



ISSN 2185-3762

Studies in Self-Access Learning Journal
<http://sisaljournal.org>

AI-Driven Vocabulary Learning: The Role of NotebookLM AI in Supporting Retention and Self-Directed Learning

Ali Al Ghaithi, Faculty of Language Studies, Sohar University, Oman

Behnam Behforouz, Department of Foreign Languages, College of Arts and Sciences, University of Nizwa, Oman

Corresponding email address:
AGhaithi@su.edu.om

Publication date: September, 2026

To cite this article

Al Ghaithi, A., & Behforouz, B. (2026). AI-driven vocabulary learning: The role of notebookLM AI in supporting retention and self-directed learning. *Studies in Self-Access Learning Journal*, 17(3), 464–481. <https://doi.org/10.37237/170310>

This article may be used for research, teaching, and private study purposes. Please contact the authors for permission to reprint elsewhere.

Scroll down for article.

AI-Driven Vocabulary Learning: The Role of NotebookLM AI in Supporting Retention and Self-Directed Learning

Ali Al Ghaithi, Faculty of Language Studies, Sohar University, Oman.

<https://orcid.org/0000-0002-9653-508X>

Behnam Behforouz, Department of Foreign Languages, College of Arts and Sciences, University of Nizwa, Oman. <https://orcid.org/0000-0002-0078-2757>

Abstract

The present study aimed at investigating the role of using advanced Artificial Intelligence (AI) technologies, particularly NotebookLM AI, vocabulary knowledge, and self-directed learning skills of students. As self-directed learning is a crucial skill for learners to plan, manage, and assess their own learning, the survey is also about its relevance to self-access learning contexts, where learners are entitled to use learning resources independently. Therefore, 70 Omani pre-intermediate English proficiency students were divided into two groups, with 35 candidates, including a control group receiving face-to-face instruction and an experimental group using NotebookLM AI as a main source of learning. To collect the required data, three sets of researcher-made tests were developed, accompanied by a self-directed learning questionnaire. The analysis of the data showed that both groups initially had an increase in scores in the posttest, but the experimental group outperformed the control group. In the delayed posttest, while the scores of the control group decreased, the experimental group showed significantly better performance. In the self-directed learning questionnaire, the experimental group showed statistically significant performance. The outcomes of the study imply that the use of AI in learning can not only augment vocabulary learning and self-regulated learning but also increase students' preparedness to cooperate more efficiently in self-access learning environments. The results of the study are helpful for self-access practitioners and users.

Keywords: Artificial Intelligence, NotebookLM AI, vocabulary, self-directed learning

Vocabulary knowledge is vital for second language learners, and it has been defined as the ability to recognize a particular word and construct its meaning (Dalimunthe & Haryadi, 2022). While vocabulary eases learning a foreign language as one of the most important elements in correctly understanding and using it, it is one of the aspects of language in which a considerable number of mistakes are made by students (Segler et al., 2002). Nation (2013) stated that the students' lack of vocabulary knowledge obstructs them in mastering the language. Recognizing these obstacles may interfere with learning (Anggara, 2023). ICT-based techniques in education have been proven to motivate students to a great extent and could be considered as one of the solutions to overcome vocabulary learning

issues. Technology-supported narrative devices encourage learners to make sense of educational concepts and to articulate their opinions within and outside the classroom (Dupain & Maguire, 2005).

Self-directed learning (SDL) is considered a process whereby learners take personal responsibility for their learning. Garrison (1997) defines SDL as a process in which learners are motivated to assume personal responsibility and collaborative control of cognitive (self-monitoring) and contextual (self-management) processes in constructing learning outcomes. Self-monitoring refers to the process in which a learner monitors their own thinking and learning. It implies cognitive and metacognitive processes. Self-management is understood as the management of one's learning resources, time, and learning environment to influence the learning task (Doo & Zhu, 2023). Language learning apps such as Duolingo and Busuu have been found to benefit self-directed language learners by providing flexibility and accessibility that support independent language study (Li & Bonk, 2023). This interconnection plays a significant role in self-access language learning, in a setting where learners are provided with materials, activities, and support systems independent of teachers and take charge of content, pace, and direction of their learning (Gardner & Miller, 1999). In this way, the digital tools that instructors familiarize students with in the traditional classroom may also serve as autonomous access points to the students when they resume using them by themselves, for example, to review vocabulary, clarify meanings, take quizzes, and self-track their progress outside of the classroom (Reinders & Benson, 2017). SDL has also been studied in online education, including MOOCs and other OERs, where learners are given more autonomy over their learning (Zhu & Bonk, 2022).

