuga, 2011). CLT differentiates among intrinsic load, which is shaped by the inherent complexity of a task; extraneous load, which stems from ineffective instructional design; and germane load, which supports the development of mental schemas. Vocabulary learning typically carries a high intrinsic load because learners must attend simultaneously to form, meaning, pronunciation, and usage (Nation, 2013).

Recent developments in artificial intelligence (AI) have introduced new possibilities for managing cognitive load in educational contexts. AI-based tools can offer adaptive instruction, individualized feedback, and targeted scaffolding, helping learners process complex material more efficiently (Mayer, 2020). Despite this potential, empirical research examining how AI-supported instruction influences intrinsic cognitive load in EFL vocabulary learning remains limited.

The present study investigates how specific teaching strategies, together with AI-enhanced support, can optimize intrinsic cognitive load and improve vocabulary acquisition among preparatory-year EFL learners.

Literature Review

Vocabulary Learning in EFL

Vocabulary learning in English as a Foreign Language (EFL) contexts is cognitively demanding because learners must integrate multiple dimensions of lexical knowledge, including orthography, phonology, semantics, collocation, and contextual usage (Nation, 2013; Laufer & Goldstein, 2004). Within the framework of Cognitive Load Theory (Sweller, Ayres, & Kalyuga, 2011), each of these dimensions constitutes an interactive element, and the high level of element interactivity can easily exceed the limits of working memory, particularly for learners with restricted prior lexical knowledge. Traditional instructional practices—such as decontextualized word lists, rote memorization, and translation exercises—often exacerbate this challenge by increasing intrinsic cognitive load and reducing efficiency (Schmitt, 2008; Webb, 2007). These approaches tend to foster shallow processing and limited retention, as learners are not encouraged to build meaningful associations or develop robust mental representations of vocabulary. In contrast, pedagogical strategies that embed words in rich contexts, integrate pronunciation and usage, and provide scaffolding have been shown to reduce cognitive demands and promote deeper, long-term acquisition (Nation, 2013; Webb & Nation, 2017).

Cognitive Load Theory (CLT)

Cognitive Load Theory highlights the limited capacity of working memory and underscores the role of instructional design in supporting effective learning (Sweller, 1988). From this perspective, well-designed instruction seeks to manage intrinsic cognitive load, minimize unnecessary or extraneous demands, and encourage germane processing that contributes to the development of stable mental schemas. In the context of vocabulary instruction, this means organizing learning activities in ways that take account of element interactivity and sequencing tasks so they align with learners' existing knowledge (Sweller et al., 2011).

Strategies for Optimizing Intrinsic Load in Vocabulary Instruction

Research in cognitive psychology and applied linguistics shows that vocabulary learning improves when instructional approaches are intentionally structured to manage cognitive load and strengthen long-term retention. One widely supported method is pre-training or lexical priming, in which key vocabulary items are introduced before learners encounter them in a text, lecture, or communicative activity. Early exposure to word forms and core meanings reduces both intrinsic and extraneous load during subsequent tasks, allowing learners to devote more working memory resources to comprehension and higher-level processing. Drawing on CLT and the Cognitive Theory of Multimedia Learning, Mayer (2020) argues that pre-training helps learners establish essential mental representations that facilitate deeper understanding when they later engage with more complex input.

Semantic grouping and chunking offer another effective means of managing cognitive demands. Organizing vocabulary into meaningful categories—whether thematic, functional, or semantic—supports learners in integrating new items into existing cognitive structures rather than treating each word as an isolated unit. Miller's (1956) work on the limits of working memory demonstrates that chunking information into coherent units increases the amount that can be processed and retained. In vocabulary learning, such grouping helps learners recognize relationships among words, which enhances recall and promotes more efficient retrieval during language use.

