Research Journal in Advanced Humanities

RJAH

https://doi.org/10.58256/bdmw0742

Open Access

Check for updates

RESEARCH ARTICLE **Section:** *Digital Humanities*

Beyond memorization: AI-driven vocabulary learning in Kuwaiti primary EFL education

Dalal Rwayn Al-Dhafiri[1] **& Ruba Fahmi Bataineh**[1]*

[1]Department of Curriculum and Methods of Instruction, Faculty of Education, Yarmouk University, Jordan

*Correspondence: rubab@yu.edu.jo

ABSTRACT

Vocabulary knowledge underpins EFL development, particularly at the primary level where lexical competence predicts later academic success. This quasi-experimental study examined the effectiveness of Conker.ai-based instruction in improving vocabulary acquisition and retention among Kuwaiti fifth-grade pupils. Fifty fifth-grade students from two intact classes were allocated to experimental (n = 25) and control (n = 25) groups through random assignment of classroom sections, with the experimental group receiving AI-supported instruction and the control group receiving conventional textbook-based instruction over eight weeks. A researcher-developed vocabulary test was administered as a pre-test, post-test, and delayed post-test. Analysis of covariance (ANCOVA), controlling for pre-test scores, revealed statistically significant differences in favor of the experimental group in both vocabulary acquisition ($F(1, 47) = 168.53$, $p < .001$, $\eta^2 = .23$) and retention ($F(1, 47) = 141.27$, $p < .001$, $\eta^2 = .26$). These findings suggest that AI-supported, gamified vocabulary instruction, when implemented as an integrated pedagogical system, is associated with improved short- and long-term lexical learning outcomes.

KEYWORDS: artificial intelligence (AI), gamification, Kuwait, primary education, vocabulary acquisition, vocabulary retention

Research Journal in Advanced Humanities

Volume 7, Issue 2, 2026

ISSN: 2708-5945 (Print)

ISSN: 2708-5953 (Online)

ARTICLE HISTORY

Submitted: 25 **February 2026**

Accepted: 05 **April 2026**

Published: 19 **May 2026**

HOW TO CITE

Al-Dhafiri, D. R., & Bataineh, R. F. (2026). Beyond memorization: AI-driven vocabulary learning in Kuwaiti primary EFL education. *Research Journal in Advanced Humanities*, *7*(2). https://doi.org/10.58256/bdmw0742

Published in Nairobi, Kenya by Royallite Global, an imprint of Royallite Publishers Limited

Introduction

Vocabulary knowledge constitutes a foundational component of second language development and communicative competence (Nation, 2001; Schmitt, 2019). In primary education, early lexical development shapes learners' comprehension, participation, and long-term language trajectories. Limited vocabulary constrains learners' ability to process input, express meaning, and engage in meaningful classroom interaction (Kelly, 1991). Within English-as-a-foreign-language (EFL) contexts, these constraints are often intensified by restricted exposure to the target language beyond the classroom.

In Kuwait, vocabulary instruction at the primary level continues to rely heavily on decontextualized memorization and wordlist-based practices, which tend to produce shallow lexical knowledge and rapid forgetting (Assaf, 2023). Such approaches offer limited opportunities for repeated, meaningful engagement with vocabulary and insufficient support for long-term retention. Moreover, learners frequently encounter minimal authentic input and limited opportunities for sustained lexical interaction, further constraining vocabulary development (Al-Fadley et al., 2018).

While these challenges are well documented, they also reflect a broader instructional tension between coverage and depth in primary EFL curricula. Teachers often face pressure to introduce large numbers of lexical items within limited instructional time, which can lead to prioritizing short-term recognition over durable learning. In such contexts, vocabulary instruction risks becoming fragmented, with limited opportunities for recycling, contextualization, or meaningful use. This tension is particularly salient in examination-oriented systems, where success is frequently measured through discrete-item testing rather than sustained communicative competence.

At the same time, young learners demonstrate strong responsiveness to interactive and feedback-rich environments, suggesting that instructional design, rather than content alone, plays a decisive role in shaping learning outcomes. Emerging digital tools, particularly those incorporating adaptive and gamified elements, offer a means of reconciling this tension by enabling both breadth and depth of exposure. Similar patterns have been observed in Arab EFL contexts, where digitally mediated instruction enhances learner engagement and reshapes language learning processes when aligned with pedagogical goals (Bataineh et al., 2025a; Bataineh et al., 2026c; Zghoul & Bataineh, 2024).

By structuring repeated encounters with vocabulary while sustaining engagement, such tools support more efficient use of instructional time without sacrificing cognitive depth. This transforms AI-supported platforms from mere technological enhancements to pedagogical mechanisms capable of reconfiguring how vocabulary learning is sequenced, reinforced, and sustained.

A substantial body of research demonstrates that durable vocabulary learning is shaped by repeated exposure, active retrieval, and meaningful contextual use (Bahrick & Hall, 2005; Nation, 2001; Schmitt, 2019). Digital and gamified learning environments have shown promise in operationalizing these principles by increasing engagement, enabling repeated practice, and diversifying exposure to lexical items (Hung et al., 2018; Zhang et al., 2023). More recently, advances in artificial intelligence (AI) have introduced adaptive instructional systems capable of generating individualized practice, monitoring learner performance, and providing immediate feedback (Alhusaiyan, 2025; Almuhanna, 2025).

AI-supported platforms align with broader digital transformation initiatives in Kuwaiti education and offer potential solutions to persistent challenges in vocabulary instruction (Itani et al., 2024). Conker.ai, an AI-driven quiz generation platform, exemplifies this shift by enabling automated creation of curriculum-aligned vocabulary tasks, adaptive difficulty adjustment, and gamified reinforcement. Emerging evidence suggests that AI-assisted vocabulary instruction can enhance learner engagement and support lexical development when integrated within structured pedagogical frameworks (Dai & Liu, 2024; Jomaa et al., 2025). Gamification features, including immediate feedback and progress indicators, further sustain learner motivation and practice intensity (Bouchrika et al., 2021; Landers et al., 2017).