Using generative AI in SDL is a developing area of research, and research studies on its relevant empirical applications are scarce (Lin, 2023). Among the limited number of existing empirical studies, Lin (2023) examined how ChatGPT can be used as a virtual tutor to provide specific learning objectives, materials, personalized feedback, and guidance for autonomous adult learners within asynchronous online contexts. Concerns have been raised about the risk of students accessing inaccurate sources and becoming overly reliant on AI, instead of critically evaluating sources. Similarly, Mogavi et al. (2024) continued this with another qualitative content analysis of 1500 posts within popular social media platforms and found that early adopters of ChatGPT for educational purposes generally highlighted its effectiveness in providing personalized feedback to learners and in creating customized learning objectives and lesson plans. However, some concerns about academic dishonesty, superficial learning, and negative effects on learners' critical thinking were reported. The

dichotomy shows that generative AI may help improve SDL but may also provide shortcuts for learners to avoid cognitive processes that are necessary for SDL. Although generative AI tools can be adapted to accommodate individuals' learning styles, needs, approaches, and goals, it is pointed out that it is important to maintain a high level of cognitive engagement and self-assessment so that learners take responsibility for evaluating and reflecting on their own learning processes and products (Ali et al., 2023).

There is still a dearth of empirical data on modern English language teaching methods due to uncertainty regarding the possibilities of technology in the field of language acquisition (Kim, 2020). A limitation of AI systems for education is that they are often more effective and usable for learners with high levels of language skill, which may make them unsuitable for learners with low language expertise (Fryer & Carpenter, 2006). Since NotebookLM AI provides a personalized, context-aware learning experience based on learner-generated materials uploaded to the tool, it was confidently selected for this study. With the provision of personalized explanations, summaries, examples, follow-up questions, and additional learner-generated material, NotebookLM AI reinforces meaningful processing, retrieval practice, and repeated exposure, contributing to long-term retention of vocabulary. Additionally, by offering adaptive feedback, multimodal delivery, and the interactive note creation process, NotebookLM AI prompts students to engage as self-directed learners, metacognitively and cognitively. The originality of this study lies in its empirical application of NotebookLM AI to an EFL context, showing the potential of an AI-supported approach to improve learners' vocabulary acquisition and self-directed learning capabilities, contributing to the practical application of generative AI in personalized language learning environments. Therefore, this study implemented NotebookLM AI in the English learning context among Omani EFL learners to measure its effect on the immediate and long-term vocabulary knowledge, as well as their skills in the self-directed learning process. The following questions will be addressed in this study:

1. Does NotebookLM AI affect vocabulary learning and retention in Omani pre-intermediate EFL learners?
2. Does NotebookLM AI affect the self-directed learning skills of Omani pre-intermediate EFL learners?

Literature Review

Artificial Intelligence and Vocabulary Retention

NG et al. (2026) tested the effectiveness of AI Copilot for learning new words on 90 pre-intermediate Omani learners, who were divided into three groups. The first group (A) used the AI Copilot in the classroom environment to learn the terms, the second group (B) used AI Copilot out of class to learn the terms, and the third group was a control group. Results showed that the first group (A) retained more vocabulary than the second group (B), and the second group (B) retained more vocabulary than the control group. The strengths of the study included three treatment groups and a delayed posttest, but the limitations included a small sample size, short treatment time, and a lack of qualitative data.

In another quasi-experimental mixed-methods study, Abdelhalim and Alsehibany (2025) intended to investigate the potential effects of using ChatGPT on the vocabulary acquisition of 71 EFL learners. There were two groups of participants. The experimental group used ChatGPT as an add-on to customary classroom instruction with real-time feedback, while the control group only received classroom instruction. Participants in the treatment group outperformed the control group in retention of productive vocabulary knowledge. The strengths include the mixed-methods design and delayed posttest. It is, however, limited by its sample size and the focus on Saudi students.

Similarly, Taj et al. (2025) studied how AI can be used to address vocabulary learning challenges that English language learners face through a quantitative survey study with 372 participants in one group. The participants used AI for practicing pronunciation, receiving feedback, and tracking vocabulary retention. Results showed that 64.5% of participants reported that they used AI for vocabulary retention, pronunciation, and thematically based vocabulary learning using a variety of mediums, such as AI-powered flashcards, chatbots, and adaptive learning systems. A large sample size and validated instruments represent strengths of the study. Limitations include a young adult sample, self-reported data, and the lack of experimental design.

Rajayi et al. (2025) intended to investigate the impact of AI-based instruction on the learning and retention of vocabulary among ESP students. A quasi-experimental study with two groups of 50 pre-intermediate accounting students used the Memrise application to teach 50 technical accounting words to the treatment group in 10 sessions. In the control group, ordinary instruction was used. Learners who used Memrise performed better than the control group in vocabulary acquisition and retention. Strengths of this study include its experimental

design and the use of a delayed posttest; limitations include the short duration of the intervention, the focus on a single field, and single-university sampling in Tehran.