Retrieval practice and spaced repetition also play a central role in long-term vocabulary development. Retrieval practice requires learners to actively recall lexical items, strengthening memory traces and improving the likelihood that words will transfer to new contexts. When paired with spaced repetition—where encounters with vocabulary are distributed over time rather than concentrated in a single session—this approach helps prevent forgetting and reduces the cognitive strain associated with massed practice. Nation (2013) emphasizes that repeated, spaced encounters with words across varied contexts are essential for consolidating both form and meaning, ultimately supporting more fluent and automatic vocabulary use. Overall, these strategies illustrate the value of aligning vocabulary instruction with established cognitive principles to enhance learning efficiency and retention.

Artificial Intelligence in Vocabulary Learning

Artificial intelligence has become an increasingly influential tool in vocabulary instruction, offering adaptive scaffolding, immediate feedback, and individualized practice opportunities. By adjusting task difficulty and pacing in response to learner performance, AI-based systems can reduce unnecessary cognitive demands and help optimize intrinsic load during vocabulary learning (Zawacki-Richter et al., 2019). When combined with principles drawn from Cognitive Load Theory, AI-supported instruction has the potential to create more efficient and responsive learning environments for EFL learners.

Recent research highlights the growing role of adaptive technologies in promoting more effective and autonomous vocabulary development. Studies consistently show that AI systems can tailor instructional pathways to individual learners, modulating challenge levels and providing targeted scaffolding that supports more efficient processing (Zawacki Richter et al., 2019). More recent work extends this understanding by demonstrating how AI driven platforms deliver immediate, context sensitive feedback and personalized practice sequences that strengthen retention and deepen semantic understanding. Emerging studies published in 2025 further emphasize that AI tools, from generative models to intelligent tutoring systems, encourage learner autonomy by enabling self directed exploration of word meanings, collocations, and pragmatic usage beyond the constraints of classroom time (Behforouz and Al Ghaithi, 2025; Yu, 2025; Li, Chen, and Wang, 2025; Xu et al., 2025). Collectively, this research suggests that integrating AI with communicative language teaching principles can create richer and more responsive instructional conditions that substantially enhance vocabulary acquisition.

Generative language models such as ChatGPT have also gained attention for their capacity to provide contextualized input and interactive feedback on demand. Unlike static digital resources, these models can offer tailored explanations, varied example sentences, and immediate corrective feedback, all of which support deeper lexical processing. Kasneci et al. (2023) note that large language models can adapt their responses to learners' questions and proficiency levels, thereby reducing extraneous cognitive load and supporting more efficient knowledge construction. Similarly, Kohnke, Zou, and Zhang (2023) argue that generative AI fosters learner autonomy by enabling self-directed practice, allowing learners to explore word meanings, collocations, and pragmatic usage independently. This type of engagement aligns with research showing that active, learner-controlled vocabulary practice contributes to stronger retention and transfer.

Research on AI-enhanced digital games and adaptive learning platforms further illustrates the motivational and cognitive benefits of real-time adaptation in vocabulary instruction. Godwin-Jones (2019) observes that AI-driven systems can dynamically adjust task difficulty, pacing, and feedback based on learner performance, maintaining an optimal level of challenge that sustains engagement. Petersen (2018) similarly demonstrates that adaptive feedback mechanisms help learners remain within their zone of proximal development, supporting long-term retention and motivation. Taken together, this body of work indicates that AI-based instructional environments not only improve vocabulary learning outcomes but also promote sustained engagement and greater learner autonomy in EFL contexts.

.

Adaptive AI and Cognitive Load

AI-enriched digital games and adaptive learning platforms have become increasingly important in managing the cognitive demands associated with vocabulary learning. Because intrinsic cognitive load is shaped by the inherent complexity of learning materials and the degree of element interactivity relative to learners' existing knowledge structures (Sweller, Ayres, & Kalyuga, 2011), vocabulary learning in EFL contexts can be particularly taxing. Learners must attend to word form, meaning, pronunciation, collocation, and contextual usage at the same time, and when these elements are introduced without careful sequencing, the resulting cognitive burden can impede learning.

