ReCALL (2025), *37*: 3, 315–331
doi:10.1017/S0958344024000259

RESEARCH ARTICLE

Examining generative AI–mediated informal digital learning of English practices with social cognitive theory: a mixed-methods study

Lihang Guan*
The Education University of Hong Kong, China (s1136696@s.eduhk.hk)

Ellen Yue Zhang
The Education University of Hong Kong, China (yuezhang@eduhk.hk)

Michelle Mingyue Gu*
The Education University of Hong Kong, China (mygu@eduhk.hk)

Abstract

This study explores the integration of generative artificial intelligence (GenAI) in informal digital learning of English (IDLE) practices, focusing on its potential to enhance language learning outcomes and addressing the technological challenges language teachers face in utilising AI-based tools to facilitate second language acquisition. Based on the research context of IDLE and holistic learning ecology and drawing on the theoretical frameworks of technological pedagogical and content knowledge and social cognitive theory, we performed a mixed-methods investigation with an empirical experiment to assess the effectiveness of GenAI followed by semi-structured interviews. The results suggest that the GenAI-mediated IDLE practices effectively improve college students' oral proficiency in English from both technological and humanistic perspectives. However, results also indicate that the GenAI conversational partner alone is not adequate to provoke continuous extramural GenAI-mediated IDLE practices. We discuss the theoretical and pragmatic significance of GenAI-mediated IDLE in educational equity and reformation.

Keywords: generative AI; oral proficiency; social cognitive theory; technological pedagogical and content knowledge; TPACK; informal digital learning of English; IDLE; holistic learning ecology

1. Introduction

Technology and pedagogical changes are transforming language instruction. In this dynamic context, generative artificial intelligence (GenAI) in education, especially in informal language learning (Godwin-Jones, 2022), has become essential to improve instructional pedagogies and learning outcomes (Labadze, Grigolia & Machaidze, 2023). Considering the holistic learning ecology (Brown, 2000; Luckin, 2008), GenAI in formal instruction provides personalised learning opportunities that strengthen learner motivation and personal agency and can provide authentic,

Cite this article: Guan, L., Zhang, E.Y. & Gu, M.M. (2025). Examining generative AI–mediated informal digital learning of English practices with social cognitive theory: a mixed-methods study. *ReCALL* 37(3): 315–331. https://doi.org/10.1017/S0958344024000259

***Corresponding authors:** Emails: s1136696@s.eduhk.hk; mygu@eduhk.hk

© The Author(s), 2024. Published by Cambridge University Press on behalf of EUROCALL, the European Association for Computer-Assisted Language Learning. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S0958344024000259 Published online by Cambridge University Press

social-like interactions in extended or virtual reality for language learning beyond coded responses (Shadiev, Sun & Huang, 2019). GenAI technology has demonstrated unparalleled capabilities in providing detailed feedback compared with human teachers (Ai, 2017) and providing real-time translation for content and language integrated learning (CLIL; Liu & Chen, 2023). However, studies have also found that only some students can benefit from using GenAI technology in learning (e.g. Niloy, Akter, Sultana, Sultana & Rahman, 2024; Ou, Stöhr & Malmström, 2024). From the language and technology perspectives, moderators such as students' English proficiency (Liu & Chen, 2023) and digital literacy (Goldenthal, Park, Liu, Mieczkowski & Hancock, 2021) significantly influence the experimental results. Furthermore, from the humanistic perspective, studies have found that language teachers at all levels of education lack the willingness and capability to utilise GenAI-based language learning tools in classes (Godwin-Jones, 2023; Ou *et al.*, 2024; Yang, Kim, Lee & Shin, 2022).

Multiple frameworks have been developed to assist instructors in improving digital literacy and incorporating technology into education (Li & Lan, 2022; Ng, 2012). This paper adopts the framework of technological pedagogical and content knowledge (TPACK) due to its comprehensiveness: it incorporates teachers' technological literacy, pedagogical capabilities, and content knowledge (Koh & Chai, 2014). Existing studies using TPACK have reported that teachers who are not native to digital technologies commonly have difficulties in using technology for content and language teaching (Koh & Chai, 2014; Miguel-Revilla, Martínez-Ferreira & Sánchez-Agustí, 2020; Tondeur, Scherer, Siddiq & Baran, 2017), which is heavily relied on their insufficient technological knowledge to integrate digital technology into language teaching and learning (Celik, 2023).

Second language acquisition (SLA) involves more than just formal education. From the viewpoint of holistic learning ecology (Brown, 2000; Lai, Liu, Hu, Benson & Lyu, 2022; Luckin, 2008), SLA depends not only on formal language instructions but also on informal language learning practices after classes (Lee, 2019a). With GenAI, such after-class practices include a range of informal digital learning of English (IDLE) activities (Liu & Ma, 2024) that seep into the daily life of English as a foreign language (EFL) learning (Liu, Darvin & Ma, 2024a). Therefore, employing GenAI-mediated IDLE practices for English learning among students could overcome the resistance of educators to use GenAI and develop a more holistic technological integration into education that fits the directional requirements of global policies (Alghamdi & Holland, 2020; Lai & Jin, 2021).

In GenAI-mediated IDLE practices, GenAI is particularly beneficial for English speaking (Chen, 2024; Yang *et al.*, 2022). Compared with the traditional classroom, which limits EFL learners' in-class practice and interaction opportunities due to the large class sizes and limited class hours (Chen, 2024), GenAI can provide personalised feedback (Escalante, Pack & Barrett, 2023) and act as an authentic conversational partner (Yang *et al.*, 2022) to increase interaction frequency (Belda-Medina & Calvo-Ferrer, 2022). When students engage in English conversations with GenAI in extracurricular situations (a category of IDLE activities; see Section 2.1), researchers have found an improved willingness to communicate (Tai & Chen, 2023) based on reduced anxiety in speaking English (Kim & Su, 2024). Moreover, this approach enhances self-regulation during out-of-class learning (García Botero, Botero Restrepo, Zhu & Questier, 2021) by promoting metacognitive learning strategies (Saadati, Zeki & Vatankhah Barenji, 2023). Therefore, GenAI-mediated IDLE activities may be able to generate better outcomes than conventional teacher-led oral English learning, thus allowing teachers to incorporate digital technology into SLA for better learning outcomes regardless of their technical competence. Furthermore, past studies have suggested that IDLE practices influenced by others (e.g. teachers) could promote extramural IDLE with autonomy (Zhang & Liu, 2022, 2023) through enjoyment (Liu, Zhang & Zhang, 2024b). Based on this view, we proposed the following research questions for our mixed-methods study:

https://doi.org/10.1017/S0958344024000259 Published online by Cambridge University Press

1. Do GenAI-mediated IDLE practices improve college students' English speaking proficiency?
2. Do GenAI-mediated IDLE practices for speaking yield better post-test results than teacher-led speaking courses?
3. In the opinion of students, what factors contribute to the changes in their speaking results?
4. Do students who practise GenAI-mediated IDLE continue to perform such activities after the experiment? Why?

Answers to the research questions should lead to significant theoretical development and practical application. Theoretically, the research could expand the TPACK framework to involve out-of-class learning into the holistic learning ecology (Brown, 2000; Luckin, 2008). Acknowledging the fundamental process of observation and imitation in language learning (Bandura, 2014), we also highlight the need for a research focus on this learning process and the factors that influence it. Practically, our research could provide insights regarding a pragmatic method to integrate technology into college EFL education so that teachers can adapt to technologies for educational purposes without being restrained by their technological knowledge and literacy in and out of class. Regarding the terminology of AI and GenAI, we acknowledge that AI and GenAI have been used interchangeably in past literature, with "AI" being used to refer to ChatGPT, Copilot, Gemini and more, all of which are GenAI (e.g. Belda-Medina & Calvo-Ferrer, 2022; Liu *et al.*, 2024a).