AI-Enhanced Self-Directed Learning

Li et al. (2025) explored the effectiveness of ChatGPT-integrated instruction for teaching research skills to 366 undergraduate college education majors who were randomly assigned to either a ChatGPT-integrated treatment group or a non-integrated control group. Participants in the experimental group were trained on prompt engineering for ChatGPT for research purposes, while the control group received conversational instruction. The study found that the group that used the AI tool demonstrated higher levels of research skills, autonomous motivation, and self-directed learning than the control group. Limitations of this study include the use of self-report questionnaires and the fact that the experimental group was selected only from one university in Pakistan. Despite these limitations, the study had a larger sample size and an empirical research design.

Another study by Umar and Purwanto (2025) investigated students' perspectives toward AI in decision support systems for self-directed learning. Thirty-six students in one group used AI technologies, such as adaptive learning systems (ALS), recommendation systems (RS), and smart tutoring systems (ITSs), to assist them with decision-making. The study found that students improved their use of learning resources, self-directed learning, cognitive load, and retention. The strength of the study included qualitative and data triangulation methods. The study's limitations include a lack of empirical methodology, its heavy reliance on AI, and its lack of emotional cognition.

Likewise, Behforouz and Al Ghaithi (2024) studied the effect of an interactive artificial intelligence (AI) chatbot on 50 Omani EFL learners' SDL. Students were randomly assigned to an experimental group that used a WhatsApp chatbot to learn vocabulary, while the control group received traditional instructions. The treatment group improved their self-directed learning and task performance more than the control group. Strengths of the study are the use of an experimental design and validated measures, while its limitations are the small number of participants and the single context in Oman.

Indriani et al. (2024) explored whether ChatGPT use and learning motivation influenced self-directed learning. The study involved a single group of 98 Indonesian undergraduate and master's students who had used ChatGPT for learning. Both ChatGPT use and learning motivation were found to influence self-directed learning considerably and positively, with learning motivation having a stronger impact. Strengths include the high

reliability and validity of the measures. Limitations include a correlational study design, convenience sampling, and the use of Indonesian participants only.

Method

Participants

The sample of the study consisted of 70 Omani learners enrolled in the foundation program, at the pre-intermediate English proficiency level as determined by the placement test and the results of the college academic assessment. The participants were native Arabic speakers between the ages of 19 and 21. Both male and female participants were included in the classes. The participants in the study were divided into two groups: a control group and an experimental group, consisting of 35 students each. The control group was taught in conventional classroom training without using any AI tools. Meanwhile, the experimental group underwent the same training, but they used NotebookLM AI as a learning facilitator.

Instruments

Self-Rating Scale of Self-Directed Learning (SRSSDL)

The SRSSDL, developed by Williamson (2007), was used in this study to assess the development of students' SDL skills. There are five general areas of SDL, each containing 12 components: awareness, learning strategies, learning activities, evaluation, and interpersonal skills. The 60-item scale is divided into five subdomains. In scoring each item, a 'never' response was reflected by a score of one, and an 'always' response received a score of five. A maximum score of 300 and a minimum score of 60 could be obtained for the SRSSDL. Subscales measuring learning, evaluation, interpersonal, and learning activities demonstrated good reliability, with the SRSSDL scale achieving overall Cronbach's alpha coefficients of 0.87, 0.93, 0.91, and 0.88 (Salem, 2022). The Arabic translation of the statements, which were adopted from a study conducted by Behforouz and Ali Al Ghaithi (2024), accompanied the English version.

Vocabulary Test

The researchers of the study developed three different vocabulary tests to be used as pretests, posttests, and delayed posttests. There were 20 items with a total score of 20 marks in each test, including five fill-in-the-blank, five matching the words with their definitions, five multiple-choice, and five cloze tests. Each correct response received 1 point, and there was no penalty for wrong answers. The words were selected from the materials that were designed by the college to be covered during the current semester. The same test was

conducted for each group in either phase of the study. Before conducting each test, the reliability index of these tests was measured using a pilot study with 30 Omani pre-intermediate EFL learners from the same university who were not in the control group or in the experimental one. The pretest was piloted before the treatment, the posttest immediately after the treatment, and the delayed posttest in the following week. All these tests were piloted before the main test distribution. Table 1 shows that the Cronbach's Alpha for the pretest, posttest, and delayed posttest was .850, .790, and .820, respectively, showing that the items are highly reliable. Additionally, the questions were reviewed by an Omani Applied Linguist with 15 years of teaching experience within the same Omani EFL learning context.