Adaptive AI systems address this challenge by continuously analyzing learner performance and adjusting task difficulty, feedback, and pacing to match individual proficiency levels (Petersen, 2018). Such systems can simplify lexical input, recycle previously encountered vocabulary, or gradually increase contextual complexity as learners' schemas develop. This real-time modulation helps maintain learners within an optimal range of intrinsic cognitive load, supporting more efficient schema construction and promoting long-term retention (Sweller et al., 2019).

Recent studies also highlight the motivational benefits of adaptive AI environments, which indirectly contribute to cognitive efficiency. Godwin-Jones (2019) notes that personalized challenges, immediate feedback, and interactive, game-like features sustain attention and encourage repeated, meaningful exposure to vocabulary. More recent research suggests that AI-supported adaptive learning not only improves vocabulary performance but also fosters autonomy and self-regulation—skills that are essential for managing cognitive resources effectively in complex learning tasks (Zawacki-Richter et al., 2019; Huang et al., 2023). Taken together, these findings indicate that adaptive AI systems enhance EFL vocabulary acquisition by strategically managing intrinsic cognitive load while simultaneously promoting sustained engagement and independent learning.

Research Questions

1. Does optimizing intrinsic cognitive load through targeted teaching strategies and

AI support improve EFL vocabulary learning?

2. Is the difference in vocabulary achievement between optimized and non-optimized instruction statistically significant?

Methodology

Research Design

A quasi-experimental within-cohort design was used for this study. The same group of preparatory-year EFL learners experienced two contrasting instructional conditions:

1. Optimized condition: 30 vocabulary items from Units 7–9 taught using AI-supported instruction combined with cognitive load–informed strategies.
2. Non-optimized condition: 30 vocabulary items from Units 3–5 taught through a traditional, teacher-centered approach.

Using a within-cohort design allowed each learner to serve as their own comparison point, thereby reducing the influence of individual differences such as proficiency level, motivation, or learning preferences. In the optimized condition, instruction incorporated AI-based scaffolding alongside strategies grounded in cognitive load theory, including pre-training, semantic chunking, and limited spaced retrieval. In contrast, the non-optimized condition relied on conventional teacher-led explanations and Arabic translations of word lists. Teaching an equivalent number of vocabulary items across two sets of units enabled a systematic comparison of instructional effectiveness while minimizing confounding variables related to content difficulty or learner characteristics.

Participants

The study involved 73 preparatory year students enrolled in an English foundation program at Umm Al-Qura University in Saudi Arabia. The same group of students participated in both the optimized and non-optimized instructional conditions, with 41 students originally assigned to each condition. Due to absences during one or more instructional sessions, three students were removed from the optimized condition and six from the non-optimized condition, resulting in final totals of 38 students in the optimized condition and 35 students in the non-optimized condition.

Instructional Materials

Instructional materials were drawn from *Cambridge Evolve 2: Special Edition*, a textbook designed for learners at the CEFR A2 level. The book follows a linear thematic progression across twelve units, each centered on communicative themes such as daily routines, work and study, invitations, and past experiences. Although the themes vary, the pedagogical structure remains consistent, offering learners a predictable sequence that supports gradual skill development.

Each unit begins with an introductory page that activates prior knowledge and presents key vocabulary. This is followed by two core lessons integrating grammar and vocabulary through contextualized activities that encourage learners to notice patterns and use language meaningfully. Additional skills-focused sections develop listening, reading, speaking, and writing abilities, while pronunciation work is embedded throughout. Units conclude with a "Time to Speak" task requiring extended interaction, problem-solving, or decision-making. Review and consolidation activities, available in print and digital formats, provide opportunities for repeated exposure and reinforcement.

