frontiers in Psychology

TYPE Original Research
PUBLISHED 23 November 2026
DOI 10.3389/fpsyg.2026.1999726

Check for updates

OPEN ACCESS

EDITED BY
Mehdi Solhi,
Istanbul Medipol University, Türkiye

REVIEWED BY
Chuandai Qiao,
Hubei University of Automotive Technology, China
Maram Othman Almaneea,
Imam Muhammad ibn Saud Islamic University, Saudi Arabia

*CORRESPONDENCE
Jiao Xiong
xiongjiao@cqut.edu.cn

RECEIVED 15 September 2026
REVISED 26 October 2026
ACCEPTED 05 November 2026
PUBLISHED 23 November 2026

CITATION
Du Y, Reynolds BL, Li J, Guo C and Xiong J (2026) Community of inquiry presences, GenAI use, and metacognitive regulation in GenAI-mediated L2 learning: evidence from SEM and fsQCA.
Front. Psychol. 17:1999726.
doi: 10.3389/fpsyg.2026.1999726

COPYRIGHT
© 2026 Du, Reynolds, Li, Guo and Xiong. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Community of inquiry presences, GenAI use, and metacognitive regulation in GenAI-mediated L2 learning: evidence from SEM and fsQCA

Yunfei Du[1,2,3], Barry Lee Reynolds[2,3,4], Jiajun Li[2], Chaoyu Guo[5] and Jiao Xiong[6]*

[1]School of Languages and Communication Studies, Chongqing University of Technology, Chongqing, China, [2]Faculty of Education, University of Macau, Taipa, Macao SAR, China, [3]Centre for Data Science, University of Macau, Taipa, Macao SAR, China, [4]Centre for Cognitive and Brain Sciences, University of Macau, Taipa, Macao SAR, China, [5]School of Mathematical Sciences, Chongqing University of Technology, Chongqing, China, [6]School of Management, Chongqing University of Technology, Chongqing, China

Generative artificial intelligence (GenAI) has made language practice readily accessible. Yet access alone tells us little about how learners engage with the technology and regulate their learning. Drawing on the Community of Inquiry (CoI) framework, this study examined the relationships among perceived social presence (SP), cognitive presence (CP), teaching presence (TP), GenAI use for second language (L2) learning (GLU), L2 speaking self-efficacy (LSS), student engagement (SE), and metacognitive regulation tendency (MRT). Survey data from 1,160 Chinese university students across seven universities were analyzed using covariance-based structural equation modeling (SEM) and fuzzy-set qualitative comparative analysis (fsQCA). SEM revealed strong interrelationships among the three CoI presences and a positive association between their shared dimension and GLU. Within the hypothesized model, GLU was indirectly associated with MRT through LSS (β = 0.108) and SE (β = 0.477). However, these indirect associations were sensitive to alternative model specifications, particularly the pathway involving student engagement. fsQCA identified the joint presence of SP, CP, and TP as a configuration associated with high GLU (consistency = 0.943; coverage = 0.853). High MRT was associated with two configurations, GLU combined with SE and LSS combined with SE, both highlighting engagement as a core condition. These findings extend the CoI framework to GenAI-mediated L2 learning and underscore the relevance of learners' psychological experiences to self-reported metacognitive regulation. GenAI may make learning opportunities abundant, while engagement could help explain how learners relate those opportunities to regulating their learning.

KEYWORDS
chatbots, community of inquiry, fsQCA, generative AI, metacognitive regulation, second language learning, speaking self-efficacy, student engagement

1 Introduction

Generative artificial intelligence (GenAI) has moved from novelty to a new infrastructure for language learning (Du and Gao, 2022; Jeon et al., 2023). In second language (L2) learning contexts, GenAI chatbots now offer multi-turn dialogue, adaptive feedback, and on-demand conversational practice at a scale difficult to reproduce in many classrooms (Lin et al., 2025; Timpe-Laughlin and Dombi, 2020). Unsurprisingly, a growing body of research reports gains in fluency, lexical diversity, communicative confidence, and reduced anxiety (e.g., Chen et al., 2025; Fathi et al., 2024; Wang C. et al., 2024). Yet the pedagogical horizon of GenAI use for L2 learning extends beyond practice performance and affective comfort. Language can be a medium for thinking together, allowing learners to examine and reconsider ideas through interaction (Mercer, 2002). The sharper question, then, is not simply whether learners use GenAI, but how such use is associated with their metacognitive regulation and the learning conditions surrounding it (Du and Reynolds, 2025b; Hwang et al., 2018).