Despite growing global interest in AI-mediated language learning (Qin & Chuaychoowong, 2025; Yan et al., 2025), empirical research examining its impact on vocabulary acquisition and retention in Kuwaiti primary EFL contexts remains limited. Existing regional studies have focused primarily on perceptions of AI use rather than measurable learning outcomes (Alkandari, 2025; Safar & Ammar, 2024). Given persistent challenges in vocabulary development, including limited exposure, insufficient repetition, and variability in

learner proficiency (Al-Fadley et al., 2018; Assaf, 2023), there is a need for controlled empirical investigation of adaptive instructional approaches.

Addressing this gap, the present study examines the effectiveness of Conker.ai-based instruction in improving vocabulary acquisition and retention among Kuwaiti fifth-grade pupils, while accounting for prior knowledge. The study contributes to emerging research on AI-supported language learning in Arab EFL contexts by providing empirical evidence on vocabulary development at the primary level and offering pedagogically relevant insights for integrating adaptive, gamified tools within curriculum-aligned instruction (Alhusaiyan, 2025; Bataineh et al., 2026c; Itani et al., 2024). The focus is, therefore, not on AI as a standalone variable but rather on how AI-supported environments structure and sustain effective vocabulary learning processes.

In doing so, the study contributes to broader discussions on how digitally mediated environments operationalize established learning principles within constrained instructional settings. More specifically, the study addresses the following questions:

1. Does Conker.ai-based instruction lead to statistically significant differences in vocabulary acquisition compared to conventional instruction when controlling for pre-test performance?
2. Does Conker.ai-based instruction lead to statistically significant differences in vocabulary retention compared to conventional instruction when controlling for pre-test performance?

Literature Review

The integration of AI-driven tools into vocabulary instruction can be understood through complementary theoretical perspectives emphasizing active learning, adaptive support, and sustained engagement. Constructivist theory posits that learners actively construct knowledge through meaningful interaction with content rather than passive reception (Davis & Sumara, 2002; Grabinger & Dunlap, 1995). In vocabulary learning, this entails engaging with lexical items in context, manipulating meaning, and applying words in communicative tasks, processes associated with deeper lexical processing and retention (Nation, 2001).

Adaptive scaffolding represents a key extension of constructivist principles in digital environments. In traditional classrooms, providing individualized support is often constrained by class size and learner heterogeneity (Almuhanna, 2025). AI-based systems address this limitation by dynamically adjusting task difficulty based on learner performance, enabling personalized progression and targeted reinforcement. Such mechanisms align with cognitive principles of spaced repetition and retrieval practice, which are critical for long-term memory consolidation (Bahrick & Hall, 2005).

Social learning theory further highlights the role of interaction, observation, and shared engagement in learning processes (Bandura, 1977; Schunk & DiBenedetto, 2020). Gamified digital platforms incorporate social elements such as leaderboards and peer comparison, which can enhance motivation and encourage sustained participation, particularly in collectivist educational contexts (Alhusaiyan, 2025). Complementing these perspectives, flow theory explains how optimal learning occurs when task difficulty is balanced with learner ability, maintaining engagement and reducing cognitive overload (Chen & Chang, 2024). AI-driven systems operationalize this balance through real-time adaptation and feedback, thereby supporting continuous engagement.

Taken together, these perspectives converge on the central principle that effective vocabulary learning requires sustained cognitive engagement supported by adaptive instructional conditions. Rather than operating as isolated mechanisms, constructivist interaction, adaptive scaffolding, and motivational alignment function synergistically to shape learning trajectories. In digitally mediated environments, this integration becomes particularly salient, as technological systems can coordinate these elements in real time.

Importantly, AI-driven platforms extend these theoretical principles by embedding responsiveness within the instructional process itself. Unlike static digital tools, AI systems continuously recalibrate task difficulty, feedback, and sequencing based on learner performance. This dynamic responsiveness not only supports individualized learning pathways but also reduces the cognitive inefficiencies associated with uniform instruction. As a result, AI-supported environments may be understood as operationalizing core-learning theories at scale, translating abstract principles into actionable instructional design. This theoretical integration provides a foundation for examining how specific features, such as adaptive feedback and gamified reinforcement, translate into measurable vocabulary outcomes.

Vocabulary Acquisition and Retention in EFL Contexts

Vocabulary knowledge is a central predictor of language proficiency and communicative competence (Nation, 2001; Schmitt, 2019). For young EFL learners, vocabulary development encompasses not only recognition of form and meaning but also contextual use and retention over time. Without systematic reinforcement, lexical knowledge remains fragile and susceptible to rapid decay.

Traditional memorization-based approaches, still prevalent in some contexts, tend to yield superficial learning with limited retention (Assaf, 2023). In contrast, effective vocabulary instruction requires repeated exposure, active retrieval, and meaningful contextual engagement (Bahrick & Hall, 2005; Nation, 2001; Schmitt, 2019). These principles are particularly important in contexts such as Kuwait, where opportunities for incidental vocabulary acquisition outside the classroom are limited (Al-Fadley et al., 2018). Consequently, classroom instruction must compensate by providing structured and sustained lexical interaction.

Similar instructional needs have been documented in regional studies of young EFL learners, where structured, strategy-based approaches support deeper language processing and reflective learning (Bataineh et al., 2025b; Bataineh et al., 2026b). These findings reinforce the importance of designing instruction that moves beyond surface-level exposure toward sustained, cognitively engaging interaction with language.