Although the textbook remains within the A2 proficiency band, linguistic and cognitive demands increase gradually across units. Early units introduce high-frequency, concrete vocabulary and simple grammatical structures, whereas later units incorporate more abstract or functional vocabulary and more complex grammar. Communicative tasks also become more demanding, shifting from short exchanges to activities requiring opinions, explanations, or narrative skills.

For this study, vocabulary items were selected from *Evolve 2: Special Edition*. The optimized condition used target words from Units 7–9, while the non-optimized condition used vocabulary from Units 3–5. Units 3–5 were intentionally chosen for the non-optimized condition because they are slightly less complex than Units 7–9. This decision reduced the likelihood that differences in vocabulary difficulty would influence outcomes. It also ensured that students in the non-optimized condition could follow the lessons without the cognitive supports provided in the optimized condition, where intrinsic load was intentionally reduced through structured strategies and AI-based assistance.

Instructional Procedure: Non-Optimized Condition

In the non-optimized instructional condition, vocabulary teaching relied primarily on traditional, teacher-led explanations supported by lists of new words accompanied by direct Arabic translations. These lists were presented in isolation, with little attention to contextualization or to the semantic relationships among items. Scaffolding was minimal, as learners received few guided activities or structured supports to aid processing or retention. Instruction proceeded at a uniform whole-class pace, a pattern that risked increasing cognitive load for some students while offering insufficient challenge for others. No AI-based tools were used in this condition, meaning that learners did not receive adaptive examples, individualized practice, or immediate feedback. Overall, the instructional approach emphasized exposure and explanation rather than processes designed to support cognitive efficiency or deeper lexical engagement.

Execution of the Lesson Plan in the Non-Optimized Condition

Vocabulary instruction in the non-optimized condition was delivered across three 90-minute sessions, each introducing 15 new words from a designated unit in *Evolve 2: Special Edition*. Unlike the optimized condition, this approach relied almost entirely on teacher explanations and translation-based word lists, with limited scaffolding and no AI support.

Session 1 (Unit 3**)** began with a 15-minute warm-up and pre-teaching phase in which the teacher introduced the first 15 target words through direct explanation and translation into Arabic. Students had already received the word list via WhatsApp and were able to review it beforehand. The subsequent 15-minute presentation phase consisted of teacher-led pronunciation practice, choral repetition (though participation was limited), individual repetition, and several cycles of pronunciation and meaning checks. During the 30-minute guided practice phase, students worked individually or in pairs on tasks focused primarily on memorizing English–Arabic and Arabic–English word pairs. Activities included matching exercises, controlled sentence writing using board-provided models, and teacher-directed analysis of sentence components such as subject–verb–object patterns. Additional tasks included word-class categorization, fill-in-the-blank exercises dictated by the teacher, and brief comprehension checks. A 15-minute consolidation phase reinforced memory through repeated oral practice and flash-card drills. The session concluded with a 15-minute wrap-up in which students wrote three to five sentences using at least five of the new words, followed by general teacher feedback. The students were encouraged to extend their practice at home.

Session 2 (Unit 4) followed a similar structure. The lesson opened with a brief review of previously taught vocabulary, after which the next 15 items were presented in list form with Arabic translations and short examples. Students again received the list in advance via WhatsApp. Pronunciation and meaning practice were teacher-led, though choral repetition remained weak, resulting in greater reliance on individual repetition. The students completed controlled written tasks—matching, fill-in-the-blank exercises, controlled sentence writing, and word-sorting activities—working individually or in pairs. Teacher questioning continued throughout, with students called on to provide meanings in English or Arabic. The session ended with a sentence writing task similar to Session 1, followed by general feedback. The students were advised to keep practicing at home.

Session 3 (Unit 5) introduced the final 15 vocabulary items using the same teacher-centered, translation-based approach. The session began with a brief review of earlier vocabulary, followed by pronunciation and meaning practice. As in previous sessions, group repetition was limited, and individual repetition became the primary mode of practice. Students completed written tasks such as matching exercises, controlled sentence writing, translation drills, and word-sorting activities. The session concluded with a short sentence writing task and general teacher feedback, and the students were encouraged to continue practicing at home.