Table 1

The Reliability Index for the Pretest, Posttest, and the Delayed Posttest

Test	Cronbach's Alpha	N of Items	N
Pretest	.850	20	30
Posttest	.790	20	30
Delayed posttest	.820	20	30

Procedures

The Research Ethics Committee approved the study; all participants signed the consent forms, and it was conducted in the first semester of the academic year 2025-2026 in a university in Oman for 6 weeks. In week 1, the students in both groups were randomly divided equally into an experimental group and a control group. All students learned 80 target words over a period of four weeks. The lesson plan focused on a set of 20 words each week, with five words presented each of four days and a fifth day reserved for teacher-created activities. Before the start of the treatment, all participants took a pretest on vocabulary knowledge and completed a self-directed learning questionnaire to check for group equivalence.

The experimental group received a session on using Google NotebookLM AI for vocabulary acquisition. Then, the experimental group was split into seven subgroups of five students. The learners deliberated the objective terms in their respective teams to guarantee that every participant comprehended their definitions and applications. The 20 weekly target words were sent to the students as PDFs via Microsoft Teams. Then, the learners used three key features of Google NotebookLM AI video generator, infographics, and quiz features to

get explanations for the words and practice the list. The students were able to use the NotebookLM AI platform to clarify what they did not understand, explain words they were not familiar with, and take vocabulary quizzes.

The video, infographic, clarification, and quiz interaction sequence were completed in 30 minutes in each session, while the teacher monitored group work. In the final 10 minutes of class, the teachers used a whole-class discussion elicited by the NotebookLM AI flashcard feature to check students' knowledge of the meanings of the words. The control group learned the same words from the teacher and peers. In the same way, the control students were subdivided into subgroups, and they talked about the weekly target words together to make sure they understood them well, but they carried out the vocabulary discussion and practice without the use of NotebookLM AI or other AI-supported features. This procedure was repeated each week, and both groups completed the vocabulary posttest and the self-directed learning questionnaire during the fifth week. In week 6, a delayed posttest for measurement of vocabulary retention was administered.

Data Analysis

This part describes the numerical data of the vocabulary tests and the self-directed learning questionnaire using SPSS 27.0. In the beginning, a normality test was conducted to determine the data distribution in the groups, and Table 2 shows the results of the Shapiro-Wilk Normality Test.

Table 2

The Results of Data Distribution in all Sets of Tests for Both Groups

	groups	Shapiro-Wilk		
		Statistic	df	Sig.
pretest	control	.971	35	.461
	experimental	.973	35	.527
posttest	control	.954	35	.148
	experimental	.923	35	.018
Delayed posttest	control	.954	35	.154
	experimental	.921	35	.016

Table 2 shows that the pretest scores were normally distributed for both the control group, $p = .461$, and the experimental group, $p = .527$. The control group also met the normality assumption at posttest, $p = .148$, and delayed posttest, $p = .154$. Conversely, the experimental group did not achieve the normality principle at posttest, $p = .018$, nor at the delayed posttest, $p = .016$. Hence, a parametric test was used to compare the performance of

the control group, and non-parametric tests were viewed as the more appropriate method for the comparisons concerning posttest and delayed posttest scores, mainly for the experimental group. Table 3 shows the performance of the control group in three tests.

Table 3

Comparison of the Control Group's Scores in All the Tests

	Paired Differences						Sig. (2-tailed)	
	Mean	Std. Dev.	Std. Error Mean	95% Confidence Interval of the Difference		t		
				Lower	Upper			
pretest - posttest	-4.028	4.176	.705	-5.463	-2.594	-5.707	34	.000
pretest – delayed posttest	-2.057	3.588	.606	-3.289	-.824	-3.391	34	.002
posttest – delayed posttest	1.971	3.966	.670	.608	3.334	2.940	34	.006

Table 3 shows that the difference from pretest to posttest was statistically significant, $p < .001$, with a mean difference of 4.03 points. The delayed posttest score was still considerably greater than the pretest score, $p = .002$, with a mean difference of 2.06. There was a small decrease in the performance from posttest to delayed posttest, $p = .006$, with a mean difference of 1.97. To compare the performance of the participants within the experimental group, a Wilcoxon Signed Rank Test was conducted, and the data are presented in Table 4.

Table 4

The Comparison of the Experimental Group's Scores in all the Tests

	posttest - pretest	delayed posttest - pretest	delayed posttest - posttest
Z	-5.134	-5.165	-3.303
Asymp. Sig. (2-tailed)	.000	.000	.001

The results in Table 4 show that the scores increased from pretest to posttest ($p < .001$), and from pretest to delayed posttest ($p < .001$). There was a statistically meaningful difference between posttest and delayed posttest, $p = .001$. Finally, the performance of both groups was compared, and the results can be seen in Table 5.

