Interactive Learning Environments

ISSN: 1049-4820 (Print) 1744-5191 (Online) Journal homepage: www.tandfonline.com/journals/nile20

The effects of artificial intelligence-based interactive scaffolding on secondary students' speaking performance, goal setting, self-evaluation, and motivation in informal digital learning of English

Feifei Wang, Xiaohua Zhou, Kangxin Li, Alan C. K. Cheung & Ming Tian

To cite this article: Feifei Wang, Xiaohua Zhou, Kangxin Li, Alan C. K. Cheung & Ming Tian (2025) The effects of artificial intelligence-based interactive scaffolding on secondary students' speaking performance, goal setting, self-evaluation, and motivation in informal digital learning of English, Interactive Learning Environments, 33:7, 4633-4652, DOI: 10.1080/10494820.2025.2470319

To link to this article: https://doi.org/10.1080/10494820.2025.2470319

INTERACTIVE LEARNING ENVIRONMENTS
2025, VOL. 33, NO. 7, 4633–4652
https://doi.org/10.1080/10494820.2025.2470319

The effects of artificial intelligence-based interactive scaffolding on secondary students' speaking performance, goal setting, self-evaluation, and motivation in informal digital learning of English

Feifei Wang[a], Xiaohua Zhou [b*], Kangxin Li[c], Alan C. K. Cheung [d] and Ming Tian[e]

[a]Centre for Information Technology in Education, The University of Hong Kong, Hong Kong, People's Republic of China; [b]Department of Applied Social Sciences, The Hong Kong Polytechnic University, Hong Kong, People's Republic of China; [c]Department of Curriculum and Instruction, The Chinese University of Hong Kong, Hong Kong, People's Republic of China; [d]Department of Educational Administration and Policy, The Chinese University of Hong Kong, Hong Kong, People's Republic of China; [e]Chenzhou Jinhai School, Chenzhou, People's Republic of China

ABSTRACT

Artificial intelligence-based interactive scaffolding (AIIS), which is the interactive process that learners communicate with AI to learn by implementing learning activities with AI's gradual assistance, has recently been reported to provide potential benefits to teachers and learners in informal digital learning of English (IDLE). While much previous research has focused on the combined effects of AIIS in facilitating interactions with teachers, learner peers, and content in higher education, this study specifically examined the discrete effects of AIIS on secondary students' speaking performance, goal setting, self-evaluation, and motivation, and the developmental trend of goal setting, self-evaluation, and motivation in IDLE. Sixty secondary students were assigned into the treatment group of interacting with AIIS and the control group of observing fixed scaffolding over ten sessions in IDLE. Results indicated that the treatment group showed more learning gains in vocabulary, goal setting, self-evaluation, and motivation than the control group. Moreover, despite the similar decline and rise developmental trajectories of goal setting, self-evaluation, and motivation, the treatment group dropped less in the middle of the sessions. This study adds to the scarce evidence that AIIS facilitates secondary students' speaking learning in IDLE and provides theoretical, pedagogical, and research implications for AI-supported interactive learning environments.

ARTICLE HISTORY
Received 10 December 2023
Accepted 15 February 2025

KEYWORDS
Interactive scaffolding; language learning; goal setting; self-evaluation; motivation; human-AI interaction

As the capacities of artificial intelligence (AI) are growing, AI accompanies the present generation of children to grow up. The proliferation of AI conversational agents (e.g. Google Assistant, Amazon Alexa, and Apple Siri) embedded in mobile technologies has facilitated the integration of AI into our lives. According to National Public Radio (2022), 35% of Americans own a smart home device, and this equals 100 million adults with an average of more than two devices per home, of which 49% acknowledge that the reason for wanting a device in media access is to be used by children. Analysts predict that the revenue of smart home devices worldwide will reach a new peak in

CONTACT Alan C. K. Cheung alancheung@cuhk.edu.hk
Supplemental data for this article can be accessed online at https://doi.org/10.1080/10494820.2025.2470319.
*Present address: Teaching and Learning Centre, Lingnan University, Hong Kong, People's Republic of China.

2029, with 29.65 billion U.S. dollars (Laricchia, 2024). Given this explosive global growth, AI conversational agents, making use of speech recognition and natural language processing, not only provide personalized scaffolding to facilitate language learning with adaptive conversational strategies, but also support learners' self-regulated learning skills such as goal setting, self-evaluation, and motivation with interactive scaffolding prompts and affective backchannels (Liu et al., 2021).

The use of AI-based interactive scaffolding (AIIS) in language learning has opened new doors to various research foci in human-AI interaction (Gonulal, 2023), including the benefits for both teachers and students in informal digital learning of English (IDLE). For example, language teachers are expected to benefit from AIIS to reduce daily workload by automating repetitive and time-consuming tasks in IDLE and making more data-driven instructional decisions for the unique needs of individual students (Ji et al., 2023). Students also benefit from AIIS for rich language input, ample opportunities to practice, personalized instruction, and immediate feedback in IDLE contexts (Godwin-Jones, 2023), which are necessary components of language learning. These benefits associated with AIIS offer promising solutions to typical intricate problems in language learning, such as teachers' lacking time and experience with various students' ideas to formulate personalized feedback to individual students, limited class time, insufficient speaking opportunities, and unwillingness to speak (Yuan, 2023).

While the current literature has illuminated the multifaceted benefits of using AIIS to facilitate IDLE, it leaves three major issues open for further investigation. First, since the effectiveness of AIIS has mostly been investigated under the influence by a complex interplay of the three types of interactions between learners with the learning environment, i.e. learner-instructor interaction (e.g. Dennen et al., 2007), learner-learner interaction (e.g. Hu et al., 2023), and learner-content interaction (e.g. Powell & Leary, 2021), the effects of AIIS itself remain unknown. AIIS is an important component of another type of interactions that has been under-researched, i.e. learner-AI interaction. AIIS is the necessary backbone that allows the aforementioned three interaction types with teachers, learner peers, and content to synergize. Understanding its discrete effect can provide invaluable insights into how to optimize the learning process and incorporate it into daily use or pedagogical designs in AI-supported interactive learning environments. Second, most studies were conducted in higher education, with single experimental designs (Yuan, 2023) and in formal classroom settings, so it is crucial to examine the effectiveness of AIIS in K-12 education and in IDLE settings (Hu et al., 2023; Moore et al., 2023). Third, there is a noteworthy dearth of knowledge about how K-12 students' self-regulated learning skills, e.g. goal setting, self-evaluation, and motivation, change over time in IDLE with AIIS (Godwin-Jones, 2023). Without such knowledge about the effects of AIIS and what happens during interaction, the mechanism of AI-assisted language learning continues to be a black box (Wang et al., 2023).

The dizzying developments of AI-based mobile technologies that are pervasive in our social life and informal digital learning settings make us wonder whether AI conversational agents embedded in these technologies are effective in facilitating language learning when K-12 students learn from the AI conversational agents independently in such informal settings (Godwin-Jones, 2023; Liu et al., 2021). In this case, examining the effects of AIIS in informal digital learning over time is essential. Therefore, this study conducted a quasi-experiment to investigate the effects of AIIS on secondary students' speaking performance, goal setting, self-evaluation, and motivation in IDLE, and the developmental trend of the three self-regulated skills over time. The findings provide theoretical and pedagogical implications for educators, instructors, researchers, and developers, and outline further research opportunities that investigate the effects of AIIS on K-12 students' language learning in informal digital learning.

Literature review

AI-based interactive scaffolding

The scaffolding theory is based on the idea that learners are to be guided and supported in learning, defining scaffolding as the process that enables learners to solve problems, complete tasks, or reach goals that would otherwise be unattainable without assistance (Wood et al., 1976). Scaffolding can be useful for teachers to provide a supportive environment for learners and simultaneously reinforce learners' independence (Puntambekar, 2022). Due to the dynamic nature of scaffolding, the concept of scaffolding needs to be modified to fit the circumstances of implementation (Steinert et al., 2024). Extended conceptualizations of scaffolding have been developed into computer-based scaffolding, defined as support from computer-based tools that provide either context-specific or generic scaffolding (Belland, 2014). Building on computer-based scaffolding, further conceptualizations have been developed within AI-supported language learning environments to incorporate AI-based interactive scaffolding (AIIS).

As a subset of computer-based scaffolding, AIIS refers to the interactive process in which learners communicate with AI to acquire knowledge or skills by carrying out learning activities that would be beyond their unassisted efforts. As AI possesses increasingly adaptive capabilities, both learners and AI share the control over learning to adapt to each other, and increased bidirectional interaction between learners and AI occurs (Liu et al., 2021). Therefore, beyond computer-based scaffolding, AIIS introduces dynamic interaction and instant feedback in contextualized learning activities (Chien et al., 2024).

