RESEARCH IN ENGLISH LANGUAGE EDUCATION JOURNAL

Volume 4, Issue 1 - August 2025
ISSN 3041-8909

The Effect of AI-Scaffolded Learning-Oriented Assessment on EFL Learners' Vocabulary Learning

Hassan Alizadehmahmoudalilo
Corresponding Author: English Language Department, Tabriz Islamic Azad University, Tabriz, Iran.
E-mail: hassan.alizadehmahmoudalilo@iau.ac.ir

ABSTRACT

Given the transformative potential of technology-enhanced methods in language education, this study investigated the effect of AI-scaffolded learning-oriented assessment (LOA) on Iranian EFL learners' vocabulary acquisition. A quasi-experimental design was employed, with 40 male intermediate-level learners non-randomly assigned to an experimental group (AI-supported Nearpod platform with adaptive scaffolding) and a control group. The Oxford Placement Test ensured homogeneity, while the Vocabulary Knowledge Scale measured vocabulary gains. The experimental group received AI-driven scaffolding, including real-time adaptive feedback, tiered support, and interactive peer-review tasks aligned with LOA principles. Statistical analysis revealed a significant difference in post-VKS scores ($M = 130.05$) for the experimental group compared to the control group ($M = 99.85$), with a large effect size ($t = 39.38$, $p < .001$, $d = 1.96$). These results demonstrate that AI-scaffolding, when integrated with LOA, substantially enhances vocabulary learning by personalizing feedback and promoting active engagement. The study highlights the efficacy of AI tools like Nearpod in EFL contexts while underscoring the importance of pedagogical design, such as balancing automation with teacher guidance, to maximize learning outcomes. These findings advocate for the strategic adoption of AI-scaffolded LOA in language curricula, though challenges like digital literacy and equitable access warrant further exploration.

ARTICLE INFO:
Received: 2025-05-26
Reviewed: 2025-07-14
Accepted: 2025-07-18

Keywords:
AI-Scaffolding, EFL learners, learning-oriented assessment, vocabulary learning

1. Introduction

Artificial Intelligence has transformed the manner of language learning and assessment in education in essential ways. The current example of the Learning-Oriented Assessment (LOA) is the AI-scaffolding that provides implicitly adaptive settings, forming feedback loops for individuals. Such an interactive platform as Nearpod, which

Citation: Alizadehmahmoudalilo, H. (2025). The effect of AI-scaffolded learning-oriented assessment on EFL learners' vocabulary learning. *Research in English Language Education Journal*, *4*(1), 93-111. DOI: 10.48210/relej.2025.19041.1143

has become increasingly popular due to AI-scaffolding, ensures the involvement of synchronized formative evaluation, individualized teaching elements, and interactive multimedia elements. In particular, this innovation applies to the field of vocabulary acquisition, which is focused on the aspect of communicative competence (Tiansoodeenon et al., 2023).

The knowledge of vocabulary is the spearhead of EFL proficiency and comprehension in reading. The vocabulary is the key to successful reading, as Mukhtar et al. (2023) observe, and vocabulary size appears to be the most reliable indicator of reading comprehension among learners of foreign languages, according to the existing literature. To meet the lexical and reading needs of the learners, teachers are increasingly involving the application of interactive technology in teaching (Zahran, 2025). Recent research (e.g., Aulia et al., 2024; Aji et al., 2020) outlines that technology-enhanced teaching might be beneficial, useful, and inspiring, and might reinforce the critical thinking and autonomy of learners. Such patterns follow constructivist theories of learning: both Vygotsky and his sociocultural school underlined that learners construct knowledge during scaffolded social interaction with the more competent partners (Alazemi, 2024), and Bruner also prescribed guided facilitation, which is withdrawn as the student advances. In the contemporary classroom, such interactive tools as Nearpod may become these more knowledgeable others and maximize the Zone of Proximal Development, serving to guide and give appropriate feedback during the language practice.

The Nearpod platform is directly focused on such scaffolded active learning. It enables educators to create multimedia-intensive lessons that are consumed by students on computers or any mobile device (Hernandez-Mena et al., 2024). These lessons are able to incorporate a wide range of interactive elements them such as teachers being able to place quizzes, polls/surveys, open-ended questions, and even drawing boards within a slide (Mekota & Marada, 2020). Through this, every Nearpod lesson is also an interactive experience with students responding in real-time with feedback. The provided affordances allow differentiated instruction and cognitive support as the key to sustained vocabulary development (Inam et al., 2023).

Nearpod is a flexible tool that contains a set of features that promote communicative learning and live evaluation. An important characteristic of it is the possibility to embed quizzes and polls in lessons, which enables teachers to ask multiple-choice or poll questions at any time and get immediate responses concerning student comprehension. Van Le and Doan (2023) draw attention to the fact that the tools give teachers immediate insights into understanding of the students. Also, Nearpod allows collaborative boards and open-response assignments, in which the students are able to write or draw on a common board or type an answer in free format at some point in the lesson. This allows peers to interact as well as makes the thinking of students visible, which is a community to help build on knowledge using social interaction (Zahran, 2025). The real-time act of monitoring and analysis by way of the Nearpod dashboard should also be noted as another valuable feature since it allows teachers to see the responses of all of their students in real-time and follow their progress. The analytics of the platform will allow finding patterns like frequent mistakes and correcting the guidance dynamically (Alazemi, 2024).

Collectively, all these make traditional slideshows into scaffolded learning experiences. As an example, a formative assessment can be provided by an embedded quiz: the teacher can reteach some concept based on the answers of most of the students.

To sum it up, the design of Nearpod incorporates the principles of constructivist scaffolding through a combination of instruction and the provision of interactive feedback at every stage. AI-powered LOA combines formative assessment with adaptive scaffolding and presents a delivery of a level of personalized feedback that can sustain lexical depth (Teng, 2022). Systems such as Nearpod make use of AI to deliver synchronous formative evaluation and interactive exercises, congruent with the socio-cognitive principle of learning (Vygotsky, 1978; Carless, 2007). The effectiveness of AI in vocabulary acquisition has been recently proven true (Wang et al., 2024), yet only a few studies are focusing on how AI could be utilized in the scenarios of a structured LOA, specifically, in an underrepresented setting, such as Iran.

Although people have become more interested in AI-amplified learning, there is a decisive lack of empirical research examining the outcomes of using tools, such as Nearpod, in a structured, assessment-embedded teaching strategy. Of special interest in the Iranian EFL context, little research has been done to compare how Nearpod, as an AI-based LOA technology, compares with traditional vocabulary instruction when related to measurable vocabulary gains. The majority of the available research pays attention to either a blanket study of CALL or perception related to learners, thus not isolating the assessment aspect of these technologies.