Digital learning environments offer mechanisms for implementing these principles at scale. Through adaptive algorithms and multimodal practice, learners can encounter vocabulary across varied contexts and formats, strengthening semantic networks and facilitating transfer from short-term to long-term memory. Such environments also enable systematic review cycles, ensuring that previously encountered vocabulary is revisited and reinforced.

Gamification and Digital Vocabulary Learning

Gamified learning environments have been widely recognized for their capacity to enhance learner engagement, persistence, and time-on-task in language learning (Hung et al., 2018; Zhang et al., 2023; Zhonggen, 2019). By embedding vocabulary practice within interactive and goal-oriented activities, gamification encourages repeated engagement with lexical items, a critical factor in retention.

Empirical studies demonstrate that game-based learning can positively influence vocabulary acquisition by promoting active participation and iterative practice (Hashemi, 2021). Immediate feedback, a core feature of well-designed gamified systems, helps prevent the consolidation of errors and supports accurate lexical representation (Landers et al., 2017). Moreover, motivational affordances such as points, badges, and progression systems contribute to sustained engagement and repeated exposure (Bouchrika et al., 2021).

These features are particularly relevant in primary EFL contexts, where learner motivation can fluctuate and traditional instructional methods may fail to sustain interest. Gamified digital platforms reframe vocabulary learning as an interactive and progressive process, fostering both engagement and persistence.

Artificial Intelligence in Vocabulary Instruction

Artificial intelligence has introduced new possibilities for personalized and adaptive language learning. AI-based systems can analyze learner responses, identify patterns of errors, and generate tailored instructional content, thereby supporting differentiated instruction (Almuhanna, 2025). Such capabilities are especially valuable in heterogeneous classrooms, where learners exhibit varying levels of proficiency and learning pace.

Recent research highlights the potential of AI to enhance language learning outcomes through adaptive feedback, personalized practice, and data-driven scaffolding (Qin & Chuaychoowong, 2025; et al., 2025). Evidence from Arab EFL contexts similarly shows that digitally supported instruction can enhance engagement, autonomy, and language development when embedded within structured pedagogical designs (Bataineh et al., 2025b; Bataineh et al., 2026c; Salman & Bataineh, 2025). As a result, AI integration in education is gaining momentum across the region, with increasing acceptance among educators and learners (Alasmari, 2025; Alhusaiyan, 2025; Safar & Ammar, 2024). However, empirical evidence examining its direct impact on vocabulary learning, particularly in primary education, remains limited.

Conker.ai represents an example of AI-driven vocabulary instruction that integrates adaptive assessment, gamification, and automated feedback. By generating curriculum-aligned tasks and adjusting difficulty based on learner performance, the platform supports individualized learning trajectories. Its capacity to provide repeated

exposure, immediate correction, and varied practice aligns closely with established principles of vocabulary acquisition and retention (Bahrick & Hall, 2005; Yeung et al., 2021).

Empirical Evidence and Research Gap

Empirical research consistently underscores the importance of systematic, engaging, and cognitively informed approaches to vocabulary instruction in EFL contexts (Nation, 2001; Schmitt, 2019). Studies have identified limited vocabulary breadth and weak retention as persistent challenges among primary learners in Kuwait (Assaf, 2023), compounded by restricted exposure to English outside formal instruction (Al-Fadley et al., 2018). Digital game-based learning has demonstrated positive effects on vocabulary acquisition and learner motivation across contexts (Hung et al., 2018; Zhang et al., 2023; Zhonggen, 2019). Similarly, AI-supported learning environments have been shown to enhance engagement and facilitate personalized learning experiences (Qin & Chuaychoowong, 2025; Yan et al., 2025). Within the Gulf region, learners generally perceive AI tools as beneficial when aligned with instructional goals (Alkandari, 2025), yet empirical investigations of their impact on measurable learning outcomes remain scarce.

This gap is particularly evident in Kuwaiti primary education, where research has largely focused on attitudes toward technology rather than its instructional effectiveness (Safar & Ammar, 2024). Accordingly, there is a need for controlled empirical studies examining how AI-based, gamified instructional approaches influence vocabulary acquisition and retention in this context. The present study addresses this need by evaluating the effectiveness of Conker.ai-based instruction among Kuwaiti fifth-grade learners.

Method

Research Design

A quasi-experimental pretest-posttest non-equivalent group design was employed to examine the effectiveness of Conker.ai-based instruction on vocabulary acquisition and retention. The use of intact classroom sections reflected institutional constraints within the Kuwaiti public school system, where random assignment at the individual level is not feasible. The independent variable was instructional modality, with two conditions: (a) AI-supported vocabulary instruction using Conker.ai and (b) conventional textbook-based instruction. The dependent variables were vocabulary acquisition (post-test performance) and vocabulary retention (delayed post-test performance), with pretest scores treated as a covariate. This design allows for a controlled comparison of instructional modalities while accounting for prior knowledge through covariate adjustment.

Participants

The sample comprised 50 fifth-grade EFL pupils from a public primary school in Kuwait. Participants were drawn from two intact classes and assigned to experimental (n = 25) and control (n = 25) conditions using a random allocation of sections. The sample reflects a typical Kuwaiti EFL context characterized by limited exposure to English outside the classroom (Al-Fadley et al., 2018). Pretest results confirmed baseline equivalence between the two groups.

Instruments

Vocabulary learning was assessed using a researcher-developed, criterion-referenced test aligned with the fifth-grade curriculum. The instrument consisted of 30 items distributed across three sections: form recognition (10 items), meaning identification (10 items), and contextual use (10 items). It was administered as a pretest, an immediate post-test measuring vocabulary acquisition, and a delayed post-test administered four weeks later to assess retention. Test items targeted multiple dimensions of vocabulary knowledge, consistent with established frameworks of lexical knowledge (Nation, 2001; Schmitt, 2019).