Table 5*The Comparison of Both Groups' Performance in All Tests*

	pretest	posttest	Delayed posttest
Mann-Whitney U	491.500	111.500	6.500
Wilcoxon W	1121.500	741.500	636.500
Z	-1.429	-5.923	-7.145
Asymp. Sig. (2-tailed)	.153	.000	.000

The results in Table 5 show that at the pretest stage, the two groups had no statistically significant differences, $U = 491.500$, $Z = -1.429$, $p = .153$. However, a statistically significant difference was found in the posttest stage, $U = 111.500$, $Z = -5.923$, $p < .001$, which was evidence of the groups' significant difference after the implementation of the intervention. Likewise, the delayed posttest results were also different at a statistically significant level, $U = 6.500$, $Z = -7.145$, $p < .001$. The absence of the difference between the groups was therefore maintained over time. To know the amount of these differences between the groups, the effect sizes were measured, and the results are observable in Table 6.

Table 6*The Results of the Effect Size Between the Two Groups in All the Tests*

		Point Estimate	95% Confidence Interval	
			Lower	Upper
pretest	Eta-squared	.033	.000	.148
posttest	Eta-squared	.488	.313	.608
delayed posttest	Eta-squared	.763	.658	.821

The results in Table 6 show that in the pretest, both groups were equivalent prior to the treatment, while the posttest indicates that the experimental group outperformed the control group. The eta-squared statistic increased from a small effect size of .033 on the pretest to a large effect size of .488 on the posttest. The delayed posttest produced the same pattern (eta-squared = .763), with the experimental group still being more affected in comparison to the control group.

The second part of the statistical analysis investigated the self-directed language learning questionnaire. A mixed-design ANOVA was conducted to compare the performance of groups together based on the self-directed learning criterion. Table 7 shows the results of the descriptive data.

Table 7

The Descriptive Analysis of the SRSSDL Questionnaire in the Pretest and Posttest of Both Groups

	groups	Mean	Std. Deviation	N
pretest	control	119.89	2.83	35
	experiment	121.00	2.71	35
	Total	120.45	2.80	70
posttest	control	120.89	2.82	35
	experiment	222.63	12.27	35
	Total	171.76	51.99	70

Table 7 shows that in the pretest, the control group scored an average of 119.89 (SD = 2.83), in contrast to the experimental group, scored 121.00 (SD = 2.71), showing similar initial self-directed learning scores. However, in the posttest, the control group, which had scored only a bit higher (from 119.89 to 120.89), thus underwent no change, while the experimental group increased from 121.00 up to 222.63 (SD = 12.27), which indicates that the intervention played a role in the experimental group's SDL, while the control group remained without almost any changes. Table 8 provides more details regarding the performance of both groups over time.

Table 8

The Results of Mixed-Design ANOVA for SRSSDL Questionnaire (All Participants)

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Time	Sphericity Assumed	92160.457	1	92160.457	3730.114	.000	.982
Time *	Sphericity Assumed	88603.457	1	88603.457	3586.147	.000	.981
groups							
Error (Time)	Sphericity Assumed	1680.086	68	24.707			

Table 8 indicates that the SDL scores of students experienced a considerable shift from the pretest to the posttest, $F(1, 68) = 3730.114$, $p < .001$, partial $\eta^2 = .982$. This consequently means that there was a clear overall increase in scores after the intervention period. Even more significantly, the Time $\times$ Group interaction was also statistically significant, $F(1, 68) = 3586.147$, $p < .001$, partial $\eta^2 = .981$, and it was such that the improvement was different for the control and the experimental groups. This means that the experimental group achieved a significantly higher score from pretest to posttest, while the control group showed only a marginal alteration. The large partial eta-squared values are consistent with the descriptive statistics in Table 7, where the posttest mean of the

experimental group surged while the control group remained nearly unchanged. To better understand the elements that were receiving more attention from the learners, the posttest of the experimental group was analyzed, and the results can be seen in Table 9.

Table 9

The Descriptive Analysis of the SRSSDL Questionnaire (Experimental Group)

	N	Minimum	Maximum	Mean	Std. Deviation
Awareness	35	41.00	52.00	46.60	2.92
Learning Strategies	35	44.00	51.00	47.17	1.93
Learning Activities	35	43.00	50.00	45.77	1.81
Evaluation	35	40.00	48.00	44.20	2.22
Interpersonal Skills	35	33.00	49.00	41.06	4.07
Valid N (listwise)	35				

Table 9 shows that all dimensions of self-directed learning were at the prominent level. The mean score for Awareness ($M = 46.60$) revealed that the learners had a developed awareness of learning needs and goals, were motivated to learn, and had a sense of personal responsibility for learning. The Learning Strategies had a mean ($M = 47.17$), which reflected that purposive strategies (group discussion, peer coaching, role-play, case-based learning, and use of interactive technologies) were frequently used, while Learning Activities were also highly rated ($M = 45.77$), indicating that the participants frequently employed some key learning behaviors (revising what they had learned, identifying the key points, concept mapping, asking relevant questions, using technology, linking theory to practice, and reflecting critically on new knowledge). Beyond that, students were strong in Evaluation ($M = 44.20$), within which were monitoring, self-assessment, giving, seeking, and using feedback, and reflective review. The Interpersonal Skills mean ($M = 41.06$) indicates that the learners communicated and collaborated, were open to the perspectives of others, and shared information with others, despite groups and cultures varying.