Four main types of AIIS have been identified: conceptual, metacognitive, procedural, and strategic scaffolding (Doo et al., 2020). Conceptual scaffolding assists learners in understanding and using concepts in language learning. It helps learners to make connections between concepts or simplify complex concepts (Kim et al., 2022). Metacognitive scaffolding assists learners in reflecting on what they have learned, evaluating their current learning outcomes against set standards, and setting goals and strategies for language improvement (Lim et al., 2023). Procedural scaffolding assists learners in knowing how to utilize resources, so it focuses on introducing specific functions of the user interface, procedures, or navigations (Kim et al., 2022). Strategic scaffolding assists learners by directly or indirectly suggesting approaches to enhance language learning and providing strategies or pathways to complete tasks (Liu et al., 2021), such as guiding them toward constructing solid and cohesive sentence structures in writing (Kim et al., 2022). The four types of AIIS scaffold language learning through various formats, including prompts, questions, feedback, guidance, hints, and expert models.

Critical to AIIS are the three key characteristics: contingency, fading, and transfer of responsibility (Belland et al., 2024; van de Pol et al., 2010). Contingency in AIIS means that AIIS needs to be provided based on a dynamic assessment of learners' current abilities and task features within specific learning contexts (Bernacki, 2018). Contingency is often related to fading, where fading refers to the gradual decrease in the intensity or frequency of support in learning (Puntambekar, 2022). In AIIS, AI dynamically assesses learners' current learning and provides sufficient scaffolding. As learners improve, they require less scaffolding, so responsibility is transferred to learners, and AIIS is finally not needed (Jennings & Muldner, 2021). Among the three characteristics, fading and transfer of responsibility have been underexplored for evidence of learners' decreased reliance on AIIS (Belland et al., 2024).

AIIS to facilitate informal digital learning of English

Drawing on AIIS to facilitate informal digital learning of English has recently gained momentum with the proliferation of AI conversational agents in mobile technologies. Informal digital learning of English (IDLE), as a subfield of computer-assisted language learning, has focused on the changing learning environment in which language learners can acquire the target language readily outside the classroom (Lee & Dressman, 2018). IDLE is defined as self-directed informal English learning by using various digital devices (e.g. smartphones and laptops) and resources (e.g. apps and social media), independent of formal contexts (Lee & Dressman, 2018). IDLE has intricate dynamics of language development, since language learners have various choices in rich digital resources, with different motives and goals in accessing learning materials and interacting with communities. The main challenges of comprehending the intricate dynamics lie in the nonlinear language development path as compared to structured language instruction, self-organizational character as learners and system elements adapt to each other, and the focus on emergent, diverse, and unpredictable outcomes (Godwin-Jones, 2018). AIIS provides opportunities to address these challenges in IDLE (Moore et al., 2023). Specifically, scaffolding is derived from the affordances of AI technologies: adaptivity (i.e. imparting knowledge based on the level of learners), adaptability (i.e. delivering contents based on learners' preferred modalities), continuous assessment, continuous data collection, recommendation of content, and evaluation of recommendation (Murtaza et al., 2022). Thus, language learners gain opportunities to make dialogic exchanges in personalized learning (Hu et al., 2023), receive instant feedback, experience reduced learning anxiety (Harley et al., 2016), and practice the target language in authentic communication environments beyond the classroom (Ji et al., 2023).

AIIS to facilitate secondary students' self-regulated speaking learning, goal setting, self-evaluation, and motivation in IDLE

Self-regulated learning (SRL) is essential for effective learning, especially in computer-based learning environments (Azevedo, 2007). SRL is the self-directive process through which learners proactively transform mental abilities into academic skills (Zimmerman, 2002). SRL can be considered as a cycle model of three main phases: forethought, performance, and self-reflection (Zimmerman, 2011). SRL is closely related to English learning performance, and learners with strong SRL abilities are able to learn independently and overcome diverse challenges (Xia et al., 2023). However, learners most often do not spontaneously regulate learning and face difficulties in adequately regulating their learning process, which is exacerbated in computer-based learning environments where monitoring and controlling learning is complex (Broadbent & Poon, 2015).

As to these learning difficulties, AIIS can be used to facilitate learners' SRL and consequently improve their learning performance (Lim et al., 2023). AIIS provides personalized support to learners through scaffolds such as prompts, feedback, or strategic guidance. AIIS is able to support the three SRL phases in forethought, performance, and self-reflection by creating resourceful, interactive, and engaging learner-centered environments based on learners' needs (Chiu, 2024). Specifically, according to Zimmerman (2002), the first phase of forethought refers to the process that occurs before learning efforts begin, including self-motivation and task analysis. Self-motivation arises from learners' beliefs about learning such as self-efficacy and interest. The motivation influences task analysis by affecting perceptions of how well the task can be performed. The second phase of performance involves self-observation and self-control. Based on self-observation, learners control their learning by using various self-control strategies, such as attention focusing and task strategies. The third phase of self-reflection includes self-judgment and self-reaction. In response to self-judgment of performance, learners' self-reaction can increase or decrease motivation to learn further. Based on their perceived achieved goals, learners evaluate their learning performance. Overall, the SRL cycle requires learners' goal setting (i.e. the process that learners set particular goals for their SRL), self-evaluation (i.e. the process that learners compare self-observed performance against some standards such as learning analytics or outcomes provided by AI), and motivation (i.e. learners' beliefs about learning such as self-efficacy and task interest) through the learning process. In relation to the facilitation of SRL, language learners' performance can be improved by providing AIIS to facilitate learners' SRL activities (Jin et al., 2023; Lim et al., 2023; Xia et al., 2023).

To date, most previous studies have focused on the application of AIIS for language learning in higher education, particularly for enhancing writing, speaking, and self-regulated skills (Doo et al., 2020; Liu et al., 2021). Results indicated that AIIS not only improved writing and speaking performance, but also enhanced self-regulated skills such as goal setting, self-evaluation, and motivation (e.g. Ardiningtyas et al., 2023; Liu et al., 2021).

AIIS for speaking learning in IDLE is receiving attention in secondary education. Secondary education poses challenges that are different from those in higher education. Secondary students have common naïve conceptions of AI (Kim et al., 2023), which might influence their interaction with AIIS. To the best of our knowledge, no study has been conducted to investigate the effectiveness of AIIS in secondary education and in IDLE settings. Given these challenges and no study on this topic, the effects of AIIS on secondary students' speaking performance, goal setting, self-evaluation, and motivation in IDLE remain unclear.

Less is known about the developmental trajectories (i.e. pathways) of secondary students' goal setting, self-evaluation, and motivation with AIIS in IDLE. Adolescence has been considered as a turning point for developing metacognitive monitoring, reflection, and regulatory skills, so it is expected that adolescents' goal setting, self-evaluation, and motivation should gradually increase (Bardach et al., 2023). However, relevant studies indicated mixed potential trajectories, especially in new learning environments. For example, when new learning environments presented cognitive learning difficulty, the levels of learners' self-rated attainment such as goal setting, self-evaluation, and motivation tended to decrease from the beginning until the third or seventh month, followed by a stabilizing trend afterwards (Koizumi & Matsuo, 1993). This decline-and-stabilizing trend reflected the process that the learners adjusted their goals based on self-evaluation until the goal scores came to approach their latest learning performance. Moreover, a rising trend was found in goal setting, self-evaluation, and motivation in online learning (Chiu, 2023). While these studies suggested potential developmental trajectories, in the specific context of using AIIS for secondary students' speaking learning in IDLE, the developmental trajectories are unknown.

Therefore, this study aims to investigate whether AIIS supports secondary students' speaking learning in IDLE by specifically addressing the two research questions (RQ):

1. In IDLE, does AI-based interactive scaffolding support secondary students' speaking performance, goal setting, self-evaluation, and motivation?
2. How do secondary students' goal setting, self-evaluation, and motivation change over time when using AI-based interactive scaffolding in IDLE?

The integrative framework for analysis of scaffolding strategies as the theoretical framework

To investigate scaffolding strategies in AIIS within the experimental design, this study draws on the integrative framework for analysis of scaffolding strategies by van de Pol et al. (2010), which offers a theoretical account of scaffolding classifications. This framework is grounded in Wood et al. (1976) for "scaffolding intentions" and Tharp and Gallimore (1988) for "scaffolding means". Scaffolding intentions refer to what is scaffolded, and scaffolding means refer to how scaffolding is taking place (van de Pol et al., 2010). The distinction in the framework includes three main scaffolding intentions (i.e. learners' metacognitive activities, cognitive activities, and affect) and six scaffolding means (i.e. feeding back, hints, instructing, explaining, modeling, and questioning). This distinction enables nuanced descriptions of scaffolding, so it has been employed to analyze computer-based scaffolding (Nickl et al., 2024) and AI-supported learning environments (Winters et al., 2023).

The study explores scaffolding strategies in AIIS and their effects on secondary students' speaking performance, goal setting, self-evaluation, and motivation in IDLE. To fit with the context of the research, as Table 1 displays, the definitions and examples of AIIS strategies have been adapted by drawing on the framework (Tharp & Gallimore, 1988; van de Pol et al., 2010; Wood et al., 1976).

Method

To address the RQs, this study employed the quasi-experimental design in a two-group setting, including an interactive scaffolding group as the treatment group (TG) and a fixed scaffolding group as the control group (CG). The study was approved by the Survey and Behavioral Research

Table 1. Definitions and examples of AI-based interactive scaffolding strategies in the study.