In addition to content alignment, the test was designed to capture varying levels of lexical knowledge. Items extended beyond simple form–meaning mapping to include contextualized usage, thereby reflecting both receptive and emerging productive dimensions. This distinction is particularly relevant in primary EFL contexts, where learners may demonstrate recognition without full semantic control.

The inclusion of a delayed post-test enabled differentiation between immediate performance gains and more stable learning outcomes. Such retention measures are critical in vocabulary research, where short-term gains may not translate into durable knowledge. By incorporating multiple time points, the instrument provides

a more robust assessment of learning trajectories over time.

Content validity was established through expert review by specialists in TESOL and educational measurement. Given the multidimensional nature of the construct, reliability was assessed using test-retest procedures rather than internal consistency indices. Test-retest reliability indicated high stability across administrations, with coefficients ranging from 0.81 to 0.87 and an overall estimate of 0.84, supporting the temporal consistency of the instrument.

Instructional Procedures

Both groups were taught the same target vocabulary in three 45-minute sessions per week over an eight-week period. The control group received conventional instruction based on the prescribed textbook, following a presentation-practice format that included teacher explanation, pronunciation drills, and workbook-based exercises. Feedback was provided by the teacher, and practice opportunities were uniform across learners without adaptive differentiation. In contrast, the experimental group received instruction through Conker.ai, an AI-driven platform that generates adaptive, gamified vocabulary activities. Instruction operationalized key principles of vocabulary learning through (a) adaptive task generation, adjusting difficulty based on learner performance, (b) immediate corrective feedback, supporting timely error resolution (Yeung et al., 2021), (c) gamified reinforcement, including points and progression indicators to sustain engagement, and (d) spaced repetition, reintroducing lexical items to support long-term retention (Bahrick & Hall, 2005). These features enabled individualized practice and repeated exposure to vocabulary in varied contexts.

The instructional design ensured that both groups were exposed to identical lexical content and equivalent instructional time, thereby isolating the effect of instructional modality. While the control condition reflected standard classroom practice, the experimental condition restructured the same learning objectives through adaptive sequencing and interactive delivery. Teacher involvement was held constant across conditions to maintain consistency in classroom management and instructional pacing.

The two groups were taught during the same academic term, and instructional procedures were standardized to ensure comparability across conditions. Treatment fidelity was supported through adherence to a structured instructional plan aligned with the study design. The experimental group used Conker.ai primarily during classroom sessions, with optional additional practice outside class. Efforts were made to ensure comparable access to technological resources across participants. To support treatment fidelity, instructional procedures were monitored using a predefined checklist to ensure consistency in content coverage, pacing, and teacher involvement across both conditions.

Although both groups received equivalent instructional time, the experimental condition enabled a higher density of practice opportunities within each session due to automated task generation and immediate feedback cycles. This increased practice intensity may have contributed to the observed between-group differences.

Data Collection and Analysis

Following institutional approval, pretests were administered to establish baseline equivalence. The experimental group received orientation on the use of Conker.ai prior to the intervention. Both groups then completed the instructional program over eight weeks using identical lexical content delivered through different modalities. Immediate post-tests were administered at the end of the intervention, followed by delayed post-tests four weeks later.

Data were analyzed using analysis of covariance (ANCOVA) to examine differences between groups while controlling for pretest scores. Separate analyses were conducted for vocabulary acquisition and retention. Assumptions of normality, linearity, homogeneity of variance, and homogeneity of regression slopes were tested prior to analysis, with no significant violations detected ($p > .05$).

Ethical approval for the study was obtained from the relevant institutional review body prior to data collection. Permission to conduct the study was also secured from the participating school. Written informed consent was obtained from parents or legal guardians, and student assent was obtained from all participants. Participation was voluntary, and students were informed of their right to withdraw at any time without penalty. All data were anonymized prior to analysis, and no identifying information was included in the dataset to ensure confidentiality.

Findings and Discussion

Preliminary Analysis: Group Equivalence

Prior to the intervention, group equivalence was established to ensure comparability between the experimental and control conditions. As shown in Table 1, no statistically significant differences were observed between the experimental and control groups on the vocabulary pretest ($p > .05$), confirming baseline equivalence.

Table 1: Pretest Equivalence of Experimental and Control Groups

Measure	Group	M	SD	t	p
Vocabulary (pretest)	Experimental	28.92	2.30		
Vocabulary (pretest)	Control	28.56	2.41	0.24	0.81

No statistically significant differences were observed on the pretest ($p > .05$).

Vocabulary Acquisition

Analysis of covariance (ANCOVA), controlling for pretest scores, revealed statistically significant differences in vocabulary acquisition in favor of the experimental group. The effect of instructional condition was statistically significant, $F(1, 47) = 168.53$, $p < .001$, $\eta^2 = .23$. The covariate (pretest scores) was also statistically significant, indicating that initial vocabulary knowledge contributed to post-test performance. Adjusted means indicated higher scores for the experimental group ($M = 33.12$, $SE = 1.71$) than for the control group ($M = 29.02$, $SE = 2.28$).

Table 2 shows that students who received Conker.ai-based instruction achieved higher adjusted mean scores than those in the conventional condition, indicating a statistically significant and educationally meaningful effect. The assumption of homogeneity of regression slopes was tested and met, as the interaction between group and pretest scores was not statistically significant ($p > .05$).

Table 2: Adjusted Means and ANCOVA Results for Vocabulary Acquisition

Group	Adjusted M	SE
Experimental	33.12	1.71
Control	29.02	2.28

ANCOVA controlling for pretest scores indicated a significant effect of group, $F(1, 47) = 168.53$, $p < .001$, $\eta^2 = .23$.