2. Literature review

2.1 IDLE and GenAI

Originating from out-of-class autonomous learning, IDLE has emerged as a crucial research concept of computer-assisted language learning. This concept addresses a research gap in English learning and technology usage that happens autonomously outside of the classroom (Soyoof, Reynolds, Vazquez-Calvo & McLay, 2023). Based on Benson's (2011) four dimensions of out-of-class learning, Lee and Dressman (2018) identified IDLE as "self-directed, informal digital English learning independent of formal contexts" (p. 436). Under this definition, IDLE has been classified as "extracurricular" and "extramural", based on the closeness between IDLE activities and formal education (Lee, 2019b), as well as "receptive" and "productive", based on the materialistic nature of the IDLE activities (Lee & Drajati, 2019).

In the literature, researchers have mostly considered GenAI usage to be an IDLE practice. Past quantitative studies have found correlations between an individual's perception towards using technology for English learning and college students' GenAI usage as an IDLE practice (Liu & Ma, 2024). Moreover, factors such as peer support and enjoyment could influence students' GenAI usage behaviour (Liu *et al.*, 2024b). From the qualitative perspective, Liu *et al.* (2024a) suggested that when GenAI mediates IDLE practices, Chinese EFL students could seek guidance from technology; moreover, they self-reported that GenAI and teachers/tutors provided similar usefulness for EFL learning. There have been similar findings in other cultural backgrounds (e.g. Lee & Drajati, 2019; Ou *et al.*, 2024). For example, a large-scale qualitative investigation into Northern European students' GenAI usage detailed students' view of such technology as "my teacher" (Ou *et al.*, 2024: 6) for they rely on GenAI for knowledge consultation, demonstrating a consistence in GenAI usage behaviour across cultures. However, this does not suggest that language teachers can be replaced, but to accentuate the significance of GenAI in IDLE practices for EFL learners, especially in oral speaking where GenAI can be used as a conversational partner (Liu *et al.*, 2024a; Yang *et al.*, 2022).

Although GenAI's application in foreign language education has been investigated quantitatively and qualitatively (Liu & Ma, 2024; Ou *et al.*, 2024), this endeavour has been confined to the IDLE discipline. Since GenAI has the potential to transform education both in

https://doi.org/10.1017/S0958344024000259 Published online by Cambridge University Press

classes and out of classes (Meniado, 2023), how to bridge IDLE to teacher-involved education remains little answered. This study, by using an experimental design that alleviates teachers' inadequacies identified by the TPACK framework, could provide an alternative to facilitate EFL speaking acquisition.

2.2 Holistic learning ecology and GenAI

A learning ecology (Brown, 2000) is a holistic and adaptive system comprising rich resources, activities, and learning practices under formal and informal learning scenarios (Brown, 2000; Luckin, 2008). Such practices are particularly sensitive to technological advancements because technology enriches resources and interactions within learning practices (Brown, 2000; Lai *et al.*, 2022; Lai, Zhu & Gong, 2015; Luckin, 2008). GenAI provides students with a personalised conversational partner for practice and feedback when learning oral English (Ai, 2017; Yang *et al.*, 2022) and a simulated culturally sensitive environment that provides relatively authentic interactions (Shadiev, Wang, Chen, Gayevskaya & Borisov, 2024) that are otherwise hard to find in a foreign country.

From the humanistic perspective, GenAI technology motivates students to conduct autonomous IDLE practice (Lai *et al.*, 2022; Tai, 2024a, 2024b). Through the simulated conversational environment, students who adopt this technology feel more motivated to engage in the conversations (Yang *et al.*, 2022), leading to deep language learning (Wang, Su, & Yu, 2020). In a large-scale qualitative text analysis, Ou *et al.* (2024) found that students treat GenAI as a significant source of information, inspiration, and teaching, which bestows an identity of "my teacher" (Ou *et al.*, 2024: 6) onto GenAI tools. This finding further stresses the significant role of GenAI in the holistic learning ecology.

2.3 TPACK and GenAI

TPACK (Koehler, Mishra & Cain, 2013) provides a sound theoretical framework for understanding how teachers integrate technology, pedagogy, and content knowledge to support student learning (Sun, Ma, Zeng, Han & Jin, 2023). It emphasises the dynamic interplay between these three domains (Dong, Chai, Sang, Koh & Tsai, 2015) and highlights the importance of teachers' ability to effectively integrate technological tools and resources into students' language learning practices while maintaining a focus on pedagogical goals and content (Saubern, Henderson, Heinrich & Redmond, 2020).

In the TPACK framework, the domain of technological knowledge refers to understanding how different technologies can be used effectively in various educational settings (Greene & Jones, 2020). It contains three elements: knowledge of existing technologies (knowing the capabilities and limitations of existing technology for teaching and learning), skills in technology use (proficiency in using technological tools), and awareness of emerging technologies (keeping up to date with technology advancements; Adipat, 2021; Haleem, Javaid & Singh, 2022). Teachers are typically aware of GenAI's potential in language teaching and learning (Jiang, Jong, Lau, Chai & Wu, 2021; Ong & Annamalai, 2024) but have technical difficulties when integrating GenAI into education (Ong & Annamalai, 2024; Zhang, Zou, Cheng & Xie, 2022). Hence, teachers show a low commitment and capability to integrate technology into EFL education (Ping, 2022), despite knowing the multifaceted benefits of GenAI in SLA (Calvo & Hartle, 2024; Godwin-Jones, 2023). We addressed this issue by contemplating the effectiveness of GenAI-mediated IDLE practices to overcome such difficulties.

2.4 Social cognitive theory and GenAI

Foreign language acquisition is multifaceted, and several theories have been developed to explain the acquisition process from different perspectives. Krashen's (1992) input hypothesis focuses on

https://doi.org/10.1017/S0958344024000259 Published online by Cambridge University Press

the learning inputs and stresses on the necessity of i+1 input in SLA for effective language learning. Moreover, Swain's (2005) output hypothesis emphasises the significance of output practices in language learning that extends beyond the suitable learning input. Social cognitive theory (SCT) transcends the discourse of input and output by focusing on the usages and practice of the materials and practices (Bandura, 2014); therefore, we adopted it as the theoretical framework for this study.

SCT (Bandura, 1986) emphasises the interaction between individuals, their behaviour, and the environment in the process of learning and development (Bandura, 2014). According to this theory, individuals are not passive recipients of information; rather, they actively engage in the learning process by setting goals, monitoring their progress, and adjusting their behaviour based on feedback and reinforcement (Ibrahim, Clark, Reese & Shingles, 2020; Liu, Huang & Wang, 2014).

SCT emphasises the significance of modelling and imitation in language learning (Chen, 2014; Deng, Wang & Xu, 2022). EFL researchers have found that learners imitate language structures, pronunciation, and communication strategies through observation from and practice with authentic and authoritative sources (LaScotte, Meyers & Tarone, 2021; Li & Somlak, 2019; Sasaki & Takeuchi, 2010). These observations and practices are rooted in the students' self-efficacy (Zhou, Chiu, Dong & Zhou, 2023) and individuals' belief in their ability to succeed in specific tasks (Bandura, 2014). GenAI can promote self-efficacy in various ways (Tseng, Chen & Lin, 2023; Zhou *et al.*, 2023). From the technological perspective, Liu, Hou, Tu, Wang and Hwang (2023) suggested that immediate and personalised feedback facilitates EFL students' writing exercises and promotes their self-efficacy. From the humanistic perspective, Chang, Hwang and Gau (2022) argued that the students' general positive perception of GenAI technology, such as convenience in obtaining information and interest in using such technology, can enhance students' self-efficacy and academic performance.