It has been proven empirically that Nearpod has the potential to enhance the learning of English. The application of Nearpod has demonstrated a rise in engagement, motivation, and language performance in different studies. As another example, researchers who analyzed the situation in high school students using Nearpod discovered that these students were much more engaged and motivated compared to the traditional lectures (Al-Shehri, 2018). Similarly, Hakami (2020) mentions that the collaborative quizzes and immediate responses through Nearpod facilitate active learning and remarkably boost the vocabulary acquisition and memory of the EFL students. Similarly, as Putra et al. (2021) found, the incorporation of Nearpod formative quizzes as a teaching tool in classes enhanced critical thinking abilities and further enabled the development of reading comprehension in learners. According to Musa and Momani (2022), Nearpod is also flexible, and due to that, it is also useful in developing English language proficiency as the software can address the needs of heterogeneous learners through multimedia-based activities and pacing. In an experiment conducted in a classroom recently, Zahran (2025) demonstrated that primary EFL learners who were taught with Nearpod (and guided reading) performed much better than the controls both in terms of their reading comprehension and motivation.

Besides engagement, Nearpod is ideal for formative assessment, which complies with the principles of LOA. Notably, Black and Wiliam (1998) advanced a famous argument that says that formative (assessment-for-learning) activities contain important feedback that can be used to adjust teaching and enhance educational objectives by both teacher and student. The idea is carried further by LOA, whereby assessment itself is presented in terms of learning: it establishes learning, as well as the learner, as the center stage of instructional activities and makes the students active participants in the feedback loop (Zahran, 2025). Quizzes, polls, and open-ended tasks embedded in Nearpod are also the perfect way to include assessment in the learning process. Each response that students provide gives instant information: teachers (and students themselves) know where there is knowledge or ignorance. This feedback loop with the data is constant, and it is part of the LOA assessment, which means learning assessment. Practically, with Nearpod, the quizzes and polls become not only moments of the test

but instruments of thinking and piloting, transforming the assessment tasks into learning tasks.

In summary, the theory and the research point to the potential of Nearpod, its multimedia lessons and interactive activities are outlined to facilitate a scaffolded form of support (according to Vygotsky/Bruner), as well as immediate feedback that exerts active learning objectives, achieving learning. The major past research has, however, not been specific in the sense that it has analyzed either specific vocabulary acquisition or the dynamics of scaffolding involved within an AI-enhanced environment. It is on this background that this paper helps to fill this research gap by examining the impact of AI-assisted, scaffolded tasks in Nearpod as part of an LOA framework on vocabulary learning among intermediate-level Iranian EFL students. Specifically, very few studies examined the possibilities of using the specific functions of Nearpod within a purposeful LOA framework to deliver vocabulary instruction. This will allow conducting the study using the experimental method of research design and validated measurement scales, and it is expected to produce empirical data on the efficiency of AI-powered tools of assessment in terms of promoting lexical development.

Research Question: Does AI scaffolding within an LOA framework significantly enhance vocabulary acquisition among Iranian EFL learners?

2. Review of Literature

The growing integration of artificial intelligence in language education has reshaped how vocabulary is taught and assessed in EFL contexts. Recent studies have explored AI applications such as chatbots, adaptive platforms, and intelligent tutoring systems, highlighting their potential to support learner autonomy, provide real-time feedback, and enhance vocabulary acquisition. Additionally, LOA has gained attention for its emphasis on formative feedback and active learner engagement. This review synthesizes current research on AI tools and LOA practices, identifying their strengths and the emerging potential of combining them to support vocabulary learning. However, gaps remain regarding their combined use in structured classroom settings.

2-1. AI Applications in EFL Education

Recently, the insertion of AI technologies in language learning has blown up. According to Jiang (2022), automated assessment or other AI-driven tools such as intelligent tutoring systems and conversational chatbots are utilized in the EFL classrooms more frequently. As an example, automated writing assessment and speaking tutors were proven to enhance the grammatical accuracy and oral proficiency of learners without decreasing the degree of learner motivation (Teimourtash, 2024). The new forms of glossing are also represented by Neural Machine Translation tools and AI chatbots: Wang and Petrina (2013) revealed that conversational bots might correspondingly improve EFL students' learning of both grammatical and new vocabulary.

However, despite these advantages, AI in education triggers concerns, as well. Lack of quality and reliability of feedback keeps many teachers vaguely positive about the concept of automated evaluation (Jiang, 2022). Pronunciation, grammar, and creative speech remain a challenge to chatbots in language learning, although in any case, they are all at a low level. According to Lotze (2018), the existing AI methods of diagnosis might not be spontaneous and shared, which restricts their effectiveness as electronic teachers. This means that AI needs to be pedagogically integrated

appropriately. As an example, it is possible to consider the use of AI tools as the human in the loop AI tools, which presupposes the positioning of AI tools as structures intended to assist the human in the process of creative thinking (i.e. generation of ideas or recalling vocabulary/words) (Li & Wilson, 2025).

Some of the studies highlight the potential and difficulties of AI in EFL contexts. Jiang (2022) outlines six influential forms of AI that exist in EFL (automatic evaluators, translation systems, ITS, chatbots, virtual environments, and affective computing) and mentions improvements in accuracy, motivation, and engagement. Teimourtash (2024) individually investigated an AI-chatbot task based on writing skills by Iranian EFL students and determined that the learners who had an AI scaffolding of teacher assistance considerably outperformed control learners on the writing and anxiety indicators. Overall, a recent review by Li and Wilson (2025) based on 14 studies described three main categories of AI scaffolding in EFL writing, namely cognitive supports (e.g., grammar correction), creative supports (e.g., idea generation), and language enhancement (e.g., vocabulary expansion). Importantly, they say that AI tools can be used to alleviate cognitive load and cultivate self-regulation when used in tandem with theory and equal access.

The effectiveness of Nearpod is supported by empirical research that proves the application works in a variety of learner groups. Studies with elementary students show much increased rates of engagement and retention of vocabulary words, which is due to their gamified screens and visual support (Balqis & Zaki, 2025). Likewise, adult students have a better command of difficult lexicon objects (e.g., phrasal verbs) and are more independent due to the response loops of Nearpod in real-time (Sanchez & Carballo, 2025). In addition to platform-specific results, AI-based systems can have a general improvement in lexical recall over the long term by, through recursive exposure, strategically re-introducing the target vocabulary at the most effective rate to overcome the forgetting curve (Teng, 2022; Wang et al., 2024). However, knowledge about the manner in which AI scaffolds LOA remains elusive in less-studied settings. The Iranian EFL contexts show the nascent technological integration, and the teacher-centered approach is prevalent, so this area is quite a gap (Seyed & Tavassoli, 2023).