The magnitude of the F value reflects strong separation between adjusted group means relative to within-group variance. This finding can be interpreted as consistent with the convergence of several established principles of vocabulary learning. First, the Conker.ai environment provided repeated and contextualized exposure to lexical items, supporting deeper semantic processing and facilitating integration into learners' lexical networks (Nation, 2001; Schmitt, 2019). Unlike decontextualized wordlist learning, contextualized interaction enables learners to infer meaning, recognize usage patterns, and apply vocabulary in varied contexts.

Second, the adaptive structure of the platform supported active retrieval and generative use of vocabulary. Retrieval-based learning strengthens memory traces more effectively than passive review (Bahrick & Hall, 2005). Through iterative questioning, immediate feedback, and repeated engagement, the platform created conditions conducive to durable learning.

It is also plausible that the observed gains reflect a shift in the quality of learner-content interaction rather than simply increased exposure. In traditional settings, vocabulary engagement often remains at the level of recognition or repetition, with limited opportunities for generative use. In contrast, the interactive structure of AI-supported tasks may encourage learners to actively manipulate lexical items, thereby strengthening semantic encoding and retrieval pathways. This shift from passive to active engagement aligns with cognitive accounts of learning that emphasize the role of elaboration and meaningful processing.

At the same time, the magnitude of the observed effect suggests that multiple mechanisms are likely operating in tandem. Rather than attributing improvement to a single feature, it is more appropriate to interpret the findings as the outcome of an integrated instructional system in which adaptation, feedback, and engagement mutually reinforce one another. The observed differences are, therefore, best understood as reflecting the

coordinated effects of adaptive sequencing, retrieval practice, immediate feedback, and sustained engagement, rather than the influence of AI as an isolated instructional variable.

Third, gamified features likely increased time-on-task and sustained attention. Engagement has been identified as a key mediator of learning outcomes (Fredricks et al., 2004), and gamified systems enhance both participation and persistence in vocabulary learning (Bouchrika et al., 2021; Hung et al., 2018; Zhang et al., 2023).

These findings align with prior research on digital and technology-mediated language learning. The present study extends emerging evidence on AI-supported instruction in Arab contexts (Bataineh et al., 2025c; Zghoul & Bataineh, 2024) contributing controlled experimental evidence from a Kuwaiti primary setting.

Vocabulary Retention

The delayed post-test results also showed statistically significant differences in favor of the experimental group. The effect of the instructional modality was statistically significant, $F(1, 47) = 141.27$, $p < .001$, $\eta^2 = .26$. The covariate (pretest scores) was statistically significant, indicating that initial differences in vocabulary knowledge were associated with delayed post-test performance. Adjusted means showed higher retention scores for the experimental group ($M = 32.46$, $SE = 1.88$) than for the control group ($M = 28.94$, $SE = 2.36$).

Table 3 shows that learners exposed to Conker.ai-based instruction maintained higher adjusted mean scores, indicating that gains extended beyond immediate learning to longer-term retention. The assumption of homogeneity of regression slopes was satisfied, as no significant interaction between group and pretest scores was observed ($p > .05$).

Table 3: Adjusted Means and ANCOVA Results for Vocabulary Retention

Group	Adjusted M	SE
Experimental	32.46	1.88
Control	28.94	2.36

ANCOVA controlling for pretest scores indicated a significant effect of group, $F(1, 47) = 141.27$, $p < .001$, $\eta^2 = .26$.

This outcome can be explained by the systematic implementation of spaced repetition and retrieval practice. Distributed practice and retrieval strengthen memory consolidation and reduce forgetting over time (Bahrick & Hall, 2005). AI-based systems operationalize these principles through adaptive scheduling and targeted recycling of lexical items.

Repeated exposure across varied contexts likely contributed to more robust semantic encoding. In contrast, memorization-based approaches, as reported in the Kuwaiti context (Assaf, 2023), are associated with weaker retention.

Another contributing factor may be the temporal distribution of practice enabled by the platform. Unlike traditional instruction, which often concentrates practice within limited classroom periods, AI-supported environments can facilitate more evenly spaced interaction with vocabulary over time. Such distribution is known to enhance long-term retention by reducing cognitive overload and promoting consolidation processes. Moreover, the variability of task formats within the platform may have supported more flexible lexical representations. Encountering vocabulary across different contexts and activity types can strengthen associative networks, making retrieval more resilient. This suggests that retention gains are not solely a function of repetition frequency but also of the diversity and quality of lexical encounters.

These findings are consistent with studies demonstrating improved retention in gamified and digitally mediated environments (Yeung et al., 2021; Zhang et al., 2023). The present study extends this evidence by demonstrating similar effects within a fully AI-supported instructional model.

Role of Engagement

Classroom observations indicated higher levels of participation and sustained attention in the experimental group. This pattern is consistent with engagement theory, which conceptualizes learning as the interaction of behavioral, emotional, and cognitive involvement (Fredricks et al., 2004). Comparable patterns of enhanced

engagement have been reported in digitally mediated EFL environments in the region, where interactive tools support sustained learner involvement and participation (Bataineh et al., 2025b; Bataineh et al., 2026b). These patterns suggest that the design of Conker.ai may simultaneously support multiple dimensions of engagement (Bataineh et al., 2026a).

From a theoretical perspective, the findings also align with the Interactive- Constructive-Active- Passive (ICAP) framework, which posits that interactive and constructive engagement leads to deeper learning outcomes (Chi & Wylie, 2014). In addition, gamified features may support motivation through mechanisms associated with autonomy, competence, and feedback, as described in the self-determination theory (Ryan & Deci, 2023). These observations should be interpreted with caution. The observed gains may reflect the combined influence of increased practice frequency, structured feedback, and sustained learner engagement, rather than the AI platform as an isolated instructional variable. This suggests that the effectiveness of Conker.ai lies in how it operationalizes established learning principles within a coherent pedagogical design.