However, three critical gaps complicate this line of inquiry. First, much of the literature treats GenAI use as a homogeneous experience, glossing over the social, cognitive, and teaching conditions associated with how learners use the technology for L2 learning. The Community of Inquiry (CoI) framework offers a coherent lens for conceptualizing these conditions (Garrison et al., 1999; Garrison and Arbaugh, 2007), yet its application to self-directed GenAI use for L2 learning remains limited—and its notion of teaching presence (TP) remains contested when no human teacher is directly involved. Second, higher-order thinking has often been examined in writing and reading contexts, where argumentation is readily observable and assessable (e.g., Shang et al., 2025; Zhang et al., 2023). Beyond the expression of reasoning, higher-order thinking also involves metacognitive regulation, through which learners monitor their understanding, consider alternative approaches, and evaluate progress toward their goals (Hwang et al., 2018). Such regulation is relevant to GenAI-mediated L2 learning, where learners encounter diverse responses and feedback that invite reflection, evaluation, and strategic adjustment (Chen and Alibakhshi, 2026). Even so, L2 research has paid less attention to how metacognitive regulation is associated with GenAI use and related psychological factors, particularly L2 speaking self-efficacy and student engagement. Third, prior methodological approaches tend to assume linear, symmetric associations (e.g., Han et al., 2025). Educational phenomena, however, are rarely so tidy. Different combinations of conditions may be associated with similar outcomes, and the absence of a factor can be as consequential as its presence.

Closing these gaps calls for examining both structural associations and configurational patterns capable of capturing complexity. This study therefore examines how perceived social, cognitive, and teaching presences are associated with GenAI use for L2 learning, and how such use relates to metacognitive regulation tendency, directly and through L2 speaking self-efficacy and student engagement, while also identifying configurations associated with high levels of GenAI use and metacognitive regulation tendency. Learners' beliefs about their capability to speak (Bandura, 1997; Mills, 2014) and their investment of effort and attention (Fredricks et al., 2016; Mercer and Dörnyei, 2020) may be relevant to metacognitive regulation. In GenAI-mediated contexts, opportunities for practice and adaptive feedback may be related to learners' speaking self-efficacy (Guan et al., 2025), while personalized support may be associated with engagement (Han et al., 2025). These contextual and psychological states may be associated with greater self-reported metacognitive regulation. Given these considerations, two research questions guided the inquiry:

> *RQ1*: To what extent are social, cognitive, and teaching presences associated with GenAI use for L2 learning, and how is such use associated with metacognitive regulation tendency, directly and indirectly through L2 speaking self-efficacy and student engagement?
>
> *RQ2*: Which configurations of perceived social, cognitive, and teaching presences are associated with high GenAI use for L2 learning, and which combinations of GenAI use, L2 speaking self-efficacy, and student engagement are associated with high metacognitive regulation tendency?

2 Literature review

2.1 GenAI use for L2 learning

GenAI use for L2 learning (GLU) refers to learners' reported use of GenAI chatbots, such as ChatGPT, Doubao, and CallAnnie, for various English learning purposes, including communication practice, listening development, information gathering, cultural understanding, and the enhancement of learning motivation and confidence (Liu et al., 2025; Wan and Moorhouse, 2025; Wang C. et al., 2024). This conceptualization draws on the Actual Use dimension of AI-mediated informal digital learning of English (Liu et al., 2025), which examines how learners employ GenAI across different language learning activities. In the present study, GLU captures learners' reported use of GenAI for these purposes, rather than the frequency, duration, or interactional quality of their engagement with the technology.

Conversational use, as one form of this broader activity, aligns with the interactionist view that language competence grows through negotiation of meaning and feedback in conversation (Gass and Mackey, 2014; Long, 1996). L2 learners no longer just retrieve information or complete drills. Instead, they prompt, receive responses, clarify, and adjust their output in real-time (Kartal, 2024; Wu et al., 2025). The experience can resemble talk with a human peer or tutor, activating similar social-cognitive processes (Guan et al., 2025; Rahman and Tomy, 2024). In line with that, GenAI-mediated learning offers opportunities for contingency, immediacy of feedback, and personalization of responses, alongside learning activities that extend beyond conversation.