Discussion

The study is intended to explore the potential role of Google NotebookLM AI as a facilitator and its possible influence on vocabulary acquisition and student SDL. The findings suggested that the students in both groups initially improved their vocabulary learning, but the experimental group achieved higher scores than the control group. The delayed posttest results showed that the control group's score decreased, while the experimental group's score continued to increase in this study context. The self-directed learning questionnaire reported higher levels of SDL than the control group. Even though the group-based word discussions

might have played a role in the vocabulary development of the students in both groups, it seems that the experimental group outperformed the others mainly due to the provision of NotebookLM AI, which allowed the learners to explain their difficult words, reread the explanations, take practical tests, and manage their learning process with independence. However, this interpretation needs to be made cautiously as the study's sample size was limited and the intervention period was short.

These findings suggest that the use of learning context in conjunction with NotebookLM AI may have supported students' vocabulary retention and SDL behavior. The score of the posttest of the control group indicated that techniques of conventional teaching may help students master vocabulary immediately. Conversely, compared to the more interactive and personalized learning approaches, they might be less effective in long-term vocabulary retention. The fact that only the experimental group showed additional gains between the posttest and delayed posttest suggests that the benefits of NotebookLM AI may be partly related to the customized learning experience of adaptive learning, allowing students to engage with vocabulary more actively by addressing their learning needs. The video feature of NotebookLM AI may be accompanied by both visual and auditory channels, which in turn leads to dual coding support. The infographic feature may support memory through pictorial representation. Finally, the platform's interactive chat feature may have increased learners' motivation and interest in completing learning activities. Real-time feedback through the quiz generation feature, rehearsal, and review of the trial sentences likely were repeated multiple times, which could have created conditions for vocabulary retrieval in the long term. The AI system may also contribute to SDL by giving learners control over their study and helping them develop metacognitive skills such as planning, monitoring, and evaluating their learning.

The results of the study on vocabulary learning are consistent with the findings of Hoşer and Aydin (2026), who found that AI-based gamification in language learning was associated with better vocabulary learning outcomes than traditional methods. Also, in the study by Maghsoudi (2025), AI-helped vocabulary instruction was reported to support both short-term and long-term vocabulary retention. Likewise, the study by Namaziandost and Çakmak (2025) found a meaningful advantage of AI over traditional vocabulary instruction in vocabulary retention with superior long-term outcomes. Furthermore, Pournabi and Ahmadi (2025) provided evidence for the efficacy of AI chatbots in vocabulary acquisition through personalized vocabulary exercises and immediate feedback.

Furthermore, the findings on SDL agree with a systematic review by Younas et al. (2025), which concluded that AI tools positively impact SDL through personalization and interactivity. Likewise, Alshammari (2024) found that ChatGPT improved Saudi university students' self-directed learning and research skills during their studies. These findings have been replicated in other studies, such as Anh (2024), who reported that ChatGPT promoted self-regulated learning in the same way for undergraduate students. These findings are also supported by Navas Bonilla (2025), who observed that AI-assisted tools supported the long-term development of self-regulation and autonomous learning through feedback.

Conclusion

This research can serve as a great asset to self-access practitioners and self-access users. Self-access practitioners may use Google NotebookLM AI to support the learners not only beyond the teacher-led instruction but also through the integration of the video generation, infographic generation, interactive chat, quizzes, and feedback on AI. These stunning tools may help to give more time to individualized attention, recommend suitable vocabulary practice to learners, and promote more effective self-basic direction learning. For self-access users, the NotebookLM AI may provide the chance for independent vocabulary learning, such as reviewing vocabulary outside of the classroom, defining new words, getting instant help, doing quizzes, and studying at their own pace. In this way, AI-supported learning may have a role in vocabulary strength, long-term retention, and the development of autonomous learning. In addition, it is of great importance to discuss the potential involvement of peer collaboration in promoting students' self-directed learning. Since the learners worked in small groups, the discussion of target words might have been the means for them to understand the meanings, compare their ideas about the words, ask questions, and to be more active in the monitoring of their learning processes. This kind of interaction might be conducive to SDL by prompting the students to take responsibility for their learning, as well as for the development of shared group knowledge. Accordingly, although NotebookLM AI was most likely the extra support with explanations, quizzes, feedback, and independent review opportunities, I also think that collaborative group activities might be responsible for students' vocabulary development and SDL growth.