Key Instructional Features of the Non-Optimized Condition

The following observations summarize the key instructional characteristics of the non- optimized condition:

1. Vocabulary was introduced in English with Arabic translations, but contextualization was minimal and offered limited support for deeper semantic processing.
2. Instruction relied heavily on teacher-led pronunciation, explanation, translation, word lists, matching tasks, sentence-pattern analysis, and fill-in-the-blank exercises, with no AI-based scaffolding to individualize learning.
3. Scaffolding overall was limited, as instruction proceeded at a uniform whole-class pace without adjustments for differing learner needs.
4. Retrieval practice occurred through cumulative review at the end of each session and at the start of the next, with additional encouragement for home practice.
5. Learners depended largely on memorization and translation for comprehension, and their production tasks remained basic.

Instructional Procedure: Optimized Condition

In the optimized instructional condition, teaching procedures were deliberately structured to reduce cognitive load and strengthen vocabulary learning through evidence-based strategies as well AI assistance. Instruction began with semantic chunking and pre-training, during which target vocabulary items were introduced in advance and grouped into meaningful semantic categories. This sequencing was intended to activate relevant background knowledge, support early schema formation, reduce cognitive load, and enable learners to process new input more efficiently when encountering the words in texts or communicative tasks.

Instruction was further supported through AI-assisted examples and exercises that provided contextualized usage, multiple exemplars, and guided practice tailored to individual learner needs. These AI-generated materials allowed students to interact with vocabulary across varied contexts, promoting deeper semantic processing. At the start of the first session, students were asked about their prior experience with artificial intelligence applications, and all indicated familiarity with such tools. They were then introduced to several mobile and web-based AI-supported vocabulary applications—including Memrise, Busuu, LingQ, and ChatGPT—which personalize vocabulary exposure through features such as spaced repetition, adaptive practice, and contextualized examples. These tools were introduced for use during class activities and, if students wished, for independent study outside the classroom.

To support long-term retention, students were encouraged to engage in spaced retrieval and independent practice at home, although completion of this practice was not monitored in the present study. Each teaching session also began with a brief review of the previous lesson, reinforcing memory consolidation through active recall. During instruction, adaptive feedback was provided in response to individual errors, ensuring that learners received immediate and focused clarification on specific lexical misunderstandings. This personalized feedback mechanism fostered self-regulation and reduced cognitive load, contributing to more efficient and sustained vocabulary learning outcomes.

Execution of the Lesson Plan in the Optimized Condition

As in the non-optimized condition, the vocabulary component of *Evolve 2: Special Edition* was delivered across three 90-minute sessions, with each session introducing 15 target words as preparation for the corresponding unit. Although a full unit in *Evolve 2* typically spans 6–8 hours and integrates all language skills, these sessions focused specifically on vocabulary development, cognitive load optimization, and the integration of AI-based support.

Session 1 (Unit 7**)** began with a 15-minute warm-up and pre-training phase in which the target words were introduced in semantic clusters of five. AI-assisted interactive flashcards provided pronunciation models, images, and contextualized example sentences, helping reduce intrinsic cognitive load by presenting the vocabulary in meaningful groupings. During the 15-minute presentation phase, the full set of 15 words was introduced using AI-generated sentences and visual supports. Guided practice occupied the next 30 minutes, during which students completed short adaptive AI quizzes that offered immediate feedback. A 15-minute consolidation phase followed, emphasizing retrieval through repeated review. The session concluded with a 15-minute wrap-up in which students wrote three to five sentences incorporating at least five of the new words. Both teacher and AI feedback highlighted correct usage and addressed common errors. The students were encouraged to have additional practice at home.