These capabilities provide several affordances relevant to long-standing challenges in L2 learning. First, GenAI expands access and language practice by providing flexible, on-demand opportunities for sustained oral production. Guan et al. (2025) identified flexible, repeatable practice as a decisive advantage of GenAI-mediated learning over teacher-led instruction, where practice opportunities were reported as comparatively fewer and unevenly distributed. Second, it delivers immediate, individualized feedback and scaffolding, which is difficult to replicate in large classes. Empirical studies suggest that instant feedback helped learners detect and correct issues in vocabulary (e.g., Wang and Reynolds, 2024) and grammar (e.g., Zheng et al., 2025). Third, it creates a low-pressure, non-judgmental interaction space that reduces affective barriers and encourages learners to speak more freely. Across studies, this stress-free interaction is linked to reduced speaking anxiety and diminished fear of negative feedback

(e.g., Dizon, 2026; Tai and Chen, 2024; Wang C. et al., 2024). Fathi et al. (2024), for instance, reported that sustained interaction with the Andy English Chatbot increased willingness to communicate by easing speaking anxiety and softening perceived evaluative pressure, which in turn encouraged more frequent and voluntary oral engagement. Fourth, GenAI-mediated learning can provide opportunities for reflection and metacognitive regulation as learners examine responses, consider alternative approaches, and monitor their understanding. For instance, Kartal (2024) observed that learners interacting with AI prompts often moved from simply providing answers to negotiating meaning, justifying claims, and re-evaluating assumptions, illustrating opportunities for reflective engagement beyond language production.

Despite the expanding research on GenAI-mediated L2 learning, much of the evidence often treats GenAI as a uniform treatment (e.g., Fathi et al., 2024; Yang et al., 2025; Zheng et al., 2025), leaving the contextual and psychological conditions associated with different learning uses insufficiently understood. More specifically, although recent research has examined disciplinary differences in university students' acceptance of GenAI (Du et al., 2026b), relatively little attention has been paid to how learners' perceived social conditions relate to their reported GenAI use, how perceived cognitive inquiry relates to the purposes for which learners use GenAI, or how instructional guidance embedded in GenAI chatbots relates to these learning activities. As a result, the field has limited explanatory leverage for understanding why learners differ in their reported GenAI use and associated metacognitive regulation tendencies. Accordingly, a more differentiated perspective is needed to examine GenAI use in relation to learners' perceived learning conditions and metacognitive regulation.

2.2 GenAI-mediated L2 learning as a community of inquiry

The Community of Inquiry (CoI) framework posits that a meaningful educational experience in online/blended environments arises from the interplay of social, cognitive, and teaching presences (Arbaugh, 2007; Garrison and Arbaugh, 2007). Social presence (SP) is the degree to which learners feel socially and emotionally connected with others in the environment (Wu et al., 2017). A strong SP creates a supportive climate for risk-taking and collaboration, which correlates positively with student engagement and satisfaction (Richardson et al., 2017). In this study, SP refers to perceived opportunities to express opinions and exchange ideas with the GenAI chatbot. Cognitive presence (CP) refers to the extent learners are able to construct and confirm meaning through sustained reflection and discourse (Garrison and Arbaugh, 2007; Wang X. et al., 2024). Associated with deeper levels of learning and reasoning, it manifests as learners engaging in exploring ideas, integrating knowledge, and applying critical thinking to solve problems or resolve confusion (Garrison et al., 2001). Teaching presence (TP) denotes the design, facilitation, and direction of cognitive and social processes toward achieving learning outcomes (Anderson et al., 2001). Traditionally, this is the role of the teacher who organizes content, guides discussion, and provides direct instruction where needed (Anderson et al., 2001). High TP has been tied to greater perceived learning and learner satisfaction (Caskurlu et al., 2020). When these three presences are optimally balanced, learners are more likely to have a rich, collaborative, and effective learning experience.

GenAI-mediated L2 learning activities, including conversational practice, inquiry, and feedback, can be conceptualized within the CoI framework, where social connection, cognitive meaning-making, and instructional guidance may be experienced through learners' interactions with a GenAI chatbot (Anderson et al., 2025; Ebadi and Amini, 2024; Wang X. et al., 2024). Anderson et al. (2025) proposed that GenAI can dynamically tune CoI-relevant conditions via personalized prompting, adaptive content delivery, and timely, individualized feedback—features that may support learner engagement in digital classrooms. Even without continuous human-teacher guidance, Wang X. et al. (2024) found that students reported higher cognitive and social presences interacted more frequently, expressed greater enjoyment, and achieved stronger learning outcomes, whereas a low-presence cluster exhibited the weakest engagement profile.