This study had some limitations to be addressed. A sample size of 70 students made generalizing the results difficult. Additionally, participants in the present study included only pre-intermediate Omani EFL learners, so the results could not be applied to learners at differing skill levels with diverse cultural or linguistic backgrounds. Another limitation of

this study was its focus on the quantitative analysis of data without quantitatively analyzing the statements in SRSSDL. Additionally, this study focused on one single skill, vocabulary, so the results cannot be applied to the other skills. Finally, there was no investigation into the role of gender and age in the performance of the students.

Future studies could improve generalizability by involving a larger sample and learners of various expertise and learning backgrounds. Future studies could also include both quantitative and qualitative investigations, especially at the item level in the SRSSDL instrument, to gain a greater understanding of the development of SDL behaviors in AI-supported environments. Future studies could focus on other language skills like reading, writing, speaking, and listening to determine the effect of NotebookLM AI on other aspects of language learning, beyond vocabulary learning. Future research could examine other demographic variables, such as gender and age, to examine vocabulary retention and self-directed learning among learners. Additionally, comparative research into other AI language-learning applications and multimodal learning environments might reveal what approach is most effective in promoting SDL.

Notes on the Contributors

Dr. Ali Al Ghaithi is an Assistant Professor of English in the Faculty of Language Studies at Sohar University, Oman. Ali earned his PhD in June 2026 from the Universidad Pública de Navarra in Spain, focusing on Applied Linguistics. Ali is interested in research studies that mainly implement Artificial Intelligence in teaching and learning processes.

Dr. Behnam Behforouz is an Assistant Professor of English in the Department of Foreign Languages at the University of Nizwa, Oman. He has been teaching English in various Iranian and Omani universities since 2009. His main areas of interest are TESOL, Applied Linguistics, and Educational Technologies.

References

- Abdelhalim, S. M., & Alsehibany, R. A. (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>
- Ali, F., Choy, D., Divaharan, S., Tay, H., & Chen, H. (2023). Supporting self-directed learning and self-assessment using TeacherGAIA, a generative AI chatbot application: Learning approaches and prompt engineering. *Learning: Research and Practice*, 9(2), 135–147. <https://doi.org/10.1080/23735082.2023.2258886>

- Alshammari, J. (2024). Revolutionizing EFL learning through ChatGPT: A qualitative study. *Amazonia Investiga*, 13(82), 208–221. <https://doi.org/10.34069/AI/2024.82.10.17>
- Anggara, S. D. (2023). Improving vocabulary mastery of junior high school students by watching digital storytelling. *Prosodi*, 17(1), 109–119.
- Anh, N. Q. (2024). The impact of ChatGPT on students. *Vietnam Social Sciences*, 4(221), 92–108. [https://doi.org/1056749/VSSR.4\(221\).92-108](https://doi.org/1056749/VSSR.4(221).92-108)
- Behforouz, B., & Al Ghaithi, A. (2024). The impact of using interactive chatbots on self-directed learning. *Studies in Self-Access Learning Journal*, 15(3). <https://doi.org/10.37237/150302>
- Dalimunthe, L., & Haryadi, R. N. (2022). The effect of learning methods and vocabulary mastery on English speaking ability. *Lingua Educationist: International Journal of Language Education*, 1(1), 1–7. <https://doi.org/10.54099/le.v1i1.58>
- Doo, M. Y., & Zhu, M. (2023). A meta-analysis of effects of self-directed learning in online learning environments. *Journal of Computer Assisted Learning*. Advance online publication. <https://doi.org/10.1111/jcal.12865>
- Dupain, M., & Maguire, L. (2005). Digital storybook projects 101: How to create and implement digital storytelling into your curriculum. In *21st Annual Conference on Distance Teaching and Learning* (p. 6).
- Gardner, D., & Miller, L. (1999). *Establishing self-access: From theory to practice*. Cambridge University Press.
- Garrison, D. R. (1997). Self-directed learning: Toward a comprehensive model. *Adult Education Quarterly*, 48(1), 18–33. <https://doi.org/10.1177/074171369704800103>
- Hoşer, S., & Aydin, S. (2026). The impact of AI-enhanced gamification on vocabulary learning in a foreign language context. In *AI-powered solutions for bilingual proficiency and communication* (pp. 249–278). IGI Global Scientific Publishing.
- Indriani, W., Nawwaf, M. N., Yundianto, D., Erikha, F., & Khatami, M. (2024). From conversation to competence: The influence of ChatGPT use and learning motivation in improving self-directed learning. *Academic Journal of Psychology and Counseling*, 5(2), 202–225. <https://doi.org/10.22515/ajpc.v5i2.8971>
- Li, Z., & Bonk, C. J. (2023). Self-directed language learning with Duolingo in an out-of-class context. *Computer Assisted Language Learning*, 36, 1–23. <https://doi.org/10.1080/09588221.2023.2206874>
- Li, Y., Sadiq, G., Qambar, G., & Zheng, P. (2025). The impact of students' use of ChatGPT on their research skills: The mediating effects of autonomous motivation, engagement, and