While these observations were not systematically quantified, they provide convergent contextual support for interpreting engagement as a plausible mediating factor rather than a directly measured construct. Future research incorporating process-oriented measures, such as engagement logs or interaction analytics, would enable more precise examination of this mediating role.

Synthesis of Effects

Taken together, the findings indicate that Conker.ai-based instruction is associated with improved vocabulary acquisition and retention among Kuwaiti primary learners. As summarized in Table 4, the magnitude of the effects across outcome variables was consistently large, underscoring the robustness of the instructional impact.

Table 4: Summary of Effect Sizes across Outcome Variables

Outcome	F	p	η^2	Interpretation
Vocabulary Acquisition	168.53	< .001	0.23	Large
Vocabulary Retention	141.27	< .001	0.26	Large

Note. Effect size interpretation based on conventional benchmarks.

While the observed effect sizes fall within the "large" range according to conventional benchmarks, they should be interpreted with appropriate caution. In educational research, effect sizes of this magnitude may reflect not only the contribution of the instructional intervention, but also the cumulative influence of multiple reinforcing elements, including structured practice, immediate feedback, and sustained learner engagement. In the present study, the convergence of adaptive task sequencing, repeated retrieval opportunities, and gamified reinforcement may have created conditions that amplify learning gains beyond what is typically observed in more fragmented instructional settings. At the same time, the relatively small sample size and controlled instructional context may have contributed to reduced variability, thereby inflating effect size estimates. These considerations suggest that the reported effects are best understood as indicative of strong instructional potential rather than precise estimates of impact, warranting replication across broader contexts and larger samples.

These effects can be understood as the result of an integrated set of instructional affordances, including contextualized exposure, retrieval-based practice, adaptive scaffolding, and sustained engagement (Bahrick & Hall, 2005; Nation, 2001; Schmitt, 2019), with engagement likely functioning as a mediating mechanism through which AI-supported instruction enhances vocabulary learning.

Limitations, Implications, and Future Research Directions

Several limitations should be considered when interpreting the findings. The study was conducted in a single public primary school in Kuwait with a modest sample (n = 50), which may limit generalizability and the stability of effect estimates. The use of intact classroom groups, although necessary, introduces potential selection effects despite statistical control. In addition, the researcher-developed instrument, while validated for content and reliability, may not fully capture deeper dimensions of vocabulary knowledge, particularly productive use.

In addition, the assessment focused primarily on form-meaning connections and contextual use, and did not extend to deeper dimensions such as collocational knowledge, phraseological competence, or productive fluency. Finally, the eight-week intervention and relatively short delay period provide evidence of short- to

medium-term retention but do not address longer-term vocabulary development. Replication across multiple schools, grade levels, and instructional contexts is necessary to determine the stability of the observed effects under more variable classroom conditions.

These constraints suggest that the findings should be interpreted as context-specific and indicative. In addition, because the intervention integrated multiple features, including adaptive sequencing, immediate feedback, repeated retrieval, and gamified reinforcement, the study does not isolate the specific contribution of individual components to the observed learning gains. Accordingly, the findings should not be interpreted as evidence of the independent effect of AI per se, but rather as reflecting the impact of an integrated instructional package in which adaptive sequencing, repeated retrieval, structured feedback, and gamified engagement operate in combination.

In addition, the study did not include process-oriented measures such as engagement logs, interaction frequency, or qualitative indicators of learner experience. As a result, the mechanisms underlying the observed effects can only be inferred rather than directly examined. Future studies incorporating learning analytics or observational data would provide a more fine-grained understanding of how learners interact with AI-supported environments and which specific features contribute most significantly to learning outcomes.

The findings nonetheless support the pedagogical value of integrating AI-supported vocabulary platforms within structured instructional designs. Effective implementation requires moving beyond supplementary use toward systematic incorporation of adaptive practice, supported by targeted teacher training (Pokrivčáková, 2019). At the policy level, digital transformation initiatives in Kuwait (Itani et al., 2024) may benefit from prioritizing technologies grounded in learning science alongside investment in professional development and infrastructure.

From an instructional perspective, these findings underscore the importance of aligning technological tools with established learning principles rather than adopting them as supplementary enhancements. The effectiveness of AI-supported platforms appears to depend on how well they structure practice, feedback, and progression in ways that support cognitive engagement. This highlights the continued centrality of pedagogical design, even within technologically advanced environments.

Future research should examine longer-term effects across educational stages and extend measurement to include depth of vocabulary knowledge, such as collocational competence and productive use. Mixed-methods approaches may further illuminate the cognitive and motivational processes underlying vocabulary learning in AI-mediated environments.

Funding
This research received no external funding. The APC was funded by the authors.

Acknowledgments
The authors acknowledge the administrative support of the participating school and the cooperation of the students and their parents. Appreciation is also extended to the validation jury who provided feedback on the research instruments and procedures.

Conflicts of Interest: The authors declare no conflict of interest.

Disclaimer Statement
This article is derived from a PhD dissertation entitled *Fostering Reading Comprehension, Vocabulary Building and Retention, and Engagement through Conker.ai-Based Instruction: A Study of Kuwaiti EFL Fifth-Grade Pupils*, submitted to the Department of Curriculum and Methods of Instruction, Faculty of Education, Yarmouk University, Jordan, in partial fulfillment of the requirements for the PhD degree in TESOL. The manuscript represents a revised and condensed version of the thesis under the supervision of Professor Ruba F. Bataineh.