2.5 SCT and TPACK

Using GenAI tools in language education can promote self-efficacy from both technological and humanistic perspectives (Liu *et al.*, 2023; Ou *et al.*, 2024), which in turn enhances the observation and imitation behaviours that influence SLA and academic performance (Bandura, 1986, 2014; Zhou *et al.*, 2023). Moreover, TPACK includes the skills that teachers should possess to integrate technology effectively to impart knowledge and stimulate learning (Greene & Jones, 2020; Sun *et al.*, 2023). Therefore, SCT could provide theoretical insights into TPACK from the perspective of the holistic learning ecology. Given that the purpose of education is to provoke learning (Robinson & Aronica, 2019), teachers' ability to use technology in education only partly constitutes the holistic learning ecology. Out-of-class autonomous learning of English (Lai *et al.*, 2015), called IDLE (Lee, 2019a), is also an essential component. It could utilise the technological knowledge of the digital native students and be carried out regardless of whether the teacher has limited technological knowledge (Ong & Annamalai, 2024). Thus, using SCT to investigate the effectiveness of GenAI-mediated IDLE practices to account for the challenging demand of teachers' technological knowledge in the TPACK framework could represent a significant step towards a more comprehensive theoretical understanding of the holistic learning ecology.

3. Methodology

To examine the role of GenAI activities on EFL learners' oral English proficiency levels and IDLE practices, we conducted an explanatory mixed-methods study comprising an experimental study supplemented with two rounds of follow-up qualitative interviews to explain the quantitative findings and to evaluate the behavioural sustainability. The experimental study used the pre- and post-test design and lasted 10 weeks. The pre- and post-tests adopted the International English

https://doi.org/10.1017/S0958344024000259 Published online by Cambridge University Press

Language Testing System (IELTS) speaking band descriptors for grading because of the communication-oriented nature of the IELTS speaking grading rubrics (Nakatsuhara, Inoue & Taylor, 2021).

3.1 Participants

This research initially included 48 undergraduate EFL students aged 18–21 years from a STEM-oriented institution in mainland China, divided into two groups of 24. One student dropped out of the experimental group owing to illness, leaving 24 students in the control group and 23 students in the experimental group. Among the 47 participants, 31 were men and 16 were women, which corresponds with the gender distribution at tech-oriented universities in China (Tencent Education, 2021). Based on the pre-test, there was no significant difference ($t = -1.88$, $df = 45$, $p = 0.851$) in English oral proficiency between the control group ($M = 5.188$, $SD = 0.548$) and the experimental group ($M = 5.217$, $SD = 0.540$). We recruited the participants through a rigorous process, with the inclusion criterion being that the participant had to have previous experience with GenAI to reduce mastery bias in the experiment (Ahn, Bong & Kim, 2017). Advertisements were posted in the university building designated for IELTS study to entice participation. We also encouraged the participants to refer others to the study. The participants received a complete experiment description and signed an informed consent form before starting.

3.2 Experimental procedures

The 10-week study comprised a 2-hour session for each group each week. We divided the participants randomly into the experimental and control group. Two experienced IELTS teachers who had scored 8.5 and 9 on the IELTS oral examination graded the pre- and post-tests, based on the IELTS speaking rubrics for pronunciation, fluency, grammar, and lexical resource (for detailed information, please refer to the "IELTS Speaking Band Descriptors" at https://ielts.org/). Before the start of the first week, the participants in the experimental group were trained on how to interact with the virtual companion (友伴) named Lucy and the digital English interpreter (英语翻译官) in iFlytek Spark (讯飞星火), a Chinese GenAI tool for academic purposes that individuals can interact with in English. We chose this virtual companion because it can generate communicative questions and responses for students to practise writing and speaking and provide feedback and sample answers that are personalised to each student's input. Moreover, the students were taught how to prompt Lucy to practise speaking, to ask for feedback, and to get sample answers when they were stuck. The training – which consisted of a brief demonstration, a student practice, and a technical consultation – lasted about 30 minutes on the first day of the experiment. Although the virtual companion may sometimes ask non-IELTS questions, the students in the experimental group had a printed question bank to ask Lucy to provide sample answers and feedback on their own answers. Because SCT describes modelling and imitation as the main ways for SLA (Chen, 2014; Deng *et al.*, 2022), we suggested that the students choose whichever feedback forms they prefer to model and imitate as a part of their IDLE practices. The typical interaction with Lucy for multimodal practice and feedback is shown in Figure 1. As the control group interacted with an impartial, experienced IELTS teacher who scored 9 on the IELTS oral examination, this group received no training in GenAI use. However, the students in the control group were encouraged to repeat the teacher's modifications of their answers in class.

During the 10-week experiment, the two groups gathered in two separate self-study rooms. The control group interacted with an IELTS teacher who did not grade the pre- or post-test. This teacher asked the students authentic IELTS oral examination questions, invited the participants to answer, and gave feedback on the answers. On the other hand, the experimental group interacted with the digital English interpreter, Lucy. There was no teacher present for the self-study sessions

https://doi.org/10.1017/S0958344024000259 Published online by Cambridge University Press

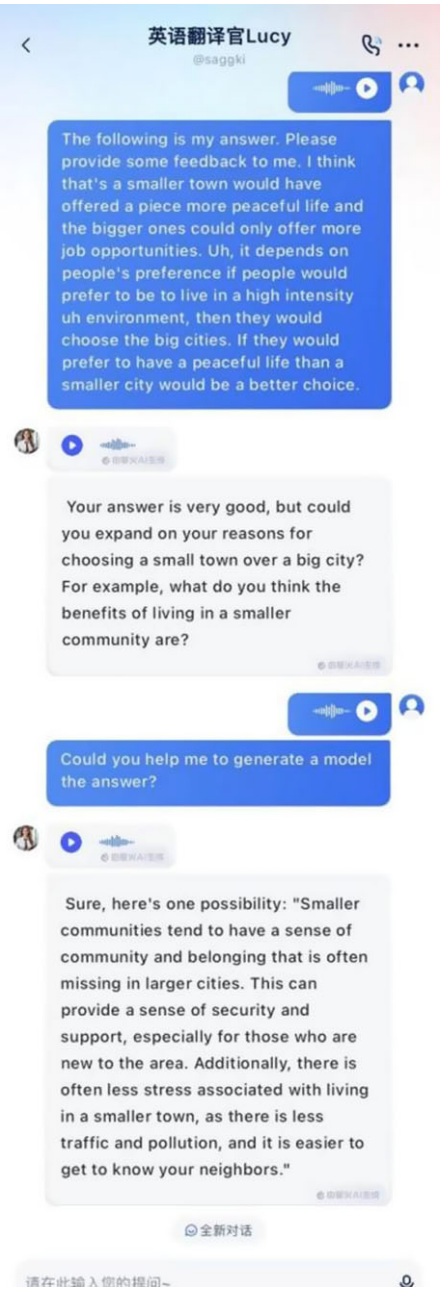


Figure 1. A typical example of practice with Lucy, including feedback and sample answers in the oral and written forms.

of the experimental group, only the first author taking attendance at the beginning and the end of each session.