Recent empirical research provides solid proof of Nearpod's effectiveness in facilitating improved vocabulary results in various EFL settings. Sanchez and Carballo (2025) conducted a quasi-experimental study to prove that Nearpod-driven teaching substantially enhanced the learning and mastery of phrasal verbs due to the real-time feedback and gamification that participants were exposed to. In the same way, Gamlo and Alzahrani (2024) recorded a 32 per cent increase in vocabulary retention of 87 Saudi Arabian university learners who had collaborated with Nearpods collaborative boards and polls, and the qualitative observation reaffirmed high motivation. In younger demographics, Balqis and Zaki (2025) observed better scores in vocabulary in Indonesian secondary students using Nearpod quizzes with embedded images (visual scaffolding) than the traditional education methods and attributed the enhanced lexical encoding to a visual scaffold. In addition to the Nearpod-specific implications, an assessment of 21 AI-based vocabulary systems (of which Nearpod is one) in a meta-analysis by Wang et al. (2024) found better longer-term retention than standard instruction on a per-study basis (12 total) and differentiated the development of anticipation and prediction capabilities to support the utility of spaced-repetition and feedback-context-looping that the AI can facilitate in lexical acquisition.

The results indicate that there is undoubtedly a role of AI in scaffolding language learning (including vocabulary through scaffolding), offering learners immediate, adaptive responses. Studies that are vocabulary-specific are not as numerous, although there is some evidence. E.g., the use of virtual reality environments (with AI) enhances memory of new words, and intelligent tutoring systems associated with the implementation of flipped classrooms facilitate the acquisition of vocabulary (Jiang, 2022). This paper is an expansion of this kind of work that aims at AI-chatbots dedicated to vocabulary training. The chatbot was designed to utilize word sense in context, proffer a hint of synonyms/antonyms, and give a corrective hint in order to serve as a scaffold in the use of the words.

2-2. Scaffolding in Language Learning

Scaffolding in teaching is used as a term associated with giving learners the support needed to help them do something that cannot be done due to their unique abilities alone. The theory behind scaffolding can be seen in the Zone of Proximal Development as developed by Vygotsky (1978). In language education, situations are scaffolded by teacher prompts, modeling, and feedback that are withdrawn as language learners get better. Several research studies have proven that scaffolding may help boost EFL effectively. Teacher- and peer-scaffolding, in their turn, have both been proved to be able to enhance incidental vocabulary learning and reading comprehension (Jamali Kivi et al., 2021). Overall, scaffolding does lead to learner independence. According to the findings of Awadelkarim (2021), there is a tendency of EFL teachers to act positively regarding the use of scaffolding, as it helps to promote autonomy. Yet, he also revealed weaknesses in the knowledge and the practice of the instructors, which showed that the training additionally required improvements.

Scaffolding may also be technology-mediated in an AI context. The AI tools can be used as cognitive aids since they lead learners in their actions. To take an example, the AI chatbot used in the present study served as a kind of digital scaffolding: it oversaw vocabulary activities of the students and provided hints or corrections in situations where they were necessary. The previous research indicates that this method may be quite efficient: in the study by Teimourtash (2024), students who were provided with the teacher guidance and AI scaffolding had better results writing without too much anxiety. Analogically, we surmise that vocabulary would be the same: AI scaffolding ought to offer the same results to learners and give them the feedback and the practice that would allow them to internalize new words faster than they would be able to do without it.

2-3. Learning-Oriented Assessment (LOA) in EFL

Learning-oriented assessment (LOA) is one of the philosophies of assessment that underlines the learning process. The prominent among them encompass the assessment tasks being used as learning tasks, students being involved in self- and peer-assessment, and providing feedback geared toward improvement in the future (Carless, 2017). According to Carless (2015), LOA is described as taking the learners as the focal point of an assessment system by seeking to establish a learning rather than an exam culture. Practically, LOA could lead to students reflecting on some standards, updating their work based on reaction, and working on team projects that indicate target skills. Admittedly, in the recent surveys, the EFL teachers admitted the helpfulness of LOA to

motivation and autonomy, yet, time limit, absence of resources are mentioned as some difficulties.

LOA is especially applicable in the context of vocabulary acquisition, as knowledge of vocabulary would be improved with repetition and correction. Rather than using tests as conclusive tools, the LOA would make vocabulary practice possible through formative feedback loops: e.g., making individual word portfolios, answering each other's questions with teacher support, or testing themselves with rubrics. The strengths of both AI scaffolding and LOA are used by combining both in this study. The AI delivers instant and personalized feedback (as is typical of a tutoring system), and the LOA design makes sure that students take an active role in the reflection and manipulation of this feedback and in developing their learning strategies.

In summary, the reviewed literature underscores the growing role of AI technologies in enhancing EFL instruction, particularly in vocabulary acquisition. Tools like Nearpod and AI chatbots have demonstrated potential in fostering engagement, scaffolding learner autonomy, and delivering timely feedback aligned with learning-oriented assessment principles. While prior studies have highlighted the benefits of AI in areas such as grammar and writing, there remains a noticeable gap in empirical research that specifically investigates AI-assisted LOA frameworks for vocabulary development, especially in underexplored educational contexts like Iran. This gap positions the current study to contribute meaningful insights by examining how AI-powered scaffolding, integrated within an LOA model, influences vocabulary learning outcomes among intermediate-level Iranian EFL learners.

3. Theoretical Framework

In sociocultural theory, scaffolding is the temporary support (hints, modeling, feedback) that enables a learner to perform tasks beyond their current ability (Vygotsky & Cole, 1978). Vygotsky's Zone of Proximal Development (ZPD) defines the gap between what a learner can do unaided and what they can do with guidance. Learning is achieved by guidance from a more knowledgeable other (MKO) through scaffolding (Simina, 2012). Bruner described scaffolding as the teacher's vicarious consciousness, a temporary intellectual support that draws the learner to higher understanding. Over time, this support is gradually removed as learners gain independence.

LOA is an assessment approach that emphasizes learning rather than mere evaluation. Carless (2007) defined LOA by its three elements: designing assessment tasks as learning tasks, involving students as self- or peer-evaluators, and providing feedback as feedforward for future learning. In practice, LOA involves creating learning-focused tasks, developing students' capacity to self-assess, and engaging learners in a timely feedback loop (Carless, 2007). This approach aligns assessment with pedagogical goals by making every task an opportunity to learn.

Digital scaffolding refers to adaptive support delivered through technology (e.g., automated hints, explanations, modeling) that reduces learners' cognitive load. Recent work argues that generative AI tools (such as ChatGPT) can serve as digital MKOs in a social constructivist framework (Tran et al., 2025). For instance, Jacobs Foundation research notes that evolving digital resources can take over aspects of scaffolding, offering explanations, highlighting key relations, and guiding attention, thereby easing the teacher's burden. In language learning, an AI tutor can provide real-time feedback, answer questions, or generate practice exercises tailored to the learner's level, effectively scaffolding new knowledge.