A key debate in positioning GenAI-mediated L2 learning within the CoI framework concerns the status of TP in self-directed learning contexts. Without a human teacher in the loop, TP may be perceived through the guidance, feedback, and task organization provided by the GenAI system (Wang X. et al., 2024). This shift raises questions about how instructional facilitation is experienced when guidance is partly embedded in an AI system (Caskurlu et al., 2020). Yet stronger TP does not necessarily yield better learning. When GenAI performs TP in an overly authoritative manner, it may constrain learner autonomy. Wang X. et al. (2024) for example, found that higher perceived TP from an AI English coach negatively predicted learning outcomes. They indicated that the coach's female teacher avatar conveyed implicit authority, which intimidated some learners and generated stress that offset the benefits of interaction (Wang X. et al., 2024). This suggests that TP in GenAI-mediated learning functions less as a fixed design feature than as a learner-interpreted experience. Empirical evidence further complicates the picture. Du and Reynolds (2025a) showed that learners often positioned Doubao as a peer rather than a teacher or authority figure. Such peer-like framing may lower affective barriers and encourage exploratory talk, even as it weakens conventional forms of instructional control. This tension motivates the present study to examine how CoI presences are perceived in self-directed GenAI use for L2 learning, where instructional guidance may be experienced differently from conventional teacher-led settings.

2.3 CoI presences and GenAI use for L2 learning

Drawing on the CoI framework, GenAI use for L2 learning (GLU) can be examined in relation to learners' perceived social, cognitive, and teaching conditions. SP captures learners' sense of having opportunities to express ideas and engage in meaningful exchanges, even when the interlocutor is artificial (Ebadi and Amini, 2024). In GenAI-mediated L2 learning, higher perceived SP—manifested through opportunities for expression and interaction—may reduce communicative inhibition and encourage learners to use GenAI for communication practice and other learning activities (Kim and Su, 2024; Wang C. et al., 2024). Prior research shows that a supportive social climate is associated with greater engagement and participation (Edumadze and Govender, 2024), and that GenAI chatbots can foster a stress-free environment in which learners feel safe to speak (Chuah and Kabilan, 2021). Empirically, Wang X. et al. (2024) found that learners who perceived stronger SP in interaction with an AI coach engaged more frequently. Therefore, we propose the following hypothesis:

> *H1*: SP is positively associated with GLU.

CP concerns learners' perceived involvement in inquiry, reflection, and meaning construction during interaction (Garrison et al., 2001; Wang et al., 2025). Learners with higher CP are more likely to treat GenAI as a cognitive catalyst rather than a passive information source, using it to explore ideas, connect information, and address learning problems (Cai et al., 2024). Empirical studies show that GenAI prompts often elicit higher-order dialogic moves, including justification of claims and re-evaluation of assumptions (Du and Reynolds, 2025a; Kartal, 2024; Zhang et al., 2025). Moreover, Wang X. et al. (2024) reported that CP was a significant predictor of learners' enjoyment and sustained use of an AI coach. Learners who perceive such opportunities for inquiry may be more inclined to use GenAI for purposeful language learning activities. Based on that, the following hypothesis is proposed:

> *H2*: CP is positively associated with GLU.

A certain level of TP may support GLU by maintaining focus, providing scaffolding, and organizing learning activities, as instructional guidance has been linked to engagement and perceived learning in CoI research (Caskurlu et al., 2020). TP encompasses instructional design, facilitation, and direct instruction (Anderson et al., 2001), yet emerging evidence in GenAI-mediated L2 learning suggests that TP may function as a double-edged condition. Wang X. et al. (2024) found that higher perceived TP from an AI English coach negatively predicted learning outcomes, attributing this to increased pressure and reduced interactional freedom when the AI adopted an authoritative teacher role. These findings suggest that the relationship between TP and GenAI use may depend on how guidance is calibrated. Rather than challenging the theoretical value of TP itself, such evidence points to the importance of distinguishing facilitative scaffolding from controlling instructional behavior. Accordingly, grounded in the CoI perspective, perceived instructional guidance may be positively associated with learners'use of GenAI for L2 learning, although its role may vary depending on how it is implemented. Based on that, the following hypothesis is proposed:

> *H3*: TP is positively associated with GLU.