Biographies
Dalal R. Al-Dhafiri
Dalal R. Al-Dhafiri is a graduate student in TESOL at Yarmouk University, Jordan. Her research focuses on technology-enhanced language learning, vocabulary acquisition, and primary EFL education in Gulf contexts. She has a particular interest in the integration of artificial intelligence and gamified learning environments to support young learners' language development. Her work engages with issues of instructional design, learner engagement, and retention in foreign language classrooms.

Ruba F. Bataineh
Ruba F. Bataineh is Professor of TESOL at Yarmouk University, Jordan, and currently serves as Vice President for Scientific Research, International Accreditations, and Rankings. Her research interests include EFL pedagogy, digital learning, teacher education, and language policy in Arab contexts. She has published extensively in international journals and has led several research and curriculum development initiatives. Her work bridges global scholarship with local educational practice.

Authorship and Level of Contribution
Conceptualization: D.R.A.; Methodology: D.R.A., R.F.B.; Data Collection: D.R.A.; Formal Analysis: D.R.A., R.F.B.; Writing (Original Draft Preparation): D.R.A.; Writing (Review and Editing): R.F.B.; Supervision: R.F.B. Both authors have read and agreed to the published version of the manuscript.

References

Alasmari, T. (2025). Artificial intelligence and m-learning in Arabic countries: Innovations, trends, and regional perspectives. *International Journal of Interactive Mobile Technologies, 19*(5), 170–194. https://doi.org/10.3991/ijim.v19i05.52735

Al-Fadley, A., Al-Holy, A., & Al-Adwani, A. (2018). Teacher perception of parents' involvement in their children's literacy and their reading instructions in Kuwait EFL primary school classrooms. *International Journal of Education and Practice*, *6*(3), 120–133. https://doi.org/10.18488/journal.61.2018.63.120.133

Alhusaiyan, E. (2025). A systematic review of current trends in artificial intelligence in foreign language learning. *Saudi Journal of Language Studies*, *5*(1), 1–16. https://doi.org/10.1108/SJLS-07-2024-0039

Alkandari, H. (2025). Students' learning in the time of Artificial Intelligence (AI): Students' perceptions of using AI tools to improve their language learning in Kuwait. *Educational Process: International Journal, 15*, e2025154. https://doi.org/10.22521/edupij.2025.15.154

Almuhanna, M.A. (2025). Teachers' perspectives of integrating AI-powered technologies in K-12 education for creating customized learning materials and resources. *Education and Information Technologies*, *30*, 10343–10371. https://doi.org/10.1007/s10639-024-13257-y

Assaf, H.M. (2023). *A comprehensive exploration of challenges that English as foreign language learners in Kuwait encounter and suggestions for enhancement* [Unpublished Master Project]. Asia e University. https://files.eric.ed.gov/fulltext/ED631196.pdf

Bahrick, H.P., & Hall, L.K. (2005). The importance of retrieval failures to long-term retention: A metacognitive explanation of the spacing effect. *Journal of Memory and Language, 52*(4), 566–577. https://doi.org/10.1016/j.jml.2005.01.012

Bandura, A. (1977). *Social learning theory*. Prentice Hall.

Bataineh, R. F., Bataineh, R. F, Al-Barakat, A. A, & AlAli, R. M. (2026a). Stratified voices: Accent, fluency, and belonging in Jordanian graduate EFL education. *Journal of Ethnic and Cultural Studies, 13*(3), 57–79. https://doi.org/10.66815/ejecs/2896

Bataineh, R. F., Bataineh, R. F., Al-Barakat, A. A., & AlAli, R. M. (2026b). WhatsApp as a mediational infrastructure: Informal parental involvement and pedagogical drift in Jordanian primary education. *Digital Education Review*, 48, 157–172. https://doi.org/10.1344/der.2026.48.157-173

Bataineh, R. F., Obeiah, S. F., & Bataineh, R. F. (2026c). Tuning in with technology: AI-enhanced listening instruction in the Jordanian EFL classroom. *Journal of Information Technology Education: Innovations in Practice, 5,* 7. https://doi.org/10.28945/5699

Bataineh, R.F., Al-Ghoul, E.N., & Bataineh, R.F. (2025a). Backed against a wall: The potential utility of self-regulated online reading instruction. *Studies in Self-Access Learning Journal, 16*(1), 25–59. https://doi.org/10.37237/202407

Bataineh, R.F., Bataineh, R.F., AlAli, R.M., Alotaibi, S.B., Al-Barakat, A.A., Al-Saud, K.M., Aboud, Y.Z., & Ibrahim, N.A. (2025b). Digital frontiers: The transformative potential of e-learning in cultivating Arab primary school learners' creativity. *SAGE Open*, *15*(4). https://doi.org/10.1177/21582440251408317

Bataineh, R.F., Kadhim, A.M., AlAli, R.M., Kanaan, E.M., Al-Barakat, A.A., Abdullatif, A.K.A., Zaher, A.M., & Aboud, Y.Z. (2025c). Nurturing young minds: SWOM-based reading as a catalyst for reflective thinking in early childhood. *International Journal of Learning, Teaching and Educational Research*, *24*(7), 1000–1019. https://doi.org/10.26803/ijlter.24.7.47

Bouchrika, I., Harrati, N., Wanick, V., & Wills, G. (2021). Exploring the impact of gamification on student engagement and involvement with e-learning systems. *Interactive Learning Environments*, *29*(8), 1244–1257. https://doi.org/10.1080/10494820.2019.1623267

Chen, C.H., & Chang, C.L. (2024). Effectiveness of AI-assisted game-based learning on science learning outcomes, intrinsic motivation, cognitive load, and learning behavior. *Education and Information Technologies*, *29*(14), 18621–18642. https://doi.org/10.1007/s10639-024-12553-x