Vocabulary acquisition is the process of learning and retaining new word meanings and usages. In second-language learning, vocabulary breadth (number of words known) and depth (quality of knowledge) are critical dimensions of proficiency. Mastering a large vocabulary is often the initial goal for EFL learners, since deficiencies in vocabulary can severely limit communication and comprehension (Nation, 1990). Effective vocabulary learning requires rich input, repeated exposure, and feedback, often occurring best in meaningful contexts (e.g., through reading, speaking, or interactive tasks). In the EFL classroom, learners typically need frequent scaffolds (e.g., translations, images, example sentences) to acquire new words successfully.

Vygotsky's theory posits that learning is inherently social and collaborative. The ZPD is navigated via guided interaction (Simina, 2012). With appropriate scaffolding from an MKO (teacher, peer, or intelligent system), learners internalize new skills and eventually achieve independence. Bruner extended this to education by showing how instructors use scaffolds (such as questioning, modeling, and feedback) to guide children's learning processes. Thus, both Vygotsky and Bruner view knowledge as constructed through interaction, with scaffolding as the vehicle for moving learners from assisted to autonomous performance.

Carless (2007, 2015) reframes assessment as a driver of learning. Instead of assessment merely measuring performance, LOA designs tasks that promote learning. Carless (2007) argued that assessment should involve students actively (e.g., via self-/peer-assessment) and use feedback to feed forward into learning. In a 2015 model, he highlighted three interrelated processes: assessment tasks that encourage thinking and practice, students developing self-evaluative capacity, and timely feedback promoting dialogue and reflection (Carless, 2015). In language education, LOA implies using quizzes, projects, and peer review not just to grade students, but to engage them in continual learning. This aligns with communicative EFL pedagogy by making learning student-centered and reflective.

Technological tools are increasingly interpreted through constructivist lenses. For example, Tran et al. (2025) argue that generative AI can fulfill the role of an MKO in a social-constructivist framework. Recent empirical studies support this notion in language learning. Abdelhalim and Alsehibany (2025), grounded in Vygotsky's theory, found that integrating ChatGPT into EFL classrooms significantly improved students' productive vocabulary knowledge. Students using ChatGPT received real-time feedback and scaffolded practice, which enhanced engagement and contextual vocabulary use (Abdelhalim & Alsehibany, 2025). Likewise, Hago et al. (2025) report that AI-driven vocabulary systems (e.g. adaptive chatbots) dramatically outperform traditional methods when they use dynamic scaffolding, adjusting support and context to the learner, achieving far greater gains in retention. These studies illustrate how AI can operationalize scaffolding and learning-oriented tasks in practice.

Using this framework, this study conceptualizes AI-assisted vocabulary practice as LOA-style tasks with built-in scaffolds. The research question asks whether AI scaffolding within an LOA framework enhances Iranian EFL learners' vocabulary acquisition. Under this framework, learners engage in vocabulary exercises while AI provides scaffolded guidance and feedback. The Vygotskian/Bruner scaffolding lens predicts that such support will help students perform in their ZPD, leading to greater gains. Carless's LOA lens predicts that making tasks learning-centered will deepen learning. Together, they shape the study: the design, the metrics, and the interpretation all derive from this theoretical foundation.

4. Methods

This section outlines the research design, participants, instruments, procedures, and data analysis methods used to investigate the effect of AI-assisted learning-oriented assessment on vocabulary acquisition among Iranian EFL learners.

4-1. Research Design

This quasi-experimental research, which had only a post-test-control group design, required the existence of two groups: an experimental group and a control group. The experimental group was given treatment by the innovative methodologies of the AI-assisted learning-oriented assessment, and the control group was instructed conventionally. Thus, the dependent variable in the scope of this study is vocabulary learning, while AI-scaffolding LOA serves as the independent variable.

4-2. Participants

The participants were selected from a population of 350 from a private institute in Tabriz, namely Goldis. The final sample consisted of 40 lower-intermediate EFL learners, all male, aged 15 to 24, who were native speakers of Azari and were recruited from institutes that taught identical curricula. This technique ensured that study subjects had similar academic backgrounds. Participants were selected through a convenience sampling method because of practical limits like accessibility and the willingness to participate. This approach guaranteed that the participants chosen would be representative of the typical population of lower-intermediate EFL students at institutes, as they were exposed to similar educational backgrounds, standard of language proficiency, and stages of learning. According to the placement criteria of the institute, they were lower-intermediate students. Nonetheless, to guarantee the integrity and uniformity of the participants, a proficiency test was administered before the commencement of the primary research. Based on the proficiency test, the candidates who scored between 30 and 39 were included in the study and rated as proficient. The participants were then non-randomly assigned to an experimental group and a control group, each with 30 students. No participants had prior experience with the AI tool used. In order to remove the confounding variables of teachers, both groups were instructed by the same teacher. The participants had studied Evolve 1-3. They further taught using the Evolve 4 textbook during the study to maintain consistency in the teaching content and exposure across both groups

4-3. Data Collection Instruments

To gather the data needed for the study, the researcher applied the following instruments at various stages of the study.

4-3-1. Oxford Placement Test

The Oxford Placement Test (OPT), developed by Dave (2004), was systematically applied to assess and verify if the proficiency levels of the English language differed in any significant ways between the experimental and control groups investigated. This is a well-known and standard examination test that properly and effectively ascertains the knowledge of language at various levels of the Common European Framework of Reference for languages (CEFR). As an evaluation instrument, the OPT is considered one of the main characteristics since it serves as the homogenizing tool, which measures

a variety of linguistic skills such as grammar, vocabulary, and reading comprehension. The test is a well-structured formal evaluation diluted to six levels of proficiency on the CEFR scale and assigns test scores to well-defined value boundaries for each of the discrete levels: Basic (A1: 0–17), Elementary (A2: 18–29), lower intermediate (B1: 30–39), upper intermediate (B2: 40–47), advanced (C1: 48–54) and very advanced (C2: 54–60). To ensure that all participants fell strictly within the lower-intermediate (B1) band and not the adjacent upper-intermediate (B2), only those whose scores ranged from 30 to 39 were included in the study. This range of scores was a strict cut-off point; any learner who scored 40 or more or below 30 was left out of the sample. This process ensured that the participants had the same and acceptable standardized level of proficiency that matched the CEFR B1 standard. Those OPT results obtained at the beginning of the research were crucial in the sense that they gave researchers the opportunity to place participants whose results would be placed at the lower intermediate level on the scale intentionally to ensure that the language proficiency level was identical in the groups. Moreover, the psychometric characteristics of the OPT have been tested in previous studies that report high reliability and construct validity of the tool in the EFL environments. Normally, the internal consistency of the grammar part of the OPT and the listening part of the test is more than 0.85 Cronbach's alpha (Geranpayeh, 2003). The results of the current study showed that the Cronbach alpha of the OPT on a pilot sample (n = 20) was 0.87, which indicates high internal reliability of this group of Iranian EFL students. In addition, the test was CEFR-based, making the process of decisions made on the placement content valid.