The pre- and post-tests were administered at Week 0 and 11, respectively. The tests simulated the IELTS oral examination, in which one examiner asks questions and records the answers from each participant. For the pre- and post-tests, questions were randomly selected from the question banks. Of note, the same student was not asked the same questions for the pre- and post-tests. The examiner also wrote comments and graded the exam. Subsequently, an additional examiner

https://doi.org/10.1017/S0958344024000259 Published online by Cambridge University Press

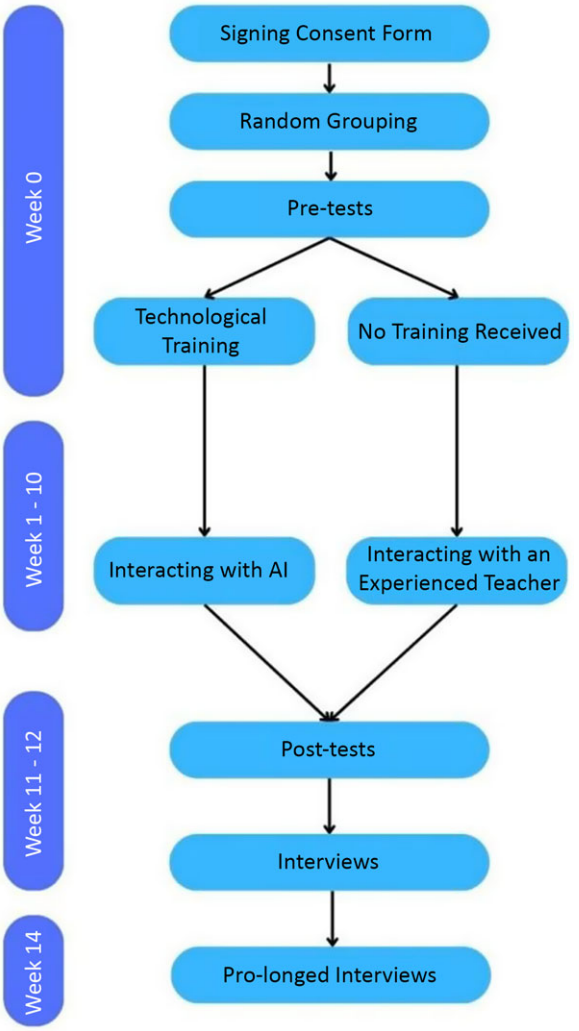


Figure 2. The experimental procedure.

played the recordings and double-checked the comments and, if needed, adjusted the grades. The inter-rater reliability was 0.93. Moreover, neither examiner took part in teaching the control group nor had any previous relationship with the participants. Figure 2 shows the experimental procedure.

3.3 Data collection and analysis

Apart from the quantitative data obtained from pre-tests and post-tests in Week 0 and 11, we collected qualitative data by interviewing the participants at two times. First, during Week 12, we interviewed seven control group participants and six experimental group participants who volunteered (a total of six women). Second, during Week 14, we interviewed 23 students in the experimental group to address RQ4. We conducted highly flexible semi-structured interviews (Brinkmann, 2020) to maximise the answers that students can give (Brinkmann, 2020; Green, Camilli & Elmore, 2012), thus facilitating qualitative data extraction. To encourage the

https://doi.org/10.1017/S0958344024000259 Published online by Cambridge University Press

Table 1. The pre- and post-test English oral proficiency for the control and experimental groups

	Pre-test		Post-test	
	M	*SD*	*M*	*SD*
Control	5.19	0.55	5.54**	0.62
Experimental	5.22	0.54	5.85**	0.35

** $p < .01$.

Table 2. Post-test English oral proficiency for the control and experimental groups

	M	*SD*	*t*	*p*	95% CI	
Control	5.54	0.62	−2.08	0.04	−0.60	−0.01
Experimental	5.85	0.35				

Note. CI = confidence interval.

participants to provide as much information as possible, we conducted the interviews in the participants' first language and subsequently translated their responses into English.

We analysed the quantitative data from the pre- and post-tests by calculating descriptive and inferential statistics with SPSS Statistics 28 to examine differences in speaking proficiency between the experimental and control groups. We used NVivo 12 to perform thematic analysis of the qualitative data and to identify recurring patterns and themes related to students' experiences and beliefs as justifications behind the quantitative findings. Specifically, we followed the five-step guidance of Braun and Clarke (2006) – data familiarisation, manual coding, thematic identification, theme reviews, and naming – to ensure the coherence, consistency, and presentation of the identified themes (Nowell, Norris, White & Moules, 2017).

4. Findings

4.1 Learning performance within the groups

As shown in Table 1, the pre- and post-test comparison indicated a significant improvement ($p < .01$) in oral proficiency in the control and experimental groups.

4.2 Learning performance between the groups

When we compared the post-test results between the groups, we found that the experimental group had better oral proficiency (Table 2), even though both groups showed similar oral proficiency before the experiment ($t = -1.88$, $df = 45$, $p = .85$). The post-test individual-sample *t*-test result indicated a significant difference ($p < .05$) between the control and experimental groups.

4.3 GenAI promotes learning performance through technological uniqueness

The interviews provide justification for GenAI's effects on English oral proficiency from technological and humanistic perspectives. From the technological viewpoint, one of the significant benefits GenAI offers is the number of practice opportunities it provides to students. In the first round of interviews, all 13 interviewees mentioned this point:

https://doi.org/10.1017/S0958344024000259 Published online by Cambridge University Press

> I wished the class size was a bit smaller, as I only had opportunities to answer about 4 to 5 questions in each session for her to correct my wrongs. (I4, control group)
>
> With iFlytec, the entire two hours is mine. It's like getting individual tuition without paying for anything. (I1, experimental group)

The students in the control group had fewer practice opportunities due to the class size, while the experimental group had many more opportunities. From the SCT perspective, students in the experimental group had more opportunities to observe and imitate proper English usage (in the first interview, 12 of the 13 interviewees hold this opinion), which contributed to the improved English learning performance:

> When the teacher explained new vocabulary to a classmate which I didn't know about, I'd write it down and try to use it in my talk. (I2, control group)
>
> You can ask for suggested answers or ways to develop your own answers from the virtual companion. And if you cannot understand it, it provides the texts in writing as well as in speaking so you can read them out loud. You can also mimic the intonations of the virtual companion, which is highly beneficial for my oral speaking. (I7, experimental group)

According to I7, the personalised feedback as well as the multimodality of the GenAI responses benefited SLA by enhancing modelling and imitation. Ten of the other interviewees agreed with this point. Furthermore, based on the qualitative data analysis, the multimodality feature of this particular GenAI tool may be especially helpful to students with low English proficiency before the experiment. A comparison between I7 (who scored a 4.5 on the pre-test) and I12 (who scored 6.0 in the pre-test) underscores this view:

> When I listen to my teachers in high school, I often needed a very long time to think about what she said and that made me not be able to follow up. But the GenAI can give me enough time to read and understand the content, and it gives the audio for the content as well so that I can model on it. I think this has helped my speaking. (I7, experimental group)
>
> It [the virtual companion] has given me a lot of suggestions regarding my answers. But when I have sought advice on using more complex sentence structures to answer some questions, the GenAI cannot provide many useful suggestions. Sometimes, it just explains why my answers are good and that is it. (I12, experimental group)

Based on the qualitative data, we found that the GenAI-based virtual companion technology can improve English oral proficiency by creating more opportunities to practise speaking and personalised feedback for students to hear and imitate. Moreover, such benefits may be more beneficial for students with low English proficiency who need help to develop and deliver their answers in English than those with above-average English proficiency who need help to construct more complex sentences when speaking.

4.4 GenAI promotes learning performance through humanistic perceptions

Another theme that emerged from the qualitative data is that GenAI benefits students' self-efficacy and learner agency by improving their willingness to communicate (mentioned by 18 of the 23 interviewed participants in the experimental group) and avoiding unconscious teacher bias (mentioned by 13 of the 23 interviewed participants in the experimental group), both of which enhance English oral proficiency. This benefit may be especially valuable for students with disadvantages related to their language proficiency or personality:

https://doi.org/10.1017/S0958344024000259 Published online by Cambridge University Press

> I feel like I didn't get enough chances to talk in class as the other students did. I think I'm being ignored because of my poor English skills. . . . The other student, a tall boy who's good at English, was asked a lot. I didn't even have half of his practice opportunities. (I8, control group)

> I'm a little introverted and I'm not good at English. So, when speaking in English to others, I would be nervous. . . . I feel like I'm being judged. . . . With an AI tool, I felt less nervous and could speak for more. . . . So I practised and improved. (I11, experimental group)

I8 may have been at a disadvantage due to the teacher's perception bias, which resulted in fewer learning opportunities, whereas I11 was encouraged by the GenAI tool because the technology improved their willingness to communicate. Hence, I8 was in a disadvantaged learning position while I11 was not. Therefore, we suggest that although specific personal characteristics may not necessarily benefit SLA, the influence of these characteristics could be mitigated by using GenAI as a mediator of IDLE practices.