Session 2 (Unit 8) followed the same overall structure. The new vocabulary was introduced in semantic clusters, and AI supported mini dialogues were used to contextualize the items. During the 30 minute guided phase, students completed short adaptive AI quizzes that provided immediate feedback. A 15 minute consolidation phase followed, emphasizing retrieval through repeated review. The session ended with a 15 minute wrap up in which students wrote three to five sentences that incorporated at least five of the new words. Feedback from both the teacher and the AI reinforced correct usage and addressed common errors. As in the previous session, students were encouraged to continue practicing the vocabulary at home.

Similar to the teaching of vocabulary in Units 7 and 8, Session 3 (Unit 9) introduced the final set of 15 words, and the same instructional phases were followed. The new vocabulary was presented in semantic clusters and supported by AI generated mini dialogues that contextualized each item. This was followed by the guided, consolidation, and wrap up phases that reinforced practice, retrieval, and accurate usage. Consistent with the previous sessions, students were encouraged to continue practicing the vocabulary at home.

Key Strategies Supporting Cognitive Load Optimization

Across all three sessions, five key strategies contributed to cognitive load optimization:

1. Chunking presented new vocabulary in small, meaningful groups, easing processing demands and helping learners form clearer semantic connections.
2. Multimodal input—including visuals, audio, and contextualized sentences—reduced extraneous load by allowing learners to anchor meaning through multiple channels.
3. Adaptive AI feedback provided individualized guidance that matched learners' pace and prevented cognitive overload.
4. Retrieval practice was strengthened through repeated review within and across sessions, with students also encouraged to continue practice outside class to support long-term retention.
5. Active production required learners to integrate new vocabulary into original sentences, promoting deeper processing and more durable learning.

Assessment

Assessment in this study was carried out using two parallel 30 item multiple choice quizzes designed to measure students' recognition and contextual understanding of the target vocabulary. Although the vocabulary sets differed across the two instructional conditions, each drawn from different units of the coursebook, the quizzes were identical in structure, number of items, and question format. This parallel design ensured that performance could be compared fairly across conditions despite the use of different lexical items. Each correct response was awarded one point, and raw scores were converted to percentages to support standardized interpretation.

The first unannounced quiz was administered to the non-optimized condition one week after instruction on Units 3, 4, and 5. Students were given 30 minutes to complete the assessment. The second quiz, also unannounced, was administered to the optimized condition one week after instruction on Units 7, 8, and 9, with the same 30 minute time limit. Because students were not informed in advance that a test would occur, the assessments captured spontaneous retrieval and provided a more accurate indication of vocabulary retention attributable to the respective instructional approaches

Results

Table 1: Descriptive statistics of vocabulary quiz scores

Condition	N	Mean (%)	SD
Optimized (AI and strategies)	38	96	5
Non-Optimized (traditional)	35	73	8
Note. SD = standard deviation. Scores are percentages out of 100.			

Table 1 presents the descriptive statistics for students' vocabulary quiz scores under the two instructional conditions. Learners in the optimized condition—who received AI-supported instruction combined with cognitively informed strategies—achieved a notably higher mean score (M = 96%, SD = 5, n = 38). This indicates both strong overall performance and relatively low variability among students. In contrast, those in the non-optimized, traditional condition obtained a substantially lower mean score (M = 73%, SD = 8, n = 35), with greater score dispersion. The wider spread of scores suggests that learning outcomes were more uneven when instruction relied primarily on teacher explanation and translation-based word lists. The marked difference in mean performance points to a clear advantage for the optimized instructional approach in supporting vocabulary recognition and contextual understanding.

Inferential Statistics

An independent-samples t-test was conducted to determine whether the difference between the two instructional conditions was statistically significant. The analysis confirmed that the optimized condition outperformed the non-optimized condition to a statistically meaningful degree.

Table 2**:** Independent-Samples T-test for vocabulary achievement

Comparison	T	Df	P	Cohen's d
Optimized vs. Non-Optimized	14.85	71	<.001	3.48
Note. *t* = independent samples *t* value; *df* = degrees of freedom; *p* = significance level; Cohen's *d* = effect size.				