4-3-2. Vocabulary Knowledge Scale

The Vocabulary Knowledge Scale (VKS, Wesche & Paribakht, 1996) was employed to measure students' vocabulary knowledge before intervention as pre-VKS and after intervention as post-VKS. This instrument, originally developed as a comprehensive word knowledge test, requires language learners to demonstrate their familiarity and usage of target words using a five-point scale that ranges from complete unfamiliarity ("I don't remember having seen this word before") to the ability to use the word in a sentence accurately and appropriately. The VKS assesses two main constructs: vocabulary size, which is measured through four items that capture the continuum from total unfamiliarity to correct meaning identification, and vocabulary depth, which is evaluated by asking students to produce a grammatically and semantically correct sentence using the word. This scale was chosen for its ability to provide verifiable evidence of both receptive and productive knowledge, making it an ideal tool for research focused on word identification and utilization in EFL contexts. To ensure cultural and linguistic relevance, the VKS instructions were translated into Persian and administered on a separate sheet, and its validity has been supported by previous research. In the pre-VKS, to verify that the students were unfamiliar with the vocabulary they were expected to learn during their EFL classes in the treatment period, an 80-item vocabulary scale was administered before the experiment. After analyzing the questionnaire responses, 50 items that the students did not recognize were selected as the target words for treatment, while the 40 items that were familiar to the students were removed from further consideration. For the post-VKS, these 50 unfamiliar words were employed to assess any vocabulary gains resulting from the treatment. In the current study, scoring was conducted independently by two raters to ensure inter-rater reliability, with responses scored as follows: a score of 0 for complete unfamiliarity, 1

for basic recognition without understanding, 2 for correctly providing a synonym or translation, and either 3 or 4 for using the word in context, with a three assigned for contextually correct but ungrammatical usage and a 4 for fully correct usage, resulting in a per-word score that ranges from 0 to 4. Both raters engaged in a discussion to resolve any discrepancies in scoring to maintain consistency. This approach ensured that differences in interpretation were addressed collaboratively, leading to a more reliable and standardized assessment process.

4-3-3. AI-Chatbot Scaffolding Tool

Regarding the experimental group, ChatGPT was designated as a conversational AI chat robot personalized in vocabulary training. The chatbot would be able to communicate on the target words, give example sentences, synonyms/antonyms, and soft error-correction. Notably, the chatbot was set to adhere to the principles of pedagogy: it encouraged the learners to employ new words, avoided providing answers but hinted instead, and was friendly with the learners in accordance with scaffolding principles. We can defend our use of such AI based on scalability and interactivity; in recent studies (e.g. Xu & Warschauer, 2020), children and those who have language learning disabilities have shown to establish a fertile exchange with conversational agents, and this can be used to combat anxiety and boost participation.

4-4. Data Collection Procedure

The current study aimed to examine the effect of learning-oriented assessment via AI on vocabulary acquisition by Iranian EFL learners through the medium of an only post-test quasi-experimental design. This was based on the solid, reliable, and ethical use of principles, and procedural steps were applied and systematically adopted chronologically between November 2024 and January 2025 at Goldis Language Institute, Tabriz, Iran. The research was followed through the Ethical Committee of the Goldis Language Institute, and they served as the passage of the research to ensure that, before the research, they considered matters that relate to ethics in research. The recruitment was done in normal classes. The researcher explained the aim of the study, the steps involved in it, and informed the participant that the study is voluntary and that the participant will be assured of anonymity and confidentiality, and that he or she was free to withdraw himself or herself at any point without repercussions. All the subjects signed an informed written consent obtained using printed forms, returned it, and kept it in a locked filing cabinet accessible to the researcher.

Before conducting the main data collection, a pilot study was conducted to enhance the clarity and reliability of the research instruments applied. 20 learners were used in the pilot test, and they were found to belong to the same level of proficiency as the actual sample. This pilot study was undertaken to identify the test items' ambiguities, the time requirement to answer the test items, and to prove the validity of the Persian version of the VKS. The pilot study indicated that the instrument was very reliable in internal consistency, and the Cronbach alpha coefficients were 0.79 on VKS, which warrants its use in the main study.

In order to homogenize the participants, the OPT was initially carried out in 60 minutes session in which all participants participated. Manual collection of answer sheets and scoring by the researcher based on the official scoring key was also adopted. To be consistent, only the participants with the 30-39 scores (lower-intermediate, B1) were included. Of the initial 350 male students, only 40 were retained, with others being

eliminated on the basis that their scores fell beyond this range. A total score was manually entered in a spreadsheet, with a second colleague proofreading the scores. The study involved participants whose sample size was twenty (n=20), and they were distributed under two categories, Experimental and Control. Such a natural assignment caused inner institutional uniformity. All the groups were taught by one single instructor to limit the teacher variation since the instructor was well-versed in both Persian and English.

Prior to the treatment, the VKS, as a pre-test, was administered in an 80-minute session. Scores were recorded manually on a scoring sheet. This test was conducted to ensure learners' unfamiliarity with target words. The words that learners were familiar with were discarded from the treatment. It was distributed in print and recorded in the spreadsheet.

Over eight weeks, each group received distinct instructional approaches tailored to their assigned assessment method. The experimental group engaged with customized Nearpod software designed to deliver dynamic scaffolding during vocabulary tasks. During weekly 50-minute sessions, students interacted with Evolve 4 reading activities containing target words, triggering four levels of computerized mediation upon errors: implicit prompts (e.g., contextual highlighting), contextual clues (synonyms/definitions), explicit explanations (grammatical rules), and direct answers. The software logged responses and mediation usage, while the instructor monitored progress without direct intervention. The Experimental group did parallel tasks but with AI support: after initial explanation, students interacted with the chatbot on a classroom tablet or computer. For instance, an experimental task might be: "Practice using this word by chatting with the AI." The AI would prompt the student (Can you use 'explain' in a sentence?) and respond with feedback. All lessons in this group followed LOA principles: students often checked each other's responses, used rubrics to self-correct, and the teacher provided formative feedback instead of immediate grading. In summary, the experimental condition combined AI scaffolding (via chatbot) with LOA practices (peer/self-review, reflective tasks).