Chi, M.T., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, *49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823

Dai, K., & Liu, Q. (2024). Leveraging artificial intelligence (AI) in English as a foreign language (EFL) classes: Challenges and opportunities in the spotlight. *Computers in Human Behavior*, *159*, 108354. https://doi.

org/10.1016/j.chb.2024.108354
Davis, B., & Sumara, D. (2002). Constructivist discourses and the field of education: Problems and possibilities. *Educational Theory*, 52(4), 409–428. https://doi.org/10.1111/j.1741-5446.2002.00409.x
Fredricks, J.A., Blumenfeld, P.C., & Paris, A.H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*(1), 59–109. https://doi.org/10.3102/00346543074001059
Grabinger, R.S., & Dunlap, J.C. (1995). Rich environments for active learning: A definition. *Research in Learning Technology*, *3*(2), https://doi.org/10.3402/rlt.v3i2.9606
Hashemi, A. (2021). The effects of using games on teaching vocabulary in reading comprehension: A case of gifted students. *Journal for the Education of Gifted Young Scientists*, *9*(2), 151–160. https://doi.org/10.17478/jegys.846480
Hung, H.T., Yang, J.C., Hwang, G.J., Chu, H.C., & Wang, C.C. (2018). A scoping review of research on digital game-based language learning. *Computers & Education*, *126*, 89–104. https://doi.org/10.1016/j.compedu.2018.07.001
Itani, M., Palmer, K., & El-Sabbagh, R. (2024). Transforming education through digital technology: Shifting the role of innovation and critical thinking to students. In S. Baroudi & M.D. Lytras (Eds.), *Transformative leadership and sustainable innovation in education: Interdisciplinary perspectives*. Emerald Publishing. https://doi.org/10.1108/978-1-83753-536-120241012
Jomaa, N., Attamimi, R., & Al Mahri, M. (2025). The use of artificial intelligence (AI) in teaching English vocabulary in Oman: Perspectives, teaching practices, and challenges. *World Journal of English Language*, *15*(3), 1–12. https://doi.org/10.5430/wjel.v15n3p1
Kelly, P. (1991). Notes and discussion. *International Review of Applied Linguistics in Language Teaching*, *29*(2), 135–170. https://doi.org/10.1515/iral.1991.29.2.135
Landers, R.N., Armstrong, M.B., & Collmus, A.B. (2017). How to use game elements to enhance learning: Applications of the theory of gamified learning. In M. Ma & A. Oikonomou (Eds.), *Serious games and edutainment applications: Volume II* (pp. 457–483). Springer. https://doi.org/10.1007/978-3-319-51645-5_21
Nation, I.S.P. (2001). *Learning vocabulary in another language*. Cambridge University Press. https://doi.org/10.1017/9781009093873
Pokrivčáková, S. (2019). Preparing teachers for the application of AI-powered technologies in foreign language education. *Journal of Language and Cultural Education*, *7*(3), 135–153. https://doi.org/10.2478/jolace-2019-0025
Qin, Z., & Chuaychoowong, M. (2025). Artificial intelligence in EFL education in China: A systematic review of trends, gaps, and future directions (2015-2024). *Language Teaching Research Quarterly*, *49*, 59–89. https://doi.org/10.32038/ltrq.2025.49.04
Ryan, R.M., & Deci, E.L. (2023). Self-determination theory: Basic psychological needs in motivation, development, and wellness. In F. Maggino (Ed.), *Encyclopedia of quality of life and well-being research*. Springer. https://doi.org/10.1007/978-3-031-17299-1_2630
Safar, P., & Ammar, H. (2024). Awareness of AI and its uses in education among general education schools teachers in Kuwait. *International Electronic Journal*, *13*(2), 451–511. https://doi.org/10.21608/ijel.2024.410548
Salman, F.A.K., & Bataineh, R.F. (2025). Learner autonomy in the Jordanian EFL context: Constraints, culture, and classroom realities. *International Journal of English Language and Literature Studies, 14*(4), 439–459. https://doi.org/10.55493/5019.v14i4.5705
Schmitt, N. (2019). Understanding vocabulary acquisition, instruction, and assessment: A research agenda. *Language Teaching, 52*(2), 261–274. https://doi.org/10.1017/S0261444819000053
Schunk, D. H., & DiBenedetto, M. K. (2020). Motivation and social cognitive theory. *Contemporary Educational Psychology, 60*, 101832. https://doi.org/10.1016/j.cedpsych.2019.101832
Yan, W., Li, B., & Lowell, V. L. (2025). Integrating artificial intelligence and extended reality in language education: A systematic literature review (2017–2024). *Education Sciences*, *15*(8), 1066. https://doi.org/10.3390/educsci15081066
Yeung, K.L., Carpenter, S.K., & Corral, D. (2021). A comprehensive review of educational technology on objective learning outcomes in academic contexts. *Educational Psychology Review*, *33*(4), 1583–1630.

https://doi.org/10.1007/s10648-020-09592-4
Zghoul, W.M., & Bataineh, R.F. (2024). Flipgrid: Unlocking the English speaking potential of Jordanian adolescent EFL learners. *Journal of Information Technology Education: Innovations in Practice, 23*, 017. https://doi.org/10.28945/5407
Zhang, R., Zou, D., & Cheng, G. (2023). Learner engagement in digital game-based vocabulary learning and its effects on EFL vocabulary development. *System*, *119*, 103173. https://doi.org/10.1016/j.system.2023.103173
Zhonggen, Y. (2019). A meta-analysis of use of serious games in education over a decade. *International Journal of Computer Games Technology*, 4797032. https://doi.org/10.1155/2019/4797032