AI interactive scaffolding	Category	Definition	Example
Scaffolding intentions	1. Goal setting	Keeping the learning process on target and maintaining learners' pursuit of particular objectives.	AI: We are going to talk about a new topic in this lesson. It is the first topic we have talked about together. Let's check out what it is.
	2. Self-evaluation	Enabling learners to compare self-observed performance against some standards such as learning analytics or outcomes provided by AI.	AI: Let's take a look at what we have learned today.
	3. Motivation	Motivating learners' beliefs about learning such as self-efficacy and task interest.	AI: Wow, you did a great job.
	4. Cognitive activities	Providing explanations and justification in understanding and practicing language.	AI: For example, I want to go to Asia. We usually use "want to do something" to express our intention.
Scaffolding means	1. Feeding back	Providing information regarding learners' performance.	ST: Brother. AI: Fantastic/wonderful/I'm impressed/ Good/Try again.
	2. Hints	Giving clues or suggestions to help learners forward.	AI: How does the boy call the girl? ST: I don't know. AI: Can you tell me what the boy should call the girl, sister or daughter?
	3. Instructing	Telling students what to do, how, and why.	ST: Son. AI: Great. But please pay attention to the pronunciation, /ʌ/. Don't open your mouth too wide. Be short.
	4. Further explaining	Providing more detailed information or clarification.	ST: Your. AI: One correct answer could be "their". Could you please repeat it?
	5. Modeling	Demonstrating behavior for imitation.	AI: Read after me, /?fæməli/. ST: Family.
	6. Further questioning	Asking learners questions to require active linguistic answers.	AI: If your friend says, "I am happy", what will you say? ST: Good for you.

Note. AI refers to the AI conversational agent. ST refers to the student.

Ethics Committee with the approval number: SBRE-22-0132. The researchers explained to the students and their parents the objective and process of this study. Written informed consent was obtained.

Participants

The participants were 60 seventh graders between 12 and 13 years old from an average junior middle school in Southern China. This school is located in a medium-sized city, with about 400 teachers and 5000 low to medium socioeconomic students. The sample consisted of 25 females (41.7%) and 35 males (58.3%). According to the pre-experiment survey of their teachers in school, the participants had little experience with speaking practice in class because their English learning was test-oriented, and their English learning tests were predominantly in the written form. Most participants did not have the experience of interacting with AIIS to learn speaking in IDLE before this study (91.1%).

The participants were from two classes taught by the same teacher who had taught English for six years. The teacher was teaching five classes with about 60 students in each class. From her five classes, two classes were randomly selected, and from each intact class, 34 students were randomly selected. Before the experiment started, 34 students from Class A were randomly selected into the TG and 34 students from Class B into the CG. During the study, because four students in the TG and four students in the CG missed some sessions, this study only included those participants who attended all the sessions. Therefore, the final numbers of participants were 30 in the TG and 30 in the CG.

Study design

Procedure

As shown in Figure 1, this study comprised three stages. In Stage One within one week, all students took pre-tests on prior speaking proficiency, goal setting, self-evaluation, and motivation (60 minutes in total, with 40 minutes for assessing prior speaking proficiency and 20 minutes for measuring goal setting, self-evaluation, and motivation). Both groups received experimental orientations separately on how to learn speaking under different conditions. Each orientation, lasting 30 minutes, provided an overview of the scaffolding environment and condition, researcher-led demonstrations of the learning process, and students' participation in the learning condition to ensure their comprehensive understanding of the app usage and the learning process.

In Stage Two over a four-week period, the participants learned five topics, each introduced by the same AI conversational agent every five to six days. The five topics were "family", "clothes", "emotion", "music", and "food". Each topic was learned in two consecutive sessions, separated by a ten-minute break. The first session covered three learning activities: (1) reading words, phrases, and sentences after the agent, (2) multiple choice and blank-filling exercises to review the expressions, and (3) real-time testing with the pictures in practical application scenarios. The second session included three more learning activities: (4) role-play with the agent to make a dialogue, (5) personalized feedback on pronunciation, grammar, vocabulary, etc., and (6) adaptive

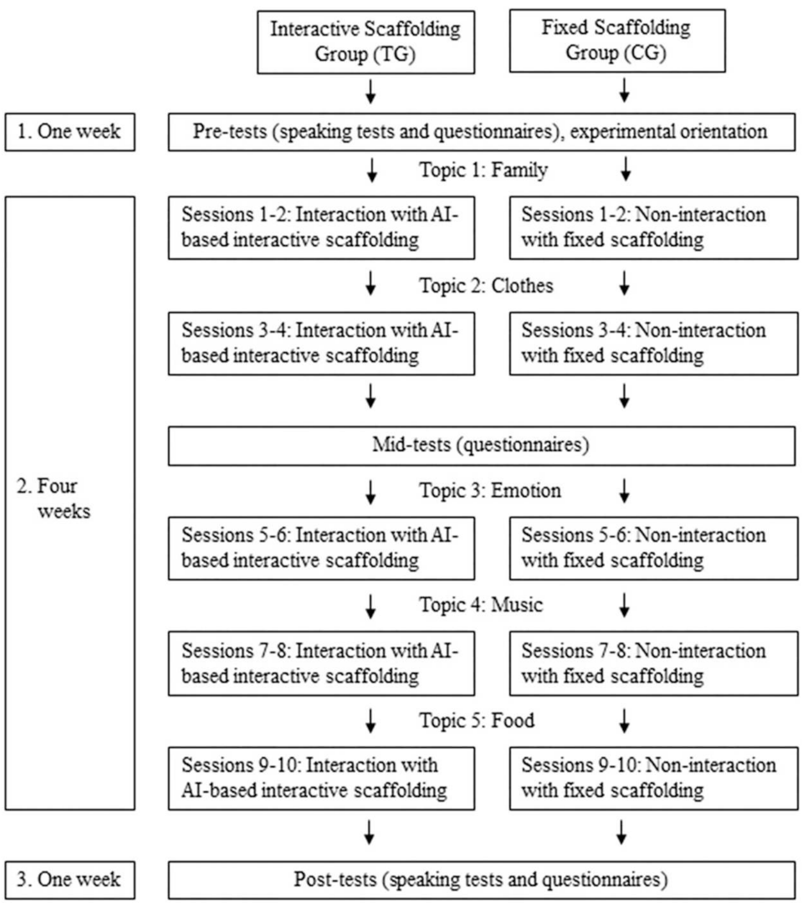

Figure 1. The experimental procedure.

instructional prompts for language errors. The six activities were sequenced in increasing difficulty to gradually scaffold speaking learning. After the second topic, mid-tests were conducted to investigate potential changes in goal setting, self-evaluation, and motivation.

In Stage Three within one week, the participants took post-tests on speaking performance, goal setting, self-evaluation, and motivation (60 minutes in total, with 40 minutes for examining speaking proficiency and 20 minutes for measuring goal setting, self-evaluation, and motivation).

All the students learned speaking independently from an AI conversational agent outside the classroom without the teacher's instruction. The teacher did not attend the process to mitigate any potential influence the teacher might exert. The research team facilitated the implementation process of the experiment by providing timely assistance for technological issues while students were using the mobile phone and observing the students' learning process to ensure consistency in two different conditions of the scaffolding environments.

Scaffolding environments and conditions

This study used an AI conversational agent developed by LAIX, as shown in Figure 2. Empowered by deep learning and adaptive learning AI technologies, the mobile app provided various speaking courses on different topics, with pronunciation and grammar correction, speech evaluation, real-time personalized feedback, and learning assessment (Liulishuo, 2022). The learning process was completed through interaction between students and AIIS. Both the TG and CG groups were provided with identical mobile phone models to ensure that the app interface they saw was exactly the same.

Under the AI-based interactive scaffolding condition, the TG had access to the AI conversational agent that provided interactive scaffolding to learn speaking and enact the self-directed learning process related to goal setting, self-evaluation, and motivation. Specifically, the students interacted orally with the agent to conduct the six aforementioned activities for each topic, supported by the six means of AIIS: feeding back, hints, instructing, further explaining, modeling, and further questioning. These means of AIIS guided the students as they responded to and posed questions to the agent. The two-way interaction, facilitated automatically by AIIS from the agent, allowed for both

Figure 2. Screenshots of AI-based interactive scaffolding.

immediate and on-demand responses when the students activated the interaction by asking questions. The students were permitted to interact with the agent to repeat the learning activities as needed, similar to their independent use of the app in IDLE. Each interactive session in the TG lasted between 20 and 30 minutes, and the average duration of each session was 23.6 minutes.

Under the fixed scaffolding condition, the CG followed similar learning activities as the TG, but without adaptive instructional prompts and personalized feedback from the same AI conversational agent. Specifically, they watched fixed scaffolding videos by the same agent about the same content to learn speaking. This experimental setting was based on the idea in vicarious online learning that learners can experience an interaction vicariously by observing and putting themselves in the shoes of one of the participants in the observed interaction (Dai & Shi, 2022; Dai & Walther, 2018). The fixed scaffolding setting was adapted from the experimental design by Chi et al. (2008) in that the students watched the video with fixed scaffolding, and practiced on their own to answer the agent's questions. When the agent in the video asked learners questions, a short period of quiet time was left for learners to answer by themselves. After the pause, the agent's scaffolding continued until all the content was delivered. In other words, students in the CG received support through the agent's fixed scaffolding.