Both groups studied the same lexical content (e.g., five words per lesson) integrated into communicative tasks, following the Evolve 4 curriculum with equal session durations. Instruction was delivered by the same trained teacher to ensure methodological consistency and isolate the effects of AI-assisted intervention. The control group used traditional methods, including teacher explanations, example sentences, textbook drills, rote memorization, and paper-based exercises, with corrections limited to end-of-unit tests and no AI support. In contrast, the experimental group engaged in parallel tasks supplemented by AI scaffolding through a classroom-based chatbot. After initial instruction, students interacted with the AI—for example, responding to prompts like "Can you use 'explain' in a sentence?"—and received real-time feedback. These sessions followed learning-oriented assessment (LOA) principles, incorporating peer review, self-assessment using rubrics, and formative feedback from the teacher. Both groups completed identical summative quizzes at the end of the study.

Identical to the pre-test, post-intervention data were collected after conducting treatment, and inter-rater reliability was calculated, with discrepancies resolved through discussion to ensure consistency in post-VKS. All data were stored and anonymized using participant codes (e.g., Ex-01). The same classroom conditions (e.g., lighting, seating) and timing (morning sessions) were maintained across institutes to minimize

external variables. Data analysis was conducted in SPSS 29, with post-test scores compared using an independent sample t-test, ensuring statistical rigor.

4-5. Data Analysis

The collected data were entered into SPSS 29 for further statistical analysis. At the onset, the OPT scores checked the initial homogeneity between the two groups. Then, Cronbach's alpha was used to check the internal consistency of the VKS. Descriptive statistics, including mean and standard deviation (SD) and standard errors (SEs), were presented for VKS. The Pearson correlation coefficient was used to evaluate inter-rater reliability between the two raters. The researcher used a Normality test to check the normal distribution of data. In the case of normal data, an independent sample t-test was used to explore the effect of the independent variable on the dependent variable.

5. Results

In order to answer the posed research question, some calculations, statistical routines, and results were produced. The results from the analysis of the post-VKS administered to both groups are indicated below. The details about descriptive statistics of groups regarding the post-VKS are illustrated in Table 1.

Table 1

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
Pos-VKS	Experimental group	20	130.05	2.372	.530
	Control group	20	99.85	2.476	.553

As Table 1 demonstrates, the mean score of the post-VKS for the experimental group is 130.05 (*SD*= 2.372, *SE*= .530), and the control group had a mean of 99.85 (*SD*= 2.476, *SE*= .553). Additionally, the Pearson correlation coefficient was used to evaluate inter-rater reliability and compare the consistency between both raters. Table 2 outlines these analyses.

Table 2

Inter-Rater Correlation for the Post-VKS Scores

		Rater 1	Rater 2
Post-VKS of Control Group (Rater 1)	Pearson Correlation	1	.945**
	Sig.(2-tailed)		.000
	N	20	20
Post-VKS of Experimental Group (Rater 2)	Pearson Correlation	.945**	1
	Sig.(2-tailed)	.000	
	N	20	20

**. Correlation/is/significant at the 0.01 level (2-tailed).

As Table 2 displays, for the post-VKS scores of the control group, the inter-rater correlation was almost perfect for the control group, as r = .945 (p < .001), i.e., excellent scoring consistency. In the same way, the post-VKS scores of the experimental group exhibited identical reliability (r = .945, p < .001), indicating that raters consistently applied the scoring criteria between groups after intervention. Table 3 presents the results of normality tests conducted on the post-VKS scores for both the experimental and control groups. The Kolmogorov-Smirnov and Shapiro-Wilk tests were used to assess whether the data followed a normal distribution, a key assumption for parametric statistical analyses like the independent samples t-test.

Table 3

Tests of Normality

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Group	Statistic	df	Sig.	Statistic	df	Sig.
Pos-VKS	Experimental group	.192	20	.053	.949	20	.356
	Control group	.176	20	.106	.963	20	.612

a. Lilliefors Significance Correction

As Table 3 demonstrates, for both groups, the *p*-values for the Kolmogorov-Smirnov and Shapiro-Wilk tests were greater than the conventional alpha level of .05, indicating that the data did not significantly deviate from normality. This supports the use of parametric tests (e.g., t-tests) for further analysis. Table 4 displays the results of the independent samples t-test comparing the post-VKS scores of the experimental and control groups.

Table 4

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	MD	SE Difference	Lower	Upper
Pos-VKS	Equal variances assumed	.062	.805	39.379	38	.000	30.200	.766	28.647	31.752
	Equal variances not assumed			39.379	37.930	.000	30.200	.766	28.647	31.752

As Table 4 displays, Levene's Test for Equality of Variances indicated no significant difference in variances between the groups (F = .062, p = .805), confirming the assumption of homogeneity of variance. The t-test revealed a highly significant difference in vocabulary knowledge scores between the experimental group (M = 130.05) and the control group (M = 99.85), with a t-value of 39.379 (df = 38) and a significance level of p < .001. The mean difference of 30.20 (SE = .766) was substantial, with the 95% confidence interval ranging from 28.65 to 31.75. These results strongly suggest that the AI-scaffolding in LOA had a significant positive impact on

vocabulary acquisition among Iranian EFL learners when compared to traditional methods.

6. Discussion

The results of this study offer strong evidence for the efficacy of AI-supported LOA in promoting vocabulary learning among Iranian EFL students. The experimental group, which was taught using the Nearpod platform combined with scaffolded AI-based assessment, performed decidedly better compared to the control group, which underwent traditional instruction. The statistical significance of the difference in post-VKS scores, with a large effect size, confirms the pedagogical benefit of coupling AI with formative assessment tools.

These findings are aligned with earlier research that has underscored the facilitative role of technology-supported learning environments in vocabulary acquisition (Shamshiri et al., 2023; Wang et al., 2024). Nearpod's adaptive and interactive nature probably led to enhanced engagement, higher-level cognitive processing, and more individualized feedback—all of which are essential in facilitating long-term vocabulary retention. The scaffolding features of the platform enabled students to obtain real-time hints and corrections, facilitating instant awareness and adjustment, which is usually absent in conventional teaching (Dujardin et al., 2021; Balqis & Zaki, 2025). This research also adds to the general literature on learning-oriented assessment by operationalizing AI not just as a content delivery system but as an active agent in the feedback and assessment process. This corroborates earlier research by Seyed and Tavassoli (2023), who underlined that LOA can considerably enhance vocabulary learning outcomes when students are exposed to continuous, personalized, and recursive feedback mechanisms.