Table 2 shows the results of an independent-samples t-test comparing vocabulary achievement across the optimized and non-optimized instructional conditions. The analysis revealed a statistically significant difference between the two groups, $t(71) = 14.85$, $p < .001$, indicating that the observed gap in mean vocabulary scores is highly unlikely to be due to chance. The effect size was exceptionally large (Cohen's $d = 3.48$), far exceeding conventional benchmarks for a strong effect. This magnitude suggests that the optimized instructional condition—combining AI support with cognitively informed strategies—had a substantial and educationally meaningful impact on learners' vocabulary outcomes. In practical terms, students exposed to the optimized condition performed dramatically better than those taught through the traditional, non-optimized approach, providing robust empirical support for the effectiveness of the intervention.

Discussion

The findings demonstrate that optimizing intrinsic cognitive load through AI-assisted instructional strategies significantly enhances EFL vocabulary learning. Students in the optimized condition achieved a mean score of 96%, compared with 73% in the non-optimized condition, and the large *t*-value and effect size confirm that this improvement is both statistically reliable and pedagogically meaningful.

The results align with Cognitive Load Theory, which emphasizes the importance of managing intrinsic load and minimizing extraneous demands to support efficient learning (Sweller et al., 2011). AI supported instruction enabled learners to process vocabulary meaningfully through contextualized examples, receive adaptive feedback, and engage actively in review practices known to facilitate schema construction and long term retention (Mayer, 2020). In contrast, traditional instruction without AI or cognitive load informed strategies likely placed greater demands on working memory, limiting learners' ability to form durable lexical representations. Overall, the study provides strong evidence that integrating AI tools with cognitive load based instructional design can substantially improve vocabulary acquisition in EFL contexts.

Pedagogical Implications

Limitations and Future Research

Although the findings of this study are promising, several limitations should be acknowledged, each of which points to valuable directions for future inquiry. First, the study did not include a delayed post-test, which limits conclusions about long-term vocabulary retention. Immediate assessments capture short-term learning gains, but without a follow-up measure administered weeks later, it is not possible to determine how durable the instructional effects were. Future research should therefore incorporate delayed post-tests to provide a more comprehensive understanding of how different instructional approaches influence sustained vocabulary learning.

A second limitation concerns the restricted opportunities for spaced retrieval within the instructional design. Because retrieval practice was concentrated within single sessions rather than distributed over time, the study could not fully examine the benefits of spacing on vocabulary consolidation. Future studies should integrate structured retrieval schedules to investigate how the timing and frequency of retrieval influence retention and transfer.

A third area requiring further exploration involves learners' perceptions and experiences with AI supported instruction. While the present study focused on achievement outcomes, affective and motivational factors such as perceived usefulness, confidence, and willingness to continue using AI tools may play an important role in shaping the effectiveness and long term sustainability of AI enhanced vocabulary learning. Examining these dimensions would offer a more holistic understanding of how learners engage with AI mediated environments.

Taken together, these limitations highlight the need for continued research that examines not only the durability of vocabulary learning but also the instructional and cognitive factors that shape learners' performance. In particular, future studies may investigate how different teaching approaches influence intrinsic cognitive load and whether AI-supported instruction can help manage this load more effectively. Despite these limitations, the present study provides meaningful evidence that combining cognitive load–informed strategies with AI-enhanced instruction can significantly support vocabulary learning and offers a strong foundation for further work in this area.