The similarities between the two groups were that they all orally performed the six learning activities for each topic scaffolded by the same agent on the same mobile phone. However, the difference lies in that the CG did not interact with the agent's interactive scaffolding, experiencing only one-way interaction from the agent to the students. The agent guided the students based on preset fixed scaffolding. The students did not respond to the agent, nor did the agent offer immediate, on-demand, and personalized responses. Compared with the six means of interactive scaffolding, the agent in this condition did not use these means of scaffolding: feeding back, hints, further explaining, and further questioning, which should be based on students' responses. Instead, the agent used only four means of scaffolding: instructing, modeling, initial explaining, and initial questioning. To mitigate the effect of unequal practice compared to the TG, the learners in the CG could choose to watch the video and practice again. Each session in the CG was between 18 and 27 minutes, and the average duration was 21.9 minutes.

Measures

The participants were individually pre-tested in November 2022 and post-tested in January 2023. This study collected two data sources: (1) speaking learning outcomes, and (2) goal setting, self-evaluation, and motivation data.

For speaking learning outcomes, this study adapted the speaking test for A1 Movers of Cambridge English Qualifications in both the pre-test and post-test (see Appendix A). The speaking test covered four sub-dimensions of speaking: fluency, vocabulary, grammar, and pronunciation (Leong & Ahmadi, 2017). An English teacher familiar with the speaking test individually tested all the participants. The testing process was audio-recorded by the research team for subsequent rating. Two raters, including another English teacher and one researcher who were both experienced in English teaching, scored the students' speaking learning performance based on these recordings. The inter-rater reliability of 93.7% was achieved. To measure goal setting, self-evaluation, and motivation, a survey was designed based on prior studies (Barnard et al., 2009; Cheng et al., 2010; Williamson, 2007), as shown in Appendix B. This survey included 23 items. The Cronbach's alpha values of goal setting, self-evaluation, and motivation reached 0.921, 0.897, and 0.859, respectively.

Data analysis

Descriptive analysis was performed to explore the scaffolding strategies provided by AIIS during AI's interaction with the students. Data for this analysis were collected from audio-recordings made by the research team about the students' learning process supported by AIIS. The recordings of AIIS for the five learning topics were transcribed. Based on the integrative framework for analysis of scaffolding strategies by van de Pol et al. (2010), three coders, including two researchers of this study and one independent researcher, coded the interactive scaffolding strategies by AI. The inter-rater reliability among the three coders reached 95.7%. When disagreement occurred, discussions were conducted until a consensus was reached.

To address RQ1, univariate analyses of covariance (ANCOVAs) were conducted. Prior to conducting ANCOVAs, this study examined the assumption of the homogeneity of regression slopes to evaluate if there was an interaction between the treatment and the covariates. For interpretation of the effects, effect sizes (Cohen's *d*) were calculated according to the adjusted differences between the TG and the CG divided by the pooled standard deviation (Cohen, 1988). For RQ2, two-way repeated measures ANOVAs were conducted at three time points: pre-test, mid-test (i.e. after Topic 2), and post-test.

Results

Descriptive results

As shown in Table 2, a total of 565 scaffolding means in AIIS were observed and coded across five topics: 125 for Topic 1 (22.12%), 120 for Topic 2 (21.24%), 114 for Topic 3 (20.18%), 111 for Topic 4 (19.65%), and 95 for Topic 5 (16.81%). There was a decreasing trend of AIIS as the students increasingly interacted with AIIS in learning speaking, and they required less scaffolding support.

Instructing (31.68%) and further questioning (19.29%) were the most commonly used scaffolding means in AIIS, followed by feeding back (15.4%), modeling (12.92%), hints (10.44%), and further explaining (10.27%). The scaffolding intentions were mostly for cognitive activities (56.11%). Motivation (17.52%), self-evaluation (13.62%), and goal setting (12.74%) were also observed.

The effects of AIIS on speaking performance, goal setting, self-evaluation, and motivation in IDLE (RQ1)

Prior to conducting ANCOVAs, this study tested the assumption of the homogeneity of regression slopes to ensure there was no interaction between the treatment and the covariates (D'Alonzo, 2004). The *p*-values of all interactions between the independent treatment variable (group) and the pre-post test scores were above 0.05, as shown in Table 3, indicating that the homogeneity of regression was not violated. Therefore, ANCOVA analyses were performed.

The results of ANCOVAs in Table 4 showed that there was no significant difference in the overall English speaking scores ($F = 1.98$, $p > .05$, Cohen's $d = 0.13$). When a set of ANCOVAs was conducted to examine the effects on the four sub-dimensions, a significant difference was found in vocabulary ($F = 4.98$, $p < .05$, Cohen's $d = 0.36$). No significant differences were found in other sub-dimensions: pronunciation ($F = 2.29$, $p > .05$, Cohen's $d = 0.26$), grammar ($F = 1.18$, $p > .05$, Cohen's $d = -0.20$), and fluency ($F = 0.13$, $p > .05$, Cohen's $d = 0.06$). These suggested that despite no significant difference in the overall speaking scores, AIIS was found to be especially beneficial for the learners' vocabulary learning.

The results also revealed that compared to the CG, the TG had significantly higher scores at the post-test on goal setting ($F = 5.14$, $p < .05$, Cohen's $d = 0.79$), after controlling for the pre-test scores. While the treatment effects were not significantly different on self-evaluation ($F = 3.21$, $p > .05$), and on motivation ($F = 0.70$, $p > .05$), the effect sizes were positive, with 0.43 for self-evaluation, and 0.22 for motivation. These suggested that while AIIS showed positive effects on goal setting, self-evaluation, and motivation, AIIS was found to be particularly beneficial for improving the learners' goal setting.

Developmental trajectories of goal setting, self-evaluation, and motivation in IDLE (RQ2)

Using two-way repeated measures ANOVAs, the developmental trajectories were summarized as the decline and rise trend on goal setting, self-evaluation, and motivation for both groups. Specifically, in

Table 2. Frequency distribution of scaffolding means and intentions in AI-based interactive scaffolding.

	Scaffolding intentions																					
	Topic 1				Topic 2				Topic 3				Topic 4				Topic 5					
Scaffolding means	G	S	M	C	G	S	M	C	G	S	M	C	G	S	M	C	G	S	M	C	Total	Percentage
Feeding back	2	2	9	4	4	2	8	3	3	3	6	5	4	2	7	6	4	3	5	5	87	15.40%
Hints	1	2	2	4	2	2	3	4	3	3	1	6	2	2	2	7	3	1	3	6	59	10.44%
Instructing	4	5	3	32	3	4	3	29	3	2	2	31	2	2	1	27	2	1	2	21	179	31.68%
Further explaining	1	2	1	2	2	3	2	3	2	4	3	3	3	4	3	5	3	3	4	5	58	10.27%
Modeling	2	2	4	7	3	4	4	7	2	3	6	5	1	4	4	5	2	2	3	3	73	12.92%
Further questioning	2	2	1	29	1	3	2	19	1	1	2	14	2	2	1	13	3	2	2	7	109	19.29%
Total	125				120				114				111				95					
Percentage	22.12%				21.24%				20.18%				19.65%				16.81%					

Note. G refers to goal setting. S refers to self-evaluation. M refers to motivation. C refers to cognitive activities.

Table 3. Homogeneity of regression slopes for treatment (group) and covariate interactions.

Test	Group interactions	
	F (1, 56)	*p*-value
Overall speaking performance	0.517	0.475
Fluency	0.290	0.592
Vocabulary	0.312	0.579
Grammar	0.264	0.609
Pronunciation	0.438	0.511
Goal setting	0.043	0.836
Self-evaluation	2.043	0.159
Motivation	0.095	0.759

Table 4. The results of ANCOVAs.

	Treatment group (*N* = 30)		Control group (*N* = 30)		*p*-value	*F*	Cohen's *d*
	Adjusted mean	*SE*	Adjusted mean	*SE*			
Goal setting	5.43	0.14	4.95	0.14	0.03	5.14*	0.79
Self-evaluation	5.04	0.10	4.78	0.10	0.08	3.21	0.43
Motivation	5.02	0.12	4.87	0.12	0.41	0.70	0.22
Overall speaking performance	69.24	0.44	68.27	0.44	0.17	1.98	0.13
Fluency	18.18	0.21	18.06	0.21	0.72	0.13	0.06
Vocabulary	17.32	0.22	16.58	0.22	0.03	4.98*	0.36
Grammar	16.33	0.22	16.71	0.22	0.28	1.18	−0.20
Pronunciation	17.58	0.21	17.09	0.21	0.14	2.29	0.26

Note. *$p < .05$. SE refers to standard error.

all the three outcome variables, the TG showed a limited decline that appeared flat, and then a gradual rise, and the CG showed a steep decline, and then a gradual rise, as shown in Figures 3–5.