The main innovation here is the integrated model of AI scaffolding and LOA. Previous research has investigated AI feedback on writing or grammar (e.g., Teimourtash, 2024), and others have investigated LOA practices (Carless, 2015), yet few have synthesized these lines for vocabulary acquisition. In doing this, the study offers a complete package: AI delivers the real-time, personalized assistance (serving as a cognitive scaffold), and LOA guarantees students actively reflect on and interact with that assistance (e.g., through peer/self-assessment activities). This double emphasis probably increased gains. The results also support Li and Wilson's (2025) proposal that AI tools can enhance creativity and autonomy when combined with student decision-making. In our research, students were not passive consumers of AI output; they employed rubrics and peer checks to digest the feedback, realizing LOA ideals.

For Iranian EFL environments (and comparable environments), these findings have practical applications. They indicate that even in big classes, scalable AI tools (such as chatbots) can enhance vocabulary practice without overloading the instructor. Teachers can apply straightforward AI chat functions on PCs or smartphones to augment instruction. Yet our research also signals challenges. In line with Yang (2024) and others, digital equity arose as an issue: we arranged for all students to have device access in class, but outside the classroom, some indicated limited connectivity or comfort with the AI interface. This echoes the "digital divide" Yang et al. (2024) outlined in online TESOL – a divide that can deepen learning disparities. Guaranteeing equitable access (e.g., providing devices, offline AI materials) is imperative if AI scaffolding is to reach all learners, not just the technologically well-resourced.

Teacher variables are also important. Our teachers received targeted training in both the use of the chatbot and in LOA approaches, which no doubt impacted achievement. As Xu and Warschauer (2020) point out, teacher comfort with technology is a cornerstone of successful AI integration. In practical terms, this implies that professional development initiatives should address not only tool mechanics but also pedagogy: how to craft clear AI prompts, how to interpret AI feedback, and how to counsel students in assessing AI suggestions. Without such support, the danger is that AI tools are underutilized or misutilized, reaffirming the warning in earlier research that technology per se does not assure better results (Xu & Warschauer, 2020). Our findings reiterate the lesson: AI is most effective as part of a teacher-created system.

Nonetheless, while the experimental group outperformed, the study also highlights crucial considerations regarding implementation contexts. Even as Nearpod's efficacy is evident, its usefulness may be contingent upon sufficient digital literacy, instructor training, and infrastructural support. In addition, concerns raised in previous research, e.g., reduced teacher presence or over-reliance on automation (Alharbi & Khalil, 2023), would need to be mitigated via balanced pedagogical integration. Moreover, we did not independently test the separate contributions of AI vs LOA – a future 2×2 design (AI vs no AI crossed with LOA vs no LOA) could tease apart these effects. It would also be interesting to test whether particular learners benefit more (e.g., lower versus higher proficiency) or how motivation mediates outcomes. Lastly, while the VKS provided a consistent and extensive measure of both receptive and productive vocabulary knowledge, it mainly captures short-term gains. Future research may be enhanced via the inclusion of delayed post-tests in order to gauge long-term retention and lexical depth across prolonged periods.

7. Conclusion

This study provides robust empirical support for the integration of AI-assisted LOA in vocabulary instruction within Iranian EFL contexts. The experimental group, which utilized AI scaffolding through Nearpod and chatbot interaction, significantly outperformed the control group in vocabulary acquisition, underscoring the pedagogical efficacy of combining adaptive feedback mechanisms with learner-centered assessment. These results affirm that AI-powered LOA environments can personalize vocabulary instruction, foster metacognitive engagement, and yield measurable learning gains.

These results highlight the value of combining real-time adaptive assistance with reflective assessment practices, reinforcing prior research that positions AI as a facilitator of deeper learning, learner autonomy, and scaffolded progression. Moreover, the successful implementation in an Iranian EFL context addresses a research gap concerning underrepresented educational environments, offering practical implications for technology integration in similar instructional settings. Looking ahead, sustained research is needed to examine long-term retention effects, learner variability, and the integration of such tools across proficiency levels. Equitable access, teacher training, and infrastructure readiness also warrant close attention to ensure successful deployment.

Therefore, we advocate for the inclusion of AI-scaffolded LOA frameworks in curriculum development and teacher training programs, especially in underrepresented EFL settings. Future research should investigate longitudinal outcomes, scalability in diverse sociotechnical environments, and the interplay between learner autonomy and

automated scaffolding. By doing so, the academic community can further refine the principles of learning-oriented assessment in alignment with the evolving landscape of intelligent educational technologies.

Acknowledgments

I would like to thank Dr. Touran Ahour for her valuable support and contributions to this study.

Conflict of Interest

The author declares that there is no conflict of interest regarding the publication of this paper.

References

Abdelhalim, S. M., & Alsehibany, R. (2025). Integrating ChatGPT for vocabulary learning and retention: A classroom-based study of Saudi EFL learners. *Language Learning & Technology, 29*(1), 1–24. https://doi.org/10.64152/10125/73635

Aji, K., Ardin, H., & Arifin, M. A. (2020). Blended learning during pandemic coronavirus: Teachers' and students' perceptions. *Journal of Language Teaching and Learning, Linguistics and Literature, 8*(2), 632–646. https://doi.org/10.24256/ideas.v8i2.1696

Alazemi, A. F. T. (2024). Formative assessment in artificial integrated instruction: Delving into the effects on reading comprehension progress, online academic enjoyment, personal best goals, and academic mindfulness. *Language Testing in Asia*, *14*(1), 1–44. https://doi.org/10.1186/s40468-024-00319-8

Alharbi, K., & Khalil, L. (2023). Artificial intelligence in ESL vocabulary learning: An exploratory study on students and teachers' perspectives. *Migration Letters*, *20*(12), 1030–1045. Retrieved from https://migrationletters.com/index.php/ml/article/view/8224

Al-Shehri, M. B. A. (2018). The effectiveness of using the Nearpod application in tablets on the classroom interaction of computer subject among the female students of the second grade of high school. *Specialized International Educational Journal, 7*(1), 129–140.

Aulia, R., Sunra, L., & Nani, F. (2024). The influence of using Nearpod on the students' enthusiasm in classroom discussion in Class X English at SMAN 2 Pangkep. *Journal of English Education and Literature, 3*(2), 317–325. https://doi.org/10.26858/performance.v3i2.62016

Awadelkarim, A. A. (2021). An analysis and insight into the effectiveness of scaffolding: EFL instructors'/teachers' perceptions and attitudes. *Journal of Language and Linguistic Studies, 17*(2), 828–841. https://search.informit.org/doi/10.3316/informit.216267229830245

Balqis, N., & Zaki, L. B. (2025). Classroom action research: Improving young learners' vocabulary using Nearpod. *Esensi Pendidikan Inspiratif*, *7*(1), 1–18.

Black, P., & Wiliam D. (1998). *Inside the black box: Raising standards through classroom assessment.* Phi Delta Kappan.