2.4 Psychological associations between GenAI use and metacognitive regulation tendency

Although research has documented benefits of GenAI-mediated L2 learning for speaking practice (Jeon et al., 2023; Lin et al., 2025), language practice represents only one dimension of learners' GenAI use. In this regard, GenAI-mediated L2 learning also offers opportunities for learners to monitor their understanding, consider alternative responses, and evaluate their progress—processes associated with metacognitive regulation tendency (MRT; Hwang et al., 2018). However, the relationship between GLU and MRT may involve learners' psychological experiences during learning. Specifically, we consider two psychologically proximal factors: L2 speaking self-efficacy (LSS) and student engagement (SE). Through opportunities for practice and feedback, GenAI-mediated L2 learning may be associated with learners' speaking confidence and sustained behavioral and cognitive engagement, both of which are relevant to reflective learning and self-regulation (Han et al., 2025; Kim and Su, 2024).

LSS refers to learners' beliefs about their capability to perform oral communication tasks in the target language (Harris, 2025; Mills, 2014). Through frequent practice, calibrated task difficulty, and immediate, supportive feedback, learners may experience incremental success and reduced anxiety, which are key sources of self-efficacy as conceptualized by Bandura (1997). Empirical studies indicate that GenAI-mediated L2 speaking practice is associated with increased willingness to speak, reduced fear of errors, and heightened confidence (Wan and Moorhouse, 2025; Wang C. et al., 2024; Guan et al., 2025). These experiences may reinforce learners' sense of competence and control over L2 speaking, which may encourage them to monitor their progress, reconsider responses, and adjust learning strategies.

SE encompasses learners' behavioral participation, emotional investment, and cognitive effort in learning activities (Fredricks et al., 2016; Reeve and Tseng, 2011). In the present study, the focus is on behavioral and cognitive engagement during GenAI-mediated L2 learning. GLU may be associated with greater engagement when GenAI-mediated activities invite sustained participation and reflection. The immediacy and contingency of GenAI-mediated interaction may encourage persistence, while personalized responses can help sustain learners' attention to learning tasks (Elkhodr et al., 2023). At the cognitive level, GenAI prompts that invite elaboration, clarification, or justification can stimulate reflective and inquiry-oriented engagement (Kartal, 2024; Zhang et al., 2025). Empirical evidence suggests that learners interacting with GenAI report higher motivation, enjoyment, and self-regulatory behaviors, and that engagement may help account for associations between GenAI use and higher-order learning-related outcomes (Cai et al., 2024; Han et al., 2025). Therefore, two hypotheses are proposed:

> *H4a*: GLU is positively associated with LSS.
>
> *H4b*: GLU is positively associated with SE.

Both speaking self-efficacy and student engagement may be relevant to learners' metacognitive regulation. Learners with higher LSS are more willing to attempt complex expression, persist through difficulty, and employ metacognitive strategies, which may support their monitoring and adjustment of learning processes (Chen et al., 2025; Mills, 2014; Raoofi et al., 2012). Likewise, engaged learners, particularly those exhibiting strong cognitive engagement, are more likely to reflect, question, and integrate information, thereby engaging in processes relevant to metacognitive monitoring and regulation (Mercer and Dörnyei, 2020). Prior research links self-efficacy to deeper processing strategies (Phan, 2010; Raoofi et al., 2012) and engagement to reflective and critical-thinking-related learning activities (Fredricks et al., 2016; Han et al., 2025). Accordingly, both constructs are expected to be positively associated with MRT.

> *H5a*: LSS is positively associated with MRT.
>
> *H5b*: SE is positively associated with MRT.

These proposed relationships suggest that GLU may be indirectly associated with MRT through LSS and SE. Empirical studies have documented opportunities for metacognitive reflection and critical engagement with ideas in GenAI-mediated learning activities (Li and Kim, 2024; Zhang et al., 2023). Han et al. (2025) further highlighted the role of learner engagement in understanding the relationship

between technology use and critical thinking. Therefore, we propose that LSS and SE represent two theoretically distinct pathways linking GLU with MRT. A direct GLU–MRT association is also examined exploratorily in the structural model. In accordance with that, we proposed the following hypotheses:

> *H6*: GLU is indirectly associated with MRT.

2.5 The present study

Building on the above hypotheses, we proposed a conceptual model that relates perceived CoI presences to GLU and examines LSS and SE as intermediate variables in the model (Figure 1). The model assumes that the three presences are associated with LSS, SE, and MRT only through GLU. Accordingly, structural equation modeling (SEM) was used to estimate direct and indirect statistical associations among the seven constructs, and fuzzy