The results of Mauchly's test of sphericity showed that sphericity was met ($p > .05$) for all the three outcome variables. Partial eta-squared (η_p^2) was employed to estimate the effect size, following the thresholds suggested by Cohen (1988): small $\eta^2 > 0.01$, medium $\eta^2 > 0.06$, and large $\eta^2 > 0.14$. The results showed that the main effect of time and the between-subjects effects of groups were statistically significant on all the three variables, as shown in Table 5, with *F* value ranging from 15.88 to 50.96, and large effect sizes ($\eta^2 > 0.14$). The time × group interaction effects on all the

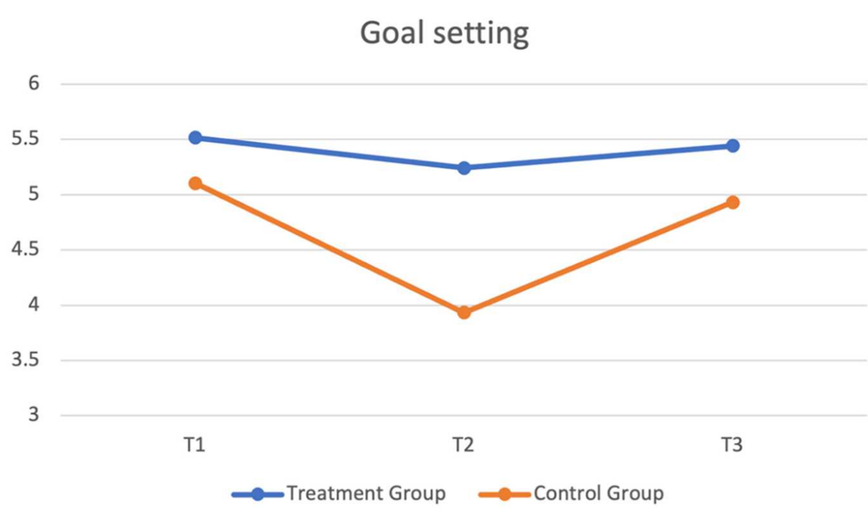

Figure 3. Developmental trajectories of goal setting.

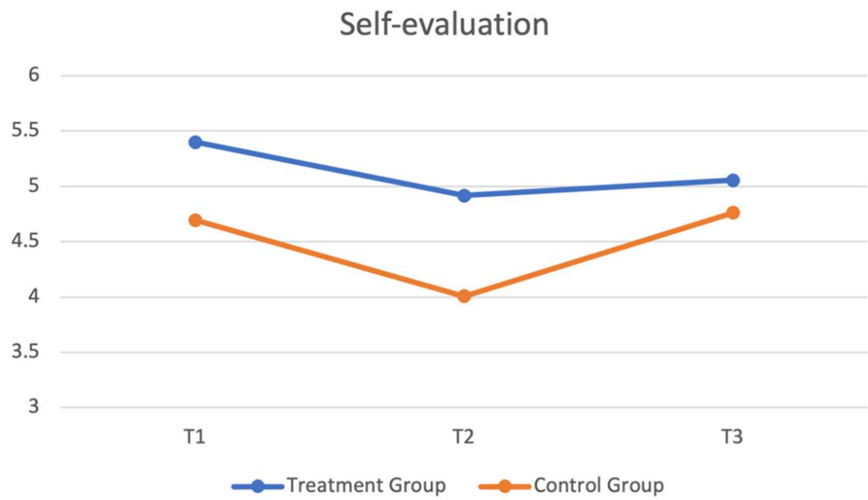

Figure 4. Developmental trajectories of self-evaluation.

three variables were also statistically significant. The effect sizes for goal setting and self-evaluation were medium ($\eta^2 > 0.06$), and the effect size for motivation was large ($\eta^2 > 0.14$).

To further investigate the developmental trajectories of goal setting, self-evaluation, and motivation, post-hoc analyses were conducted using the Bonferroni test. As displayed in Table 6, the post-hoc analyses revealed the following significant differences: (1) goal setting: students' goal setting significantly decreased from T1 to T2 in both the TG and CG. Students' goal setting significantly increased from T2 to T3 in the CG. (2) Self-evaluation: students' self-evaluation decreased from T1 to T2 in both the TG and CG. Students' self-evaluation significantly decreased from T1 to T3 in the TG. Students' self-evaluation significantly increased from T2 to T3 in the CG. (3) Motivation: students' motivation decreased from T1 to T2 in both the TG and CG. Students' motivation significantly decreased from T1 to T3 in the TG. Students' motivation significantly increased from T2 to T3 in the CG.

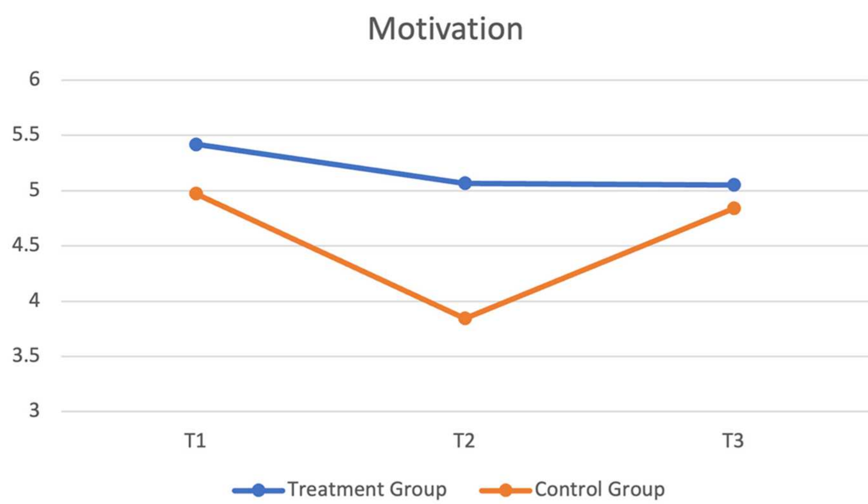

Figure 5. Developmental trajectories of motivation.

Table 5. The two-way repeated measures ANOVA.

	Goal setting			Self-evaluation			Motivation		
	Time	Group	Time × Group	Time	Group	Time × Group	Time	Group	Time × Group
F	21.04***	40.84***	8.61***	15.88***	50.96***	4.16*	20.68***	30.20***	9.82***
η^2	0.27	0.42	0.13	0.22	0.47	0.07	0.27	0.35	0.15

Note. *$p < 0.05$, ***$p < 0.001$

Despite the significant differences, due to the small sample sizes of the two groups, the results should be interpreted with caution. However, the results showed different developmental trajectories for the two groups, as shown in Figures 3–5. Goal setting, self-evaluation, and motivation showed the decline and rise trajectories, i.e. decreasing in the middle of the experiment after Topic 2, and increasing at the end of the experiment. Moreover, the trajectories of the TG dropped less than the CG on goal setting, self-evaluation, and motivation in the middle of the experiment.

Discussion

As noted in the results, scaffolding means in AIIS facilitated the students' speaking learning in IDLE, with a decreasing trend in the number of scaffolding means across five topics. This aligns with the key characteristics of AIIS of fading and transfer of responsibility, meaning that scaffolding tends to withdraw gradually as learners' levels of development and competence increase, and responsibility for task performance is gradually transferred to learners as they take more control (Belland et al., 2024; van de Pol et al., 2010). In AIIS, AI dynamically evaluates learners' progress, providing appropriate scaffolding that gradually diminishes as learners improve and take on more responsibility, reducing the need for AIIS (Jennings & Muldner, 2021).

Instructing and further questioning were identified as the most common scaffolding means in AIIS. The result is in line with Winters et al. (2023) in that instruction and questioning are the most common scaffolding means for secondary students. Possible reasons are that instruction and skillful questioning foster meaningful learning by encouraging students to express their views and ask follow-up questions in interactive communication (Pun & Macaro, 2019). Of all the scaffolding intentions identified, the intention of supporting cognitive activities was the most frequent. This echoes the positive correlations with instruction and further questioning in supporting

Table 6. Post-hoc Bonferroni analysis of overtime changes in goal setting, self-evaluation, and motivation in the treatment and control groups.

	Treatment group					Control group				
				95% Confidence Interval					95% Confidence Interval	
Time point	Mean difference	*SE*	*p*-value	Lower bound	Upper bound	Mean difference	*SE*	*p*-value	Lower bound	Upper bound
Goal setting										
T1→T2	−0.27	0.11	0.047	−0.54	−0.00	−1.19	0.19	0.000	−1.66	−0.72
T2→T3	0.20	0.09	0.111	−0.03	0.43	1.00	0.24	0.001	0.38	1.62
T1→T3	−0.07	0.12	1.000	−0.37	0.23	−0.19	0.22	1.000	−0.76	0.38
Self-evaluation										
T1→T2	−0.48	0.12	0.001	−0.80	−0.17	−0.67	0.18	0.002	−1.12	−0.22
T2→T3	0.14	0.06	0.084	−0.01	0.29	0.74	0.21	0.004	0.21	1.26
T1→T3	−0.35	0.11	0.008	−0.61	−0.08	0.07	0.19	1.000	−0.42	0.56
Motivation										
T1→T2	−0.35	0.10	0.006	−0.62	−0.09	−1.14	0.20	0.000	−1.64	−0.63
T2→T3	−0.02	0.07	1.000	−0.19	0.16	0.97	0.25	0.002	0.33	1.60
T1→T3	−0.37	0.11	0.007	−0.65	−0.09	−0.17	0.20	1.000	−0.69	0.34

Note. SE refers to standard error. T1 refers to Time 1. T2 refers to Time 2. T3 refers to Time 3.

cognitive activities. In other words, the scaffolding strategies in AIIS are primarily utilized to support learner cognition, followed by metacognition in self-regulated learning such as motivation, self-evaluation, and goal setting (Järvelä et al., 2023; van de Pol et al., 2010).