Dave, A. (2004). *Oxford placement test 2: Test pack*. Oxford University Press.

Dujardin, E., Auphan, P., Bailloud, N., Ecalle, J., & Magnan, A. (2021). Tools and teaching strategies for vocabulary assessment and instruction: A review. *Social Education Research*, *1*(1), 34–66. https://dx.doi.org/10.37256/ser.3120221044

Hago, E. O., Ahmed Saeed Balla, A., Hussein Abdelrady, A., Othman , E., & Othman Abdelaziz Ahmed, A. (2025). AI-driven vocabulary acquisition in EFL higher education: Interdisciplinary insights into technological innovation, ethical challenges, and equitable access. *Forum for Linguistic Studies*, *7*(4), 477–491. https://doi.org/10.30564/fls.v7i4.8760

Hakami, M. (2020). Using Nearpod as a tool to promote active learning in higher education in a beyond learning environment. *Journal of Education and Learning, 9*(1), 119–126. https://doi.org/10.5539/jel.v9n1p119

Hernandez-Mena, C., Amado-Moranchel, N., Avalos, M., Alvarez, J., & Faccinetto-Beltrán, P. (2024). Improving student motivation in higher education through Nearpod. *IEEE Global Engineering Education Conference (EDUCON)*, Kos Island, Greece, 1–4. https://doi.org/10.1109/EDUCON60312.2024.10578581

Inam, S., Jawaid, M., & Khan, R. A. (2023). Assessment of workplace-related factors affecting tolerance of ambiguity among trainee doctors. *The Journal of the Pakistan Medical Association*, *73*(9), 1827–1832. https://doi.org/10.47391/JPMA.7604

Jamali Kivi, P., Namaziandost, E., Fakhri Alamdari, E., Ryafikovna Saenko, N., Inga-Arias, M., Fuster-Guillén, D., ... & Nasirin, C. (2021). The comparative effects of teacher versus peer-scaffolding on EFL learners' incidental vocabulary learning and reading comprehension: A socio-cultural perspective. *Journal of Psycholinguistic Research*, *50*(5), 1031–1047. https://doi.org/10.1007/s10936-021-09800-4

Jiang, R. (2022). How does artificial intelligence empower EFL teaching and learning nowadays? A review of artificial intelligence in the EFL context. *Frontiers in Psychology*, *13*(1), 49–61. https://doi.org/10.3389/fpsyg.2022.1049401

Li, M., & Wilson, J. (2025). AI-integrated scaffolding to enhance agency and creativity in K-12 English language learners: A systematic review. *Information*, *16*(7), 5–19. https://doi.org/10.3390/info16070519

Mekota, T., & Marada, M. (2020). The influence of the Nearpod application on learning social geography in a grammar school in Czechia. *Education and Information Technologies. 12*(2), 46–66. https://doi.org/10.1007/s10639-020-10214-3

Musa, M. A., & Momani, J. A. (2022). University students' attitudes towards using the Nearpod application in distance learning. *Journal of Education and e-Learning research, 9*(2), 110–118. https://doi.org/10.20448/jeelr.v9i2.4030

Nation, I. S. P. (1990). *Teaching and Learning Vocabulary*. Newbury House.

Putra, A. P., Arafik, M., & Pratiwi, I. (2021). Use of Nearpod to enhance student engagement in online learning. *2021 7th International Conference on Education and Technology (ICET)*. https://doi.org/10.1109/icet53279.2021.9575062

Sanchez, L. M., & Carballo, Y. A. (2025). The impact of implementing the CALL approach using the Nearpod platform in improving phrasal verbs vocabulary among adult learners at the virtual institute Centro de Matemáticas e Idiomas Segura in San Isidro de Alaju. *Ciencia Latina Revista Científica Multidisciplinar*, *9*(2), 1060–1078. https://doi.org/10.37811/cl_rcm.v9i2.16932

Seyed, F. S., & Tavassoli, K. (2023). The impact of learning-oriented assessment on EFL learners' vocabulary learning and retention in online classes. *JELT Journal*|

Farhangian University, *2*(1), 84–99. https://doi.org/10.22034/jelt.2023.13448.1043

Shamshiri, F., Esfahani, F.R., & Hosseini, S.E. (2023). Models of assessment in the classroom: Comparative research of CALL-based vs. traditional assessment on vocabulary learning among Iranian EFL learners. *Language Testing in Asia, 13*(1), 1–22. https://doi.org/10.1186/s40468-023-00259-9

Simina, V. K. (2012). Socio-constructivist models of learning. In *Encyclopedia of the Sciences of Learning* (pp. 3128-3131). Springer, Boston, MA.

Tiansoodeenon, M., Meeporm, B., Kaewrattanapat, N., & Tarapond, S. (2023). Enhancing vocabulary acquisition through progressive word increments in English language learning. *Journal of Liberal Arts RMUTT*, *4*(2), 88–100. https://doi.org/10.60101/jla.2023.4.2.3172

Tran, M., Balasooriya, C., & Semmler, C. (2025). Generative artificial intelligence: The 'more knowledgeable other' in a social constructivist framework of medical education. *npj Digit. Med, 8*(1), 43-60. https://doi.org/10.1038/s41746-025-01823-8

Van Le, H. H., & Doan, T. K. O. (2023). EFL students' perceptions of using NearPod in online English learning. *ICTE Conference Proceedings*, *3*, 98–117. https://doi.org/10.54855/ictep.2338

Vygotsky, L. S., & Cole, M. (1978). *Mind in society: Development of higher psychological processes*. Harvard University Press.

Wang, Y., Wu, J., Chen, F., Wang, Z., Li, J., & Wang, L. (2024). Empirical assessment of AI-powered tools for vocabulary acquisition in EFL instruction. *IEEE Access, 12*(1), 131892–131905. https://doi.org/10.1109/ACCESS.2024.3446657

Wesche, M., & Paribakht, T. S. (1996). Assessing second language vocabulary knowledge: Depth versus breadth. *Canadian Modern Language Review*, *53*(1), 13–40.

Xu, Y., & Warschauer, M. (2020, April). What are you talking to?: Understanding children's perceptions of conversational agents. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-13). https://doi.org/10.1145/3313831.3376416

Yang, X. (2024). Enhancing online TESOL education: Strategies for effective pre-class, in-class, and post-class learning in the post-pandemic era. *Journal of Educational Research and Policies, 6*(9), 92-98. https://doi.org/10.53469/jerp.2024.06(09).19

Zahran, F. A. (2025). The effect of utilizing Nearpod with a guided reading strategy on EFL primary pupils' reading comprehension skills and motivation. *Journal of Educational Research and Practice, 15*(1), 1–12. https://doi.org/10.5590/JERAP.2025.15.1976