4.5 GenAI alone is not enjoyable enough to foster extramural IDLE

During Week 14, the 23 students in the experimental group who we interviewed reported engaging in extramural IDLE with GenAI to some degree (mentioned by seven of the 23 interviewed participants). However, these seven participants struggled to sustain it, resulting in the abandonment of activities outside of their usual routine:

> When I was using GenAI at the dorm, it was easy to be disturbed by others and it was easy to disturb them. When I tried to find a room for self-study, it was difficult to find an entire classroom that has no one else in it. So, eventually, I dropped it before the experiment ended. (I7, experimental group)

The relinquishment of such behaviours may be caused by changes in learning environment (mentioned by five of the seven participants who reported using GenAI after the experiment) as well as a lack of enjoyment (mentioned by five of the seven participants who reported using GenAI after the experiment). In general, the participants in the experimental group found GenAI to be useful or practical, but not necessarily enjoyable:

> The instant feedback and the plentiful practice opportunities can help me to improve my oral speaking for sure, but it's a bit boring to study with it. . . . Because the content is not exciting and the feedback modality is not interactive enough. (I7, experimental group)

> Yes, I felt less nervous when talking to Lucy. But I didn't enjoy it. I actually found it to be tiresome because I had to control every conversation and sometimes Lucy couldn't understand my need and I had to think about different prompts to get what I needed, which is unlikely to happen with teachers. (I11, experimental group)

5. Discussion

5.1 The quantitative findings

We found that GenAI could improve EFL college students' English oral proficiency. This result corresponds with the previous findings that GenAI technology used as a conversational partner can benefit EFL student's language learning (e.g. Belda-Medina & Calvo-Ferrer, 2022; Liu *et al.*, 2024a; Yang *et al.*, 2022). As suggested by Yang *et al.* (2022), GenAI chatbots could facilitate students' language learning in informal settings by enhancing their understanding and ability to

https://doi.org/10.1017/S0958344024000259 Published online by Cambridge University Press

complete the tasks and, subsequently, improving their ability to use language correctly in future exams (Yang *et al.*, 2022). The responses and feedback based on large academic corpora and presented in forms of natural language align with SCT, which stresses the significance of modelling and imitation of appropriate language usage while learning a language (Chen, 2014; Deng *et al.*, 2022).

Our study provides empirical evidence that GenAI-mediated IDLE practices can lead to significantly better outcomes in English oral proficiency than learning in traditional teacher-centred classrooms. This finding challenges the role of teachers in Education 5.0, which focuses on "learner-centredness" (Meniado, 2023: 467) supported by "human-machine interaction technologies" (Meniado, 2023: 466). Technological advancements such as GenAI can add value and effectiveness to improve learning (Ong & Annamalai, 2024) and have the potential to revolutionise "the L2 teaching-learning ecosystem" (Meniado, 2023: 471), introducing new policies, theoretical conceptualisations, and pragmatic practices in this new era of education (Ng *et al.*, 2023). Our findings advocate for increased adoption of IDLE practices with GenAI technology.

5.2 The qualitative findings

We have demonstrated how GenAI promotes learning through technological perspectives, such as more learning opportunities and personalised feedback. By increasing students' chances to speak and by generating responses to their learning needs, GenAI strengthens the bond between modelling and imitation, underscoring the significance of SCT in SLA. In addition, GenAI may be especially beneficial for disadvantaged learners by providing constructive suggestions to answers and reducing negative aspects such as L2 anxiety and teacher perception bias.

The benefits of GenAI in providing constructive suggestions and enhancing learning motivation have been discussed previously (Chiu, 2023; Li & Kim, 2024). Teacher bias is a widely reported phenomenon (Copur-Gencturk, Cimpian, Lubienski & Thacker, 2020; Denessen, Hornstra, van den Bergh & Bijlstra, 2022; Dian & Triventi, 2021; Umansky & Dumont, 2021). However, it is not yet clear how GenAI can counteract or prevent teacher bias to improve students' learning. As suggested by Starck, Riddle, Sinclair and Warikoo (2020), "teachers are people too" (p. 273). Indeed, educators are also subjective to perception biases such as colour (Copur-Gencturk *et al.*, 2020), weight (Dian & Triventi, 2021), and social stereotypes (Denessen *et al.*, 2022). In practice, there are many factors for teachers to consider, such as students' flow of experience in teacher-centred education (Ateş & Garzón, 2022; Wagner, Holenstein, Wepf & Ruch, 2020). Therefore, even assuming every teacher possesses strong teacher agency, their teaching may not be equally beneficial to every student. From the perspective of achieving educational equality, we suggest that GenAI-mediated IDLE practices could promote equity for disadvantaged English learners to improve their oral proficiency. This step towards learner-centred SLA, facilitated by GenAI technologies in informal learning settings, ensures the holistic learning ecology of second language education.

Fully autonomous extramural IDLE practices are integral to achieve the holistic learning ecology for EFL learners. That said, our findings indicate that the use of GenAI as a conversational partner for IDLE practices may not provide students with adequate enjoyment, leading to inefficient intrinsic motivation (Deci & Ryan, 2012; Liu *et al.*, 2024b) for interest-based learning with GenAI. Although students find it useful to have GenAI as a conversational partner in EFL learning, these extramural IDLE practices may not be suitable for different learning environments or provide human likeness in the interactions. These issues need to be addressed before GenAI can be applied to attain the holistic learning ecology.

https://doi.org/10.1017/S0958344024000259 Published online by Cambridge University Press

6. Conclusion and suggestions for further research

From the TPACK framework, teachers' adequate technological knowledge is essential to benefit students' learning with digital technologies. Moreover, although the SLA processes involve multifaceted factors such as cognitive and social skill development and the integration of linguistic knowledge with cultural context, observing and imitating the language's proper usage has been argued to be one of the fundamental practices within SLA (Bandura, 2014). Therefore, it would be possible to present observation and imitation opportunities to learners using technological means without the influence of teachers' technological knowledge by providing GenAI-mediated oral IDLE practices. Our quantitative findings support this theoretical view. We found that GenAI technology could represent an advantageous alternative for students to practise speaking English and could yield a significant proficiency improvement between the pre- and post-tests (RQ1) and between the control and experimental groups (RQ2).

Based on the qualitative data, such improvements may derive from the technological and humanistic perspectives (RQ3). GenAI tools provide more practice opportunities and personalised feedback catered to the learner's personal learning needs. Humanistically, such technology could advance educational equity by preventing student characteristics from negatively interacting with and influencing the learning resources and environments. We recommend wider adoption of learner-centred GenAI-facilitated SLA in informal settings to achieve the holistic learning ecology. However, our qualitative results also suggest that students are not likely to continue such actions in the long run (RQ4). The experimental group participants generally found the use of GenAI in their self-study sessions to be helpful, but not enjoyable. Hence, they would be less willing to use GenAI as a conversational partner to improve their English fluency.

Under the guidance of SCT as the theoretical framework, we analysed the EFL learning as a practice of observation and imitation of authentic materials and practice dialogues. Acknowledging the multifaceted influences of SLA, it would be useful to investigate how GenAI contributes to input and output materials such as personalised feedback to gain a more holistic comprehension of technology and language acquisition. Moreover, we recommend further investigation regarding the potential strategies that could help transform extracurricular IDLE into extramural IDLE to gain a more holistic understanding of the dynamics of these activities.

Ethical statement and competing interests. This research is supported by JC_AI research fund (Project number: 02186), funded by the Education University of Hong Kong and the Hong Kong University of Science and Technology, Hong Kong, China. All participants were voluntary, and anonymity was ensured with codes assigned to each of the participants. Furthermore, the authors declare no competing interests. The authors declare no use of generative AI.