The effects of AIIS on speaking performance, goal setting, self-evaluation, and motivation in IDLE (RQ1)

Results in RQ1 demonstrated that while no significant difference was found in the overall score of speaking performance, in its sub-dimension of vocabulary a significant difference was revealed. Specifically, the learners who interacted with AIIS learned more vocabulary than those in the fixed scaffolding group. The results align with the literature that AIIS could improve vocabulary learning (Shi & Tsai, 2022). A possible reason is that AIIS provides dynamic assessment for vocabulary learning, i.e. diagnostic information about the learning process, and promotes learning development as a mediated process of learner transition from other-regulation to self-regulation (Lantolf & Thorne, 2006). Another reason may be that when AIIS presents the form, meaning, and use of words in authentic contexts (Godwin-Jones, 2018), young learners' curiosity, interest, and concentration are high. Therefore, the sub-dimension of vocabulary improves through AIIS in dynamic feedback and concrete demonstrations of authentic learning contexts.

Conversely, no significant difference was found in pronunciation, grammar, and fluency. This may be caused by the limited number of interactive sessions in this study, given that a long-term interaction is expected to improve pronunciation, grammar, and fluency. These results are consistent with the findings that revealed no development in articulation rate and pause frequency over time (Hanzawa, 2024), and a gradual improvement in fluency over one year (Saito & Hanzawa, 2018). Furthermore, the results may be due to the trade-off effect between students' lexical accuracy and errors related to pronunciation, grammar, and fluency. Based on the Trade-off Hypothesis, learners' attentional resources are limited, and allocating attention to one componential construct of linguistic performance may generate a restrictive effect on another in the learning process (Khatib & Farahanynia, 2020). Since AIIS presents words within authentic learning contexts, students tend to memorize vocabulary through interactions with AI. Consequently, students' attentional resources allotted to pronunciation, grammar, and fluency might be restricted, potentially leading to errors in these areas of their utterances (Jiang et al., 2023).

Moreover, results showed that the students who learned with AIIS outperformed those who learned with fixed scaffolding in goal setting, self-evaluation, and motivation, with a significant difference in goal setting. Specifically, in IDLE, AIIS supported the secondary learners' goal setting in particular, besides self-evaluation and motivation. This may be because adaptive learning environments support goal setting in fostering positive emotions. When metacognitive prompts of AIIS are adapted to individual characteristics, positive emotions are found to be more activated than negative emotions (Harley et al., 2016). Therefore, goal setting tends to be enhanced. Additionally, self-evaluation and motivation are also supported by AIIS. An explanation is that AIIS provides feedback on students' learning performance to help them improve their monitoring processes (Ardiningtyas et al., 2023). The enhanced monitoring leads to accurate self-evaluation and changes motivation orientations toward different stimuli from the interactive environment (Liu et al., 2021). Thus, AIIS supports self-evaluation and motivation in addition to goal setting.

Developmental trajectories of goal setting, self-evaluation, and motivation in IDLE (RQ2)

The results indicated that goal setting, self-evaluation, and motivation in both groups decreased from the beginning to the middle, and then increased from the middle to the end of the experiment. Moreover, when the two groups were compared, the TG dropped less than the CG in the middle of the experiment. The overall decline and rise trend for both groups may be, first, due to the learners' unprepared learning conceptions from passively receiving and absorbing knowledge transmitted from teachers, to actively understanding and gaining meaning in interaction with AIIS, i.e. an adaptation between adolescents' expectations for more autonomy and control and the support provided by the learning environment (Bardach et al., 2023). Second, the decline and rise trajectories may be interpreted as the process of attaining dynamic equilibrium between the students and the new learning environment, following perturbations (Wapner et al., 1973). Students tend to experience a gradual decline in self-rated attainment, emotion, and motivation, until their goals, self-evaluation, and motivation adjust to become more realistic over time, after which these measures stabilize or increase (Chiu, 2023; Koizumi & Matsuo, 1993).

Despite the similar decline and rise trends of both groups, the TG showed a less sharp drop in the middle, compared to the CG. In other words, AIIS supported the secondary learners' goal setting, self-evaluation, and motivation in IDLE. This finding is consistent with the study by Lim et al. (2023) in that an important factor related to greater self-regulated learning includes using an online learning environment to track learners' own progress in achieving goal setting and self-evaluation. Moreover, Ebadi and Amini (2022) suggest that the social presence and human-likeness of the AI conversational agent positively affect learners' motivation, which indirectly influences goal setting and self-evaluation. Overall, when secondary learners interact with AIIS, their goal setting, self-evaluation, and motivation are likely to be supported.

Limitations and future research

Four limitations of this study warrant consideration as they pave the way for future research. First, the small sample size recruited for each group might limit the results of the study (Cheung & Slavin, 2016), so future research may consider using larger sample sizes. Second, the results were gathered after ten sessions with an immediate post-test, so a longer-term experimental design with a delayed post-test is expected to generalize the results or reveal a longer-term effect of AIIS on secondary students' speaking learning in IDLE. Third, due to the sudden outbreak of the COVID-19 pandemic at the end of the intervention, the experimental school was closed, and the formal schooling environment of the participating students was transitioned from instruction in the classroom to distance learning at home. Although this study has managed to maintain consistent experimental conditions for both groups until the end, the unexpected transition of the formal schooling environment might have slowed the pace of learning progress in the intervention. Considering that younger students rely more on cognitive scaffolding due to their less developed self-regulation skills and they might be vulnerable to stress and challenges associated with the pandemic (Tomasik et al., 2021), future research may investigate the influence of formal schooling environments on secondary students' self-regulated language learning assisted with AIIS in IDLE. Fourth, this study included seventh-graders only. Since younger and older children show different features in cognitive development, interests, and digital literacy (Moore et al., 2023), future studies might examine the effects of AIIS on different phases of secondary or K-12 students in IDLE, and explore how individual differences of K-12 students shape their interactions with AIIS (Chai et al., 2022).

Theoretical and pedagogical implications

Both theoretical and pedagogical implications can be drawn from the findings. Theoretically, this study extends the application of the integrative framework for analysis of scaffolding strategies into AIIS during learner-AI interaction by revealing scaffolding means and intentions of AIIS. The findings enhance our understanding of scaffolding strategies of AI's interactive scaffolding in secondary students' English learning, particularly in speaking, in informal digital learning. Future research may examine the effectiveness of each scaffolding means in AIIS for supporting different scaffolding intentions. Moreover, this study adds to the literature by identifying the two underexplored characteristics of AIIS: fading and transfer of responsibility. Future research may further explore the nuanced relationship between features related to tasks and learners, and fading of scaffolding and/or transfer of responsibility in AIIS.

Pedagogically, first, this study shows that AIIS facilitates secondary students' speaking learning in IDLE in vocabulary, goal setting, self-evaluation, and motivation. Therefore, secondary students, specifically seventh graders, may have the ability to interact with AIIS to learn speaking in IDLE. Teachers may trust them to use such technology appropriately, especially if given guidance. Second, the findings also suggest that learners' goal setting, self-evaluation, and motivation may decline from the onset of using such technology, and rise from the time when the learners have sufficient interactive sessions with AIIS. Thus, if teachers are able to guide students to familiarize the technology and train self-regulated learning skills before, during, or after the task, the decline phase may be mitigated. Third, since AIIS offers data that reflect the process of learning, if teachers trace and evaluate the data properly, the learning process tends to be enhanced with individualized instruction.

Conclusion

Despite the proliferation of using AIIS to facilitate speaking learning, goal setting, self-evaluation, and motivation in the formal educational contexts of higher education, scant knowledge has been revealed about the effects of AIIS on secondary students' speaking learning in IDLE. Therefore, by comparing the TG of interacting with AIIS to the CG of observing fixed scaffolding, this study investigated whether AIIS could support secondary students' speaking performance, goal setting, self-evaluation, and motivation in IDLE over ten sessions, and how goal setting, self-evaluation, and motivation developed during the process. Findings indicate that AIIS may support secondary students' vocabulary achievement, goal setting, self-evaluation, and motivation. Furthermore, while both groups have shown a decline and rise trend in goal setting, self-evaluation, and motivation in the learning process with AIIS, the TG tends to drop less in the middle of the sessions. These findings provide evidence for the benefits of AIIS in supporting secondary students' speaking learning in IDLE, particularly for improving vocabulary, goal setting, self-evaluation, and motivation.