- self-directed learning. *Education and Information Technologies*, 30(4), 4185–4216.
<https://doi.org/10.1007/s10639-024-12981-9>
- Lin, X. (2023). Exploring the role of ChatGPT as a facilitator for motivating self-directed learning among adult learners. *Adult Learning*. Advance online publication.
<https://doi.org/10.1177/10451595231184928>
- Maghsoudi, N. (2025). Leveraging artificial intelligence for vocabulary development: Effects on recall and retention in English for specific purposes contexts. *Research Article*, 6(2), 53–76. <https://doi.org/10.71703/CURE.2025.1209154>
- Mogavi, R., Deng, C., Kim, J., Zhou, P., Kwon, Y. D., Metwally, A. H. S., Tlili, A., Bassanelli, S., Bucchiarone, A., Gujar, S., Nacke, L. E., & Hui, P. (2024). ChatGPT in education: A blessing or a curse? A qualitative study exploring early adopters' utilization and perceptions. *Computers in Human Behavior: Artificial Humans*, 2(1), Article 100027. <https://doi.org/10.1016/j.chbah.2023.100027>
- Namaziandost, E., & Çakmak, F. (2025). Impact of AI-generated storytelling vs. gamified learning on vocabulary retention and engagement in CALL environments. *Computers and Education: Artificial Intelligence*, 9. <https://doi.org/10.1016/j.caeai.2025.100505>
- Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9781139858656>
- Navas Bonilla, C. del R., Viñan Carrasco, L. M., Gaibor Pupiales, J. C., & Murillo Noriega, D. E. (2025). The future of education: A systematic literature review of self-directed learning with ai. *Future Internet*, 17(8), 1–22. <https://doi.org/10.3390/fi17080366>
- NG, M. L., Al Ghaithi, A., & Behforouz, B. (2026). Digital scaffolding as learning opportunities: Enhancing vocabulary knowledge through Copilot AI. *Educational Process: International Journal*, 20, 1–16. <https://doi.org/10.22521/edupij.2026.20.12>
- Pournabi, M., & Ahmadi, S. (2025). The effect of an artificial intelligence chatbot on vocabulary retention by Iranian intermediate EFL Learners: A mixed methods approach. *Journal of Mixed Methods Studies in English Language Teaching*, 1(4), 128–147. <https://doi.org/10.71873/mslt.2025.1205641>
- Rajayi, S., Piri Damagh, H., & Saed, M. (2025). The effect of teaching vocabulary through artificial intelligence (AI) on ESP students' vocabulary learning and retention with focusing on accounting students. *Journal of English for Specific Purposes Pedagogy*, 2(1), 40–53. <https://doi.org/10.22034/jespp.2025.513535.1000>
- Reinders, H., & Benson, P. (2017). Research agenda: Language learning beyond the classroom. *Language Teaching*, 50(4), 561–578.

<https://doi.org/10.1017/S0261444817000192>

Salem, A. A. M. S. (2022). Multimedia presentations through digital storytelling for sustainable development of EFL learners' argumentative writing skills, self-directed learning skills and learner autonomy. *Frontiers in Education*, 7, 884709.

<https://doi.org/10.3389/feduc.2022.884709>

Segler, T. M., Pain, H., & Sorace, A. (2002). Second language vocabulary acquisition and learning strategies in ICALL environments. *Computer Assisted Language Learning*, 15(4), 409–422. <https://doi.org/10.1076/call.15.4.409.8272>

Umar, U., & Purwanto, M. B. (2025). AI and decision assistance for enhancing self-directed learning. *Eternal: English Teaching Journal*, 16(2), 457–465.

<https://doi.org/10.26877/eternal.v16i2.1524>

Williamson, S. N. (2007). Development of a self-rating scale of self-directed learning. *Nurse Researcher*, 14, 66–83.

<http://dx.doi.org/10.7748/nr2007.01.14.2.66.c6022>

Younas, M., Abdel Salam El-Dakhs, D., & Jiang, Y. (2025). A comprehensive systematic review of AI-driven approaches to self-directed learning. *IEEE Access*, 13, 38387–38403. <https://doi.org/10.1109/ACCESS.2025.3546319>

Zhu, M., & Bonk, C. J. (2022). Guidelines and strategies for fostering and enhancing self-directed online learning. *Open Learning: The Journal of Open, Distance and e-Learning*, 1–17. <https://doi.org/10.1080/02680513.2022.2141105>