References

Adipat, S. (2021) Developing technological pedagogical content knowledge (TPACK) through technology-enhanced content and language-integrated learning (T-CLIL) instruction. *Education and Information Technologies*, 26(5): 6461–6477. https://doi.org/10.1007/s10639-021-10648-3

Ahn, H. S., Bong, M. & Kim, S. (2017) Social models in the cognitive appraisal of self-efficacy information. *Contemporary Educational Psychology*, 48: 149–166. https://doi.org/10.1016/j.cedpsych.2016.08.002

Ai, H. (2017) Providing graduated corrective feedback in an intelligent computer-assisted language learning environment. *ReCALL*, 29(3): 313–334. https://doi.org/10.1017/S095834401700012X

Alghamdi, J. & Holland, C. (2020) A comparative analysis of policies, strategies and programmes for information and communication technology integration in education in the Kingdom of Saudi Arabia and the Republic of Ireland. *Education and Information Technologies*, 25(6): 4721–4745. https://doi.org/10.1007/s10639-020-10169-5

Ateş, H. & Garzón, J. (2022) Drivers of teachers' intentions to use mobile applications to teach science. *Education and Information Technologies*, 27(2): 2521–2542. https://doi.org/10.1007/s10639-021-10671-4

Bandura, A. (1986) *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs: Prentice-Hall.

Bandura, A. (2014) Social-cognitive theory. In Ewen, R. B. (eds.), *An introduction to theories of personality* (7th ed.). New York: Psychology Press, 341–360.

https://doi.org/10.1017/S0958344024000259 Published online by Cambridge University Press

Belda-Medina, J. & Calvo-Ferrer, J. R. (2022) Using chatbots as AI conversational partners in language learning. *Applied Sciences*, 12(17): Article 8427. https://doi.org/10.3390/app12178427

Benson, P. (2011) Language learning and teaching beyond the classroom: An introduction to the field. In Benson, P. & Reinders, H. (eds.), *Beyond the language classroom*. London: Palgrave Macmillan UK, 7–16. https://doi.org/10.1057/9780230306790_2

Braun, V. & Clarke, V. (2006) Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2): 77–101. https://doi.org/10.1191/1478088706qp063oa

Brinkmann, S. (2020) Unstructured and semistructured interviewing. In Leavy, P. (eds.), *The Oxford handbook of qualitative research* (2nd ed.). New York: Oxford University Press, 424–456. https://doi.org/10.1093/oxfordhb/9780190847388.013.22

Brown, J. S. (2000) Growing up: Digital: How the web changes work, education, and the ways people learn. *Change*, 32(2): 11–20. https://doi.org/10.1080/00091380009601719

Calvo, L. C. S. & Hartle, L. C. (2024) Investigating pre-service teachers from Brazil and the US in a virtual exchange project: Benefits and challenges of student-selected and required technologies. *Education and Information Technologies*, 29(4): 5169–5187. https://doi.org/10.1007/s10639-023-12000-3

Celik, I. (2023) Towards intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. *Computers in Human Behavior*, 138: Article 107468. https://doi.org/10.1016/j.chb.2022.107468

Chang, C.-Y., Hwang, G.-J. & Gau, M.-L. (2022) Promoting students' learning achievement and self-efficacy: A mobile chatbot approach for nursing training. *British Journal of Educational Technology*, 53(1): 171–188. https://doi.org/10.1111/bjet.13158

Chen, Y.-C. (2014) An empirical examination of factors affecting college students' proactive stickiness with a web-based English learning environment. *Computers in Human Behavior*, 31: 159–171. https://doi.org/10.1016/j.chb.2013.10.040

Chen, Y.-C. (2024) Effects of technology-enhanced language learning on reducing EFL learners' public speaking anxiety. *Computer Assisted Language Learning*, 37(4): 789–813. https://doi.org/10.1080/09588221.2022.2055083

Chiu, T. K. F. (2023) The impact of generative AI (GenAI) on practices, policies and research direction in education: A case of ChatGPT and Midjourney. *Interactive Learning Environments*. Advance online publication. https://doi.org/10.1080/10494820.2023.2253861

Copur-Gencturk, Y., Cimpian, J. R., Lubienski, S. T. & Thacker, I. (2020) Teachers' bias against the mathematical ability of female, Black, and Hispanic students. *Educational Researcher*, 49(1): 30–43. https://doi.org/10.3102/0013189X19890577

Deci, E. L. & Ryan, R. M. (2012) Self-determination theory. In Van Lange, P. A. M., Kruglanski, A. W. & Higgins, E. T. (eds.), *Handbook of theories of social psychology* (Vol. 1). London: SAGE Publications, 416–436. https://doi.org/10.4135/9781446249215.n21

Denessen, E., Hornstra, L., van den Bergh, L. & Bijlstra, G. (2022) Implicit measures of teachers' attitudes and stereotypes, and their effects on teacher practice and student outcomes: A review. *Learning and Instruction*, 78: Article 101437. https://doi.org/10.1016/j.learninstruc.2020.101437

Deng, X., Wang, C. & Xu, J. (2022) Self-regulated learning strategies of Macau English as a foreign language learners: Validity of responses and academic achievements. *Frontiers in Psychology*, 13: Article 976330. https://doi.org/10.3389/fpsyg.2022.976330

Dian, M. & Triventi, M. (2021) The weight of school grades: Evidence of biased teachers' evaluations against overweight students in Germany. *PLOS ONE*, 16(2): Article e0245972. https://doi.org/10.1371/journal.pone.0245972

Dong, Y., Chai, C. S., Sang, G.-Y., Koh, J. H. L. & Tsai, C.-C. (2015) Exploring the profiles and interplays of pre-service and in-service teachers' technological pedagogical content knowledge (TPACK) in China. *Journal of Educational Technology & Society*, 18(1): 158–169. https://www.jstor.org/stable/jeductechsoci.18.1.158

Escalante, J., Pack, A. & Barrett, A. (2023) AI-generated feedback on writing: Insights into efficacy and ENL student preference. *International Journal of Educational Technology in Higher Education*, 20(1): Article 57. https://doi.org/10.1186/s41239-023-00425-2

García Botero, G., Botero Restrepo, M. A., Zhu, C. & Questier, F. (2021) Complementing in-class language learning with voluntary out-of-class MALL. Does training in self-regulation and scaffolding make a difference? *Computer Assisted Language Learning*, 34(8): 1013–1039. https://doi.org/10.1080/09588221.2019.1650780

Godwin-Jones, R. (2022) Partnering with AI: Intelligent writing assistance and instructed language learning. *Language Learning & Technology*, 26(2): 5–24. http://doi.org/10125/73474

Godwin-Jones, R. (2023) Presence and agency in real and virtual spaces: The promise of extended reality for language learning. *Language Learning & Technology*, 27(3): 6–26. https://hdl.handle.net/10125/73529

Goldenthal, E., Park, J., Liu, S. X., Mieczkowski, H. & Hancock, J. T. (2021) Not all AI are equal: Exploring the accessibility of AI-mediated communication technology. *Computers in Human Behavior*, 125: Article 106975. https://doi.org/10.1016/j.chb.2021.106975

Green, J. L., Camilli, G. & Elmore, P. B. (eds.) (2012) *Handbook of complementary methods in education research*. Washington, DC: Routledge.

https://doi.org/10.1017/S0958344024000259 Published online by Cambridge University Press

Greene, M. D. & Jones, W. M. (2020) Analyzing contextual levels and applications of technological pedagogical content knowledge (TPACK) in English as a second language subject area. *Educational Technology & Society*, 23(4): 75–88. https://www.jstor.org/stable/26981745

Haleem, A., Javaid, M. & Singh, R. P. (2022) An era of ChatGPT as a significant futuristic support tool: A study on features, abilities, and challenges. *BenchCouncil Transactions on Benchmarks, Standards and Evaluations*, 2(4): Article 100089. https://doi.org/10.1016/j.tbench.2023.100089

Ibrahim, A., Clark, K., Reese, M. J. & Shingles, R. (2020) The effects of a teaching development institute for early career researchers on their intended teaching strategies, course design, beliefs about instructors' and students' knowledge, and instructional self-efficacy: The case of the Teaching Institute at Johns Hopkins University. *Studies in Educational Evaluation*, 64: Article 100836. https://doi.org/10.1016/j.stueduc.2020.100836