To explore scaffolding strategies of AIIS in learner-AI interaction, this study theoretically extends the integrative framework for analysis of scaffolding strategies into AIIS, thereby offering educators, instructors, researchers, and developers insightful criteria for analyzing and designing scaffolding strategies in AIIS in terms of scaffolding means and intentions. In this regard, our approach to analyzing scaffolding strategies in AIIS and examining the effects of AIIS on secondary students' self-regulated speaking learning in IDLE marks a step forward in revealing the black box of learner-AI interaction (Jin et al., 2023; Wang et al., 2024). Overall, this study serves as a valuable addition to understanding interactive scaffolding generated by AI to support secondary students' self-regulated speaking learning in IDLE, identifying further research opportunities, and offering theoretical and pedagogical implications for AI-supported interactive learning environments.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Feifei Wang is a postdoctoral fellow at the Centre for Information Technology in Education, The University of Hong Kong, Pokfulam, Hong Kong, China. Her research focuses on language education, educational technology, and evidence-based reform.

Xiaohua Zhou is a senior research officer at the Teaching and Learning Centre, Lingnan University, Hong Kong, China.

Kangxin Li is a doctoral student at the Department of Curriculum and Instruction, The Chinese University of Hong Kong, Hong Kong, China.

Alan C. K. Cheung is a professor and department chairperson at the Department of Educational Administration and Policy, The Chinese University of Hong Kong, Shatin, Hong Kong, China. His research focuses on bilingual education, educational technology, and evidence-based reform.

Ming Tian is a teacher of English at Chenzhou Jinhai School in Chenzhou City, China.

ORCID

Xiaohua Zhou http://orcid.org/0000-0002-6567-3543
Alan C. K. Cheung http://orcid.org/0000-0001-5831-2320

References

Ardiningtyas, S. Y., Butarbutar, R., Weda, S., & Nur, S. (2023). Online scaffolding behavior for speaking EFL improvement: Narrative inquiry issues. *Interactive Learning Environments*, 1–11. https://doi.org/10.1080/10494820.2023.2207608

Azevedo, R. (2007). Understanding the complex nature of self-regulatory processes in learning with computer-based learning environments: An introduction. *Metacognition and Learning*, *2*, 57–65. https://doi.org/10.1007/s11409-007-9018-5

Bardach, L., Yanagida, T., Goetz, T., Jach, H., & Pekrun, R. (2023). Self-regulated and externally regulated learning in adolescence: Developmental trajectories and relations with teacher behavior, parent behavior, and academic achievement. *Developmental Psychology*, *59*(7), 1327–1345. https://doi.org/10.1037/dev0001537

Barnard, L., Lan, W. Y., To, Y. M., Paton, V. O., & Lai, S.-L. (2009). Measuring self-regulation in online and blended learning environments. *The Internet and Higher Education*, *12*(1), 1–6. https://doi.org/10.1016/j.iheduc.2008.10.005

Belland, B. R. (2014). Scaffolding: Definition, current debates, and future directions. In J. Spector, M. Merrill, J. Elen, & M. Bishop (Eds.), *Handbook of research on educational communications and technology* (pp. 505–518). Springer. https://doi.org/10.1007/978-1-4614-3185-5_39

Belland, B. R., Kim, C., Dinç, E., & Zhang, A. Y. (2024). Transfer of responsibility from scaffolding to preservice early childhood teachers learning to debug. *Educational Technology Research and Development*, *72*, 1439–1464. https://doi.org/10.1007/s11423-024-10347-z

Bernacki, M. L. (2018). Examining the cyclical, loosely sequenced, and contingent features of self-regulated learning: Trace data and their analysis. In D. H. Schunk, & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (pp. 370–387). Routledge. https://doi.org/10.4324/9781315697048

Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education*, *27*, 1–13. https://doi.org/10.1016/j.iheduc.2015.04.007

Chai, C. S., Teo, T., Huang, F., Chiu, T. K. F., & Wang, X. W. (2022). Secondary school students' intentions to learn AI: Testing moderation effects of readiness, social good and optimism. *Educational Technology Research and Development*, *70*(3), 765–782. https://doi.org/10.1007/s11423-022-10111-1

Cheng, S.-F., Kuo, C.-L., Lin, K.-C., & Lee-Hsieh, J. (2010). Development and preliminary testing of a self-rating instrument to measure self-directed learning ability of nursing students. *International Journal of Nursing Studies*, *47*(9), 1152–1158. https://doi.org/10.1016/j.ijnurstu.2010.02.002

Cheung, A. C., & Slavin, R. E. (2016). How methodological features affect effect sizes in education. *Educational Researcher*, *45*(5), 283–292. https://doi.org/10.3102/0013189x16656615

Chi, M. T., Roy, M., & Hausmann, R. G. (2008). Observing tutorial dialogues collaboratively: Insights about human tutoring effectiveness from vicarious learning. *Cognitive Science*, *32*(2), 301–341. https://doi.org/10.1080/03640210701863396

Chien, C. C., Chan, H. Y., & Hou, H. T. (2024). Learning by playing with generative AI: Design and evaluation of a role-playing educational game with generative AI as scaffolding for instant feedback interaction. *Journal of Research on Technology in Education*, 1–20. https://doi.org/10.1080/15391523.2024.2338085

Chiu, T. K. (2023). Student engagement in K-12 online learning amid COVID-19: A qualitative approach from a self-determination theory perspective. *Interactive Learning Environments*, *31*(6), 3326–3339. https://doi.org/10.1080/10494820.2021.1926289

Chiu, T. K. (2024). A classification tool to foster self-regulated learning with generative artificial intelligence by applying self-determination theory: A case of ChatGPT. *Educational Technology Research and Development*, *72*, 2401–2416. https://doi.org/10.1007/s11423-024-10366-w

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

Dai, Y., & Shi, J. (2022). Vicarious interactions in online support communities: The roles of visual anonymity and social identification. *Journal of Computer-Mediated Communication*, *27*(3), https://doi.org/10.1093/jcmc/zmac006

Dai, Y., & Walther, J. B. (2018). Vicariously experiencing parasocial intimacy with public figures through observations of interactions on social media. *Human Communication Research*, *44*(3), 322–342. https://doi.org/10.1093/hcr/hqy003

D'Alonzo, K. T. (2004). The Johnson-Neyman procedure as an alternative to ANCOVA. *Western Journal of Nursing Research*, *26*(7), 804–812. https://doi.org/10.1177/0193945904266733

Dennen, V. P., Aubteen Darabi, A., & Smith, L. J. (2007). Instructor–learner interaction in online courses: The relative perceived importance of particular instructor actions on performance and satisfaction. *Distance Education*, *28*(1), 65–79. https://doi.org/10.1080/01587910701305319

Doo, M. Y., Bonk, C., & Heo, H. (2020). A meta-analysis of scaffolding effects in online learning in higher education. *International Review of Research in Open and Distributed Learning*, *21*(3), 60–80. https://doi.org/10.19173/irrodl.v21i3.4638

Ebadi, S., & Amini, A. (2022). Examining the roles of social presence and human-likeness on Iranian EFL learners' motivation using artificial intelligence technology: A case of CSIEC chatbot. *Interactive Learning Environments*, 1–19. https://doi.org/10.1080/10494820.2022.2096638

Godwin-Jones, R. (2018). Chasing the butterfly effect: Informal language learning online as a complex system. *Language Learning & Technology*, *22*(2), 8–27.

Godwin-Jones, R. (2023). Smart devices and informal language learning. In D. Toffoli, G. Sockett, & M. Kusyk (Eds.), *Language learning and leisure: Informal language learning in the digital Age* (pp. 69–88). Walter de Gruyter. https://doi.org/10.1515/9783110752441-004

Gonulal, T. (2023). Investigating EFL learners' humorous interactions with an intelligent personal assistant. *Interactive Learning Environments*, *31*(7), 4521–4534. https://doi.org/10.1080/10494820.2021.1974489

Hanzawa, K. (2024). Development of second language speech fluency in foreign language classrooms: A longitudinal study. *Language Teaching Research*, *28*(3), 816–838. https://doi.org/10.1177/13621688211008693

Harley, J. M., Carter, C. K., Papaionnou, N., Bouchet, F., Landis, R. S., Azevedo, R., & Karabachian, L. (2016). Examining the predictive relationship between personality and emotion traits and students' agent-directed emotions: Towards emotionally-adaptive agent-based learning environments. *User Modeling and User-Adapted Interaction*, *26*(2), 177–219. https://doi.org/10.1007/s11257-016-9169-7

Hu, X., He, W., Chiu, T. K., & Zhao, L. (2023). Using a teacher scheme for educational dialogue analysis to investigate student–student interaction patterns for optimal group activities in an artificial intelligence course. *Education and Information Technologies*, *28*(7), 8789–8813. https://doi.org/10.1007/s10639-022-11556-w

Järvelä, S., Nguyen, A., & Molenaar, I. (2023). Advancing SRL research with artificial intelligence. *Computers in Human Behavior*, *147*, 107847. https://doi.org/10.1016/j.chb.2023.107847