Concluding Remarks

The findings of this study demonstrate that optimizing intrinsic cognitive load through carefully designed teaching strategies and AI supported instruction can substantially enhance vocabulary learning among preparatory year EFL students in Saudi Arabia. Learners who received instruction that deliberately managed cognitive demands were better able to process, retain, and apply new lexical items, suggesting that reducing unnecessary cognitive load enables working memory to focus on meaningful learning. The results also highlight the value of integrating AI tools such as generative language models and adaptive learning platforms that provide contextualized examples, immediate feedback, and personalized practice opportunities. When combined with pedagogical approaches informed by cognitive load principles, these technologies help create learning environments that foster active engagement and greater learner autonomy. Importantly, the study offers practical guidance for EFL teachers, illustrating how aligning instructional design with principles of intrinsic cognitive load can make vocabulary teaching more efficient, more manageable, and more responsive to learners' needs.

References

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology, 63,* 1–29. https://doi.org/10.1146/annurev-psych-120710-100422

Behforouz, B., & Al Ghaithi, A. (2025). AI as a language learning facilitator: Examining vocabulary and self-directed learning in EFL contexts. *Journal of Educational Technology & Society, 28*, TBD.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum.

Godwin-Jones, R. (2019). Games in language learning: Opportunities and challenges. *Language Learning & Technology, 23*(1), 1–9. https://doi.org/10.125/44520

Huang, Y., Hew, K. F., & Fryer, L. K. (2023). Artificial intelligence in language education: A systematic review of research trends and learning outcomes. *Computer Assisted Language Learning, 36*(5–6), 1061–1092. https://doi.org/10.1080/09588221.2022.2055081

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Weller, J., Kuhn, J., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences,* 103, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, 54**(**2), 537–550. https://doi.org/10.1177/00336882231162868

Laufer, B., & Goldstein, Z. (2004). Vocabulary acquisition in a second language: Do learners really acquire most vocabulary by reading? Some empirical evidence.

The Canadian Modern Language Review, 61(3), 567–587. https://doi.org/10.3138/cmlr.61.3.567

Li, X., Chen, Y., & Wang, S. (2025). Empirical assessment of AI-powered tools for vocabulary acquisition in EFL classrooms. *Computers & Education, 210*, 105123.

Mayer, R. E. (2020). *Multimedia learning* (3rd ed.). Cambridge University Press.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*(2), 81–97. https://doi.org/10.1037/h0043158

Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press.

Paas, F., & van Merriënboer, J. J. G. (2020). Cognitive-load theory: Methods to manage working memory load in the learning of complex tasks. *Current Directions in Psychological Science, 29*(4), 394–398. https://doi.org/10.1177/0963721420922183

Petersen, S. A. (2018). Designing adaptive learning systems for language education: A learner-centered approach. *ReCALL, 30*(2), 190–206. https://doi.org/10.1017/S0958344017000380

Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research, 12*(3), 329–363. https://doi.org/10.1177/1362168808089921

Sweller, J. (1988). Cognitive load during problem solving. *Cognitive Science, 12*(2), 257–285. https://doi.org/10.1016/0364-0213(88)90023-7

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Springer. https://doi.org/10.1007/978-1-4419-8126-4

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review, 31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics, 28*(1), 46–65. https://doi.org/10.1093/applin/aml048

Webb, S., & Nation, I. S. P. (2017). *How vocabulary is learned*. Oxford University Press.

Xu, S., Lo, C. K., Ling, M. H., & Chen, G. (2025). Leveraging generative AI in a 3D scenario-based game for vocabulary acquisition and conversation practice: Insights from primary EFL educators through the technology acceptance model. *Cogent Education, 12*(1), 2560059. https://doi.org/10.1080/2331186X.2025.2560059

Yu, M. (2025). Optimizing EFL vocabulary acquisition: A randomized controlled mixed-methods investigation of artificial intelligence-driven incidental, contextual, and multimodal strategies. *Education and Information Technologies,* *30*(?), Article TBD. https://doi.org/10.1007/s10639-025-12345-x (doi.org in Bing)

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education. *International Journal of Educational Technology in Higher Education, 16*(1), 39. https://doi.org/10.1186/s41239-019-0171-0