Jiang, M. Y.-C., Jong, M. S.-Y., Lau, W. W.-F., Chai, C.-S. & Wu, N. (2021) Using automatic speech recognition technology to enhance EFL learners' oral language complexity in a flipped classroom. *Australasian Journal of Educational Technology*, 37(2): 110–131. https://doi.org/10.14742/ajet.6798

Kim, A. & Su, Y. (2024) How implementing an AI chatbot impacts Korean as a foreign language learners' willingness to communicate in Korean. *System*, 122: Article 103256. https://doi.org/10.1016/j.system.2024.103256

Koehler, M. J., Mishra, P. & Cain, W. (2013) What is technological pedagogical content knowledge (TPACK)? *Journal of Education*, 193(3): 13–19. https://doi.org/10.1177/002205741319300303

Koh, J. H. L. & Chai, C. S. (2014) Teacher clusters and their perceptions of technological pedagogical content knowledge (TPACK) development through ICT lesson design. *Computers & Education*, 70: 222–232. https://doi.org/10.1016/j.compedu.2013.08.017

Krashen, S. (1992) The input hypothesis: An update. In Alatis, J. E. (eds.), *Linguistics and language pedagogy: The state of the art*. Washington, DC: Georgetown University Press, 409–431.

Labadze, L., Grigolia, M. & Machaidze, L. (2023) Role of AI chatbots in education: Systematic literature review. *International Journal of Educational Technology in Higher Education*, 20: Article 56. https://doi.org/10.1186/s41239-023-00426-1

Lai, C. & Jin, T. (2021) Teacher professional identity and the nature of technology integration. *Computers & Education*, 175: Article 104314. https://doi.org/10.1016/j.compedu.2021.104314

Lai, C., Liu, Y., Hu, J., Benson, P. & Lyu, B. (2022) Association between the characteristics of out-of-class technology-mediated language experience and L2 vocabulary knowledge. *Language Learning & Technology*, 26(1): 1–24. https://hdl.handle.net/10125/73485

Lai, C., Zhu, W. & Gong, G. (2015) Understanding the quality of out-of-class English learning. *TESOL Quarterly*, 49(2): 278–308. https://doi.org/10.1002/tesq.171

LaScotte, D., Meyers, C. & Tarone, E. (2021) Voice and mirroring in SLA: Top-down pedagogy for L2 pronunciation instruction. *RELC Journal*, 52(1): 144–154. https://doi.org/10.1177/0033688220953910

Lee, J. S. (2019a) Quantity and diversity of informal digital learning of English. *Language Learning & Technology*, 23(1): 114–126. http://hdl.handle.net/10125/44675

Lee, J. S. (2019b) Informal digital learning of English and second language vocabulary outcomes: Can quantity conquer quality? *British Journal of Educational Technology*, 50(2): 767–778. https://doi.org/10.1111/bjet.12599

Lee, J. S. & Drajati, N. A. (2019) English as an international language beyond the ELT classroom. *ELT Journal*, 73(4): 419–427. https://doi.org/10.1093/elt/ccz018

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://www.jstor.org/stable/44986999

Li, L. & Kim, M. (2024) It is like a friend to me: Critical usage of automated feedback systems by self-regulating English learners in higher education. *Australasian Journal of Educational Technology*, 40(1): 1–18. https://doi.org/10.14742/ajet.8821

Li, P. & Lan, Y.-J. (2022) Digital language learning (DLL): Insights from behavior, cognition, and the brain. *Bilingualism: Language and Cognition*, 25(3): 361–378. https://doi.org/10.1017/S1366728921000353

Li, Y. & Somlak, T. (2019) The effects of articulatory gestures on L2 pronunciation learning: A classroom-based study. *Language Teaching Research*, 23(3): 352–371. https://doi.org/10.1177/1362168817730420

Liu, C., Hou, J., Tu, Y.-F., Wang, Y. & Hwang, G.-J. (2023) Incorporating a reflective thinking promoting mechanism into artificial intelligence-supported English writing environments. *Interactive Learning Environments*, 31(9): 5614–5632. https://doi.org/10.1080/10494820.2021.2012812

Liu, G. L., Darvin, R. & Ma, C. (2024a) Exploring AI-mediated informal digital learning of English (AI-IDLE): A mixed-method investigation of Chinese EFL learners' AI adoption and experiences. *Computer Assisted Language Learning*. Advance online publication. https://doi.org/10.1080/09588221.2024.2310288

Liu, G. L., Zhang, Y. & Zhang, R. (2024b) Examining the relationships among motivation, informal digital learning of English, and foreign language enjoyment: An explanatory mixed-method study. *ReCALL*, 36(1): 72–88. https://doi.org/10.1017/S0958344023000204

https://doi.org/10.1017/S0958344024000259 Published online by Cambridge University Press

Liu, G. & Ma, C. (2024) Measuring EFL learners' use of ChatGPT in informal digital learning of English based on the technology acceptance model. *Innovation in Language Learning and Teaching*, 18(2): 125–138. https://doi.org/10.1080/17501229.2023.2240316

Liu, P.-L. & Chen, C.-J. (2023) Using an AI-based object detection translation application for English vocabulary learning. *Educational Technology & Society*, 26(3): 5–20. https://www.jstor.org/stable/48734318

Liu, S., Huang, J. L. & Wang, M. (2014) Effectiveness of job search interventions: A meta-analytic review. *Psychological Bulletin*, 140(4): 1009–1041. https://doi.org/10.1037/a0035923

Luckin, R. (2008) The learner centric ecology of resources: A framework for using technology to scaffold learning. *Computers & Education*, 50(2): 449–462. https://doi.org/10.1016/j.compedu.2007.09.018

Meniado, J. C. (2023) Digital language teaching 5.0: Technologies, trends and competencies. *RELC Journal*, 54(2): 461–473. https://doi.org/10.1177/00336882231160610

Miguel-Revilla, D., Martínez-Ferreira, J. M. & Sánchez-Agustí, M. (2020) Assessing the digital competence of educators in social studies: An analysis in initial teacher training using the TPACK-21 model. *Australasian Journal of Educational Technology*, 36(2): 1–12. https://doi.org/10.14742/ajet.5281

Nakatsuhara, F., Inoue, C. & Taylor, L. (2021) Comparing rating modes: Analysing live, audio, and video ratings of IELTS speaking test performances. *Language Assessment Quarterly*, 18(2): 83–106. https://doi.org/10.1080/15434303.2020.1799222

Ng, D. T. K., Lee, M., Tan, R. J. Y., Hu, X., Downie, J. S. & Chu, S. K. W. (2023) A review of AI teaching and learning from 2000 to 2020. *Education and Information Technologies*, 28(7): 8445–8501. https://doi.org/10.1007/s10639-022-11491-w

Ng, W. (2012) Can we teach digital natives digital literacy? *Computers & Education*, 59(3): 1065–1078. https://doi.org/10.1016/j.compedu.2012.04.016

Niloy, A. C., Akter, S., Sultana, N., Sultana, J. & Rahman, S. I. U. (2024) Is Chatgpt a menace for creative writing ability? An experiment. *Journal of Computer Assisted Learning*, 40(2): 919–930. https://doi.org/10.1111/jcal.12929

Nowell, L. S., Norris, J. M., White, D. E. & Moules, N. J. (2017) Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*, 16(1): 1–13. https://doi.org/10.1177/1609406917733847

Ong, Q. K. L. & Annamalai, N. (2024) Technological pedagogical content knowledge for twenty-first century learning skills: The game changer for teachers of industrial revolution 5.0. *Education and Information Technologies*, 29(2): 1939–1980. https://doi.org/10.1007/s10639-023-11852-z

Ou, A. W., Stöhr, C. & Malmström, H. (2024) Academic communication with AI-powered language tools in higher education: From a post-humanist perspective. *System*, 121: Article 103225. https://doi.org/10.1016/j.system.2024.103225

Ping, W. (2022) Revisiting English as a foreign language teachers' professional identity and commitment in social media-focused professional development. *Frontiers in Psychology*, 13: Article 992038. https://doi.org/10.3389/fpsyg.2022.992038

Robinson, K. & Aronica, L. (2019) *You, your child, and school: Navigate your way to the best education*. New York: Penguin Books.

Saadati, Z., Zeki, C. P. & Vatankhah Barenji, R. (2023) On the development of blockchain-based learning management system as a metacognitive tool to support self-regulation learning in online higher education. *Interactive Learning Environments*, 31(5): 3148–3171. https://doi.org/10.1080/10494820.2021.1920429

Sasaki, A. & Takeuchi, O. (2010) EFL students' vocabulary learning in NS-NNS e-mail interactions: Do they learn new words by imitation? *ReCALL*, 22(1): 70–82. https://doi.org/10.1017/S0958344009990206

Saubern, R., Henderson, M., Heinrich, E. & Redmond, P. (2020) TPACK – Time to reboot? *Australasian Journal of Educational Technology*, 36(3): 1–9. https://doi.org/10.14742/ajet.6378

Shadiev, R., Sun, A. & Huang, Y.-M. (2019) A study of the facilitation of cross-cultural understanding and intercultural sensitivity using speech-enabled language translation technology. *British Journal of Educational Technology*, 50(3): 1415–1433. https://doi.org/10.1111/bjet.12648

Shadiev, R., Wang, X., Chen, X., Gayevskaya, E. & Borisov, N. (2024) Research on the impact of the learning activity supported by 360-degree video and translation technologies on cross-cultural knowledge and attitudes development. *Education and Information Technologies*, 29(7): 7759–7791. https://doi.org/10.1007/s10639-023-12143-3

Soyoof, A., Reynolds, B. L., Vazquez-Calvo, B. & McLay, K. (2023) Informal digital learning of English (IDLE): A scoping review of what has been done and a look towards what is to come. *Computer Assisted Language Learning*, 36(4): 608–640. https://doi.org/10.1080/09588221.2021.1936562

Starck, J. G., Riddle, T., Sinclair, S. & Warikoo, N. (2020) Teachers are people too: Examining the racial bias of teachers compared to other American adults. *Educational Researcher*, 49(4): 273–284. https://doi.org/10.3102/0013189X20912758

Sun, J., Ma, H., Zeng, Y., Han, D. & Jin, Y. (2023) Promoting the AI teaching competency of K-12 computer science teachers: A TPACK-based professional development approach. *Education and Information Technologies*, 28(2): 1509–1533. https://doi.org/10.1007/s10639-022-11256-5

Swain, M. (2005) The output hypothesis: Theory and research. In Hinkel, E. (ed.), *Handbook of research in second language teaching and learning*. New York: Routledge, 471–483.

https://doi.org/10.1017/S0958344024000259 Published online by Cambridge University Press

Tai, T.-Y. (2024a) Comparing the effects of intelligent personal assistant-human and human-human interactions on EFL learners' willingness to communicate beyond the classroom. *Computers & Education*, 210: Article 104965. https://doi.org/10.1016/j.compedu.2023.104965

Tai, T.-Y. (2024b) Effects of intelligent personal assistants on EFL learners' oral proficiency outside the classroom. *Computer Assisted Language Learning*, 37(5–6): 1281–1310. https://doi.org/10.1080/09588221.2022.2075013

Tai, T.-Y. & Chen, H. H.-J. (2023) The impact of Google Assistant on adolescent EFL learners' willingness to communicate. *Interactive Learning Environments*, 31(3): 1485–1502. https://doi.org/10.1080/10494820.2020.1841801

Tencent Education (2021, September 6) *Women tongjile bai yu suo gaoxiao 2021 xinsheng dashuju, faxian zhexie xuexiao nannv bili chaju da* [We compiled big data on first-year students from over a hundred universities in 2021 and found that there is a significant gender ratio disparity in these schools]. Tencent. https://new.qq.com/rain/a/20210906A03GYL00

Tondeur, J., Scherer, R., Siddiq, F. & Baran, E. (2017) A comprehensive investigation of TPACK within pre-service teachers' ICT profiles: Mind the gap! *Australasian Journal of Educational Technology*, 33(3). https://doi.org/10.14742/ajet.3504

Tseng, Y.-C., Chen, M.-R. A. & Lin, Y.-H. (2023) An investigation of the effects of EFL students' self-efficacy in an asynchronous online course with interactive contents. *Educational Technology & Society*, 26(4): 1–13. https://www.jstor.org/stable/48747517

Umansky, I. M. & Dumont, H. (2021) English learner labeling: How English learner classification in kindergarten shapes teacher perceptions of student skills and the moderating role of bilingual instructional settings. *American Educational Research Journal*, 58(5): 993–1031. https://doi.org/10.3102/0002831221997571

Wagner, L., Holenstein, M., Wepf, H. & Ruch, W. (2020) Character strengths are related to students' achievement, flow experiences, and enjoyment in teacher-centered learning, individual, and group work beyond cognitive ability. *Frontiers in Psychology*, 11: Article 1324. https://doi.org/10.3389/fpsyg.2020.01324

Wang, D., Su, J. & Yu, H. (2020) Feature extraction and analysis of natural language processing for deep learning English language. *IEEE Access*, 8: 46335–46345. https://doi.org/10.1109/ACCESS.2020.2974101

Yang, H., Kim, H., Lee, J. H. & Shin, D. (2022) Implementation of an AI chatbot as an English conversation partner in EFL speaking classes. *ReCALL*, 34(3): 327–343. https://doi.org/10.1017/S0958344022000039

Zhang, R., Zou, D., Cheng, G. & Xie, H. (2022) Implementing technology-enhanced collaborative writing in second and foreign language learning: A review of practices, technology and challenges. *Education and Information Technologies*, 27(6): 8041–8069. https://doi.org/10.1007/s10639-022-10941-9

Zhang, Y. & Liu, G. (2022) Revisiting informal digital learning of English (IDLE): A structural equation modeling approach in a university EFL context. *Computer Assisted Language Learning*. Advance online publication. https://doi.org/10.1080/09588221.2022.2134424

Zhang, Y. & Liu, G. L. (2023) Examining the impacts of learner backgrounds, proficiency level, and the use of digital devices on informal digital learning of English: An explanatory mixed-method study. *Computer Assisted Language Learning*. Advance online publication. https://doi.org/10.1080/09588221.2023.2267627

Zhou, S., Chiu, M. M., Dong, Z. & Zhou, W. (2023) Foreign language anxiety and foreign language self-efficacy: A meta-analysis. *Current Psychology*, 42(35): 31536–31550. https://doi.org/10.1007/s12144-022-04110-x

About the authors

Ellen Yue Zhang is an assistant professor in the Department of English Language Education at the Education University of Hong Kong. Her research interests include L2 motivation, identity and investment, CALL, IDLE, and critical pedagogies. She has published in *Computer Assisted Language Learning*; *Journal of Multilingual and Multicultural Development*; *TESOL Quarterly*; *System*; *ReCALL*; *Journal of Language, Identity, and Education*; *Language Awareness*; and *Chinese Journal of ESP*.

Mingyue Michelle Gu is a professor and the dean of the Graduate School at the Education University of Hong Kong. Her research interests include E-medium instruction in higher education, multilingualism and mobility, family language policy, and identity and digital literacies studies, and she has published widely in these fields. She is listed as one of the the world's top 2% scientists by Stanford University (2022).

Lihang Guan is a PhD student in the Department of English Language Education at the Education University of Hong Kong. His research interests include computer-assisted language learning (CALL), informal digital learning of English, and AI in education.

https://doi.org/10.1017/S0958344024000259 Published online by Cambridge University Press