Jennings, J., & Muldner, K. (2021). When does scaffolding provide too much assistance? A code-tracing tutor investigation. *International Journal of Artificial Intelligence in Education*, *31*, 784–819. https://doi.org/10.1007/s40593-020-00217-z

Ji, H., Han, I., & Ko, Y. (2023). A systematic review of conversational AI in language education: Focusing on the collaboration with human teachers. *Journal of Research on Technology in Education*, *55*(1), 48–63. https://doi.org/10.1080/15391523.2022.2142873

Jiang, M. Y. C., Jong, M. S. Y., Lau, W. W. F., Chai, C. S., & Wu, N. (2023). Exploring the effects of automatic speech recognition technology on oral accuracy and fluency in a flipped classroom. *Journal of Computer Assisted Learning*, *39*(1), 125–140. https://doi.org/10.1111/jcal.12732

Jin, S. H., Im, K., Yoo, M., Roll, I., & Seo, K. (2023). Supporting students' self-regulated learning in online learning using artificial intelligence applications. *International Journal of Educational Technology in Higher Education*, *20*(1), 37. https://doi.org/10.1186/s41239-023-00406-5

Khatib, M., & Farahanynia, M. (2020). Planning conditions (strategic planning, task repetition, and joint planning), cognitive task complexity, and task type: Effects on L2 oral performance. *System*, *93*, 102297. https://doi.org/10.1016/j.system.2020.102297

Kim, M. K., Kim, N. J., & Heidari, A. (2022). Learner experience in artificial intelligence-scaffolded argumentation. *Assessment & Evaluation in Higher Education*, *47*(8), 1301–1316. https://doi.org/10.1080/02602938.2022.2042792

Kim, K., Kwon, K., Ottenbreit-Leftwich, A., Bae, H., & Glazewski, K. (2023). Exploring middle school students' common naive conceptions of artificial intelligence concepts, and the evolution of these ideas. *Education and Information Technologies*, 1–28. https://doi.org/10.1007/s10639-023-11600-3

Koizumi, R., & Matsuo, K. (1993). A longitudinal study of attitudes and motivation in learning English among Japanese seventh-grade students. *Japanese Psychological Research*, *35*(1), 1–11. https://doi.org/10.4992/psycholres1954.35.1

Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development*. Oxford University Press.

Laricchia, F. (2024). Smart speaker market revenue worldwide 2019-2029. Retrieved November 25, 2024, from https://www.statista.com/forecasts/1367982/smart-speaker-market-volume-worldwide

Lee, J. S., & Dressman, M. (2018). When IDLE hands make an English workshop: Informal digital learning of English and language proficiency. *TESOL Quarterly*, *52*(2), 435–445. https://doi.org/10.1002/tesq.422

Leong, L. M., & Ahmadi, S. M. (2017). An analysis of factors influencing learners' English speaking skill. *International Journal of Research in English Education*, *2*(1), 34–41. https://doi.org/10.18869/acadpub.ijree.2.1.34

Lim, L., Bannert, M., van der Graaf, J., Singh, S., Fan, Y., Surendrannair, S., Rakovic, M., Molenaar, I., Moore, J., & Gašević, D. (2023). Effects of real-time analytics-based personalized scaffolds on students' self-regulated learning. *Computers in Human Behavior*, *139*, 107547. 1–18. https://doi.org/10.1016/j.chb.2022.107547

Liu, C., Hou, J., Tu, Y. F., Wang, Y., & Hwang, G. J. (2021). Incorporating a reflective thinking promoting mechanism into artificial intelligence-supported English writing environments. *Interactive Learning Environments*, 1–19. https://doi.org/10.1080/10494820.2021.2012812

Liulishuo. (2022). A leading "technology + education" company. Retrieved October 6, 2022, from https://www.liulishuo.com/en/aboutus.html

Moore, R. L., Jiang, S., & Abramowitz, B. (2023). What would the matrix do? A systematic review of K-12 AI learning contexts and learner-interface interactions. *Journal of Research on Technology in Education*, *55*(1), 7–20. https://doi.org/10.1080/15391523.2022.2148785

Murtaza, M., Ahmed, Y., Shamsi, J. A., Sherwani, F., & Usman, M. (2022). AI-based personalized e-learning systems: Issues, challenges, and solutions. *IEEE Access*, *10*, 81323–81342. https://doi.org/10.1109/ACCESS.2022.3193938

National Public Radio. (2022). The smart audio report. Retrieved November 12, 2023, from https://www.nationalpublicmedia.com/uploads/2022/06/The-Smart-Audio-Report-Spring-2022.pdf

Nickl, M., Sommerhoff, D., Radkowitsch, A., Huber, S. A., Bauer, E., Ufer, S., Plass, J. L., & Seidel, T. (2024). Effects of real-time adaptivity of scaffolding: Supporting pre-service mathematics teachers' assessment skills in simulations. *Learning and Instruction*, *94*, 101994. https://doi.org/10.1016/j.learninstruc.2024.101994

Powell, S. T., & Leary, H. (2021). Measuring learner–content interaction in digitally augmented learning experiences. *Distance Education*, *42*(4), 520–546. https://doi.org/10.1080/01587919.2021.1986369

Pun, J., & Macaro, E. (2019). The effect of first and second language use on question types in English medium instruction science classrooms in Hong Kong. *International Journal of Bilingual Education and Bilingualism*, *22*(1), 64–77. https://doi.org/10.1080/13670050.2018.1510368

Puntambekar, S. (2022). Distributed scaffolding: Scaffolding students in classroom environments. *Educational Psychology Review*, *34*(1), 451–472. https://doi.org/10.1007/s10648-021-09636-3

Saito, K., & Hanzawa, K. (2018). The role of input in second language oral ability development in foreign language classrooms: A longitudinal study. *Language Teaching Research*, *22*(4), 398–417. https://doi.org/10.1177/13621688211008693

Shi, Y.-S., & Tsai, C.-Y. (2022). Fostering vocabulary learning: Mind mapping app enhances performances of EFL learners. *Computer Assisted Language Learning*, 1–47. https://doi.org/10.1080/09588221.2022.2052905

Steinert, S., Krupp, L., Avila, K. E., Janssen, A. S., Ruf, V., Dzsotjan, D., Schryver, C. D., Karolus, J., Ruzika, S., Joisten, K., Lukowicz, P., Kuhn, J., Wehn, N., & Küchemann, S. (2024). Lessons learned from designing an open-source automated feedback system for STEM education. *Education and Information Technologies*, 1–42. https://doi.org/10.1007/s10639-024-13025-y

Tharp, R. G., & Gallimore, R. (1988). *Rousing minds to life: Teaching, learning, and schooling in social context*. Cambridge University Press.

Tomasik, M. J., Helbling, L. A., & Moser, U. (2021). Educational gains of in-person vs. distance learning in primary and secondary schools: A natural experiment during the COVID-19 pandemic school closures in Switzerland. *International Journal of Psychology*, *56*(4), 566–576. https://doi.org/10.1002/ijop.12728

van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher–student interaction: A decade of research. *Educational Psychology Review*, *22*, 271–296. https://doi.org/10.1007/s10648-010-9127-6

Wang, F., Cheung, A. C., Chai, C. S., & Liu, J. (2024). Development and validation of the perceived interactivity of learner-AI interaction scale. *Education and Information Technologies*, 1–32. https://doi.org/10.1007/s10639-024-12963-x

Wang, X., Liu, Q., Pang, H., Tan, S. C., Lei, J., Wallace, M. P., & Li, L. (2023). What matters in AI-supported learning: A study of human-AI interactions in language learning using cluster analysis and epistemic network analysis. *Computers & Education*, *194*, 104703. https://doi.org/10.1016/j.compedu.2022.104703

Wapner, S., Kaplan, B., & Cohen, S. B. (1973). An organismic-developmental perspective for understanding transactions of men in environments. *Environment and Behavior*, *5*(3), 255.

Williamson, S. N. (2007). Development of a self-rating scale of self-directed learning. *Nurse Researcher*, *14*(2), 66–83. https://doi.org/10.7748/nr2007.01.14.2.66.c6022

Winters, S., Farnsworth, K., Berry, D., Ellard, S., Glazewski, K., & Brush, T. (2023). Supporting middle school students in a problem-based makerspace: Investigating distributed scaffolding. *Interactive Learning Environments*, *31*(6), 3396–3408. https://doi.org/10.1080/10494820.2021.1928709

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

Xia, Q., Chiu, T. K., Chai, C. S., & Xie, K. (2023). The mediating effects of needs satisfaction on the relationships between prior knowledge and self-regulated learning through artificial intelligence chatbot. *British Journal of Educational Technology*, *54*(4), 967–986. https://doi.org/10.1111/bjet.13305

Yuan, Y. (2023). An empirical study of the efficacy of AI chatbots for English as a foreign language learning in primary education. *Interactive Learning Environments*, 1–16. https://doi.org/10.1080/10494820.2023.2282112

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, *41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2

Zimmerman, B. J. (2011). Motivational sources and outcomes of self-regulated learning and performance. In D. H. Schunk, & B. Zimmerman (Eds.), *Handbook of self-regulation of learning and performance* (pp. 49–64). Routledge. https://doi.org/10.4324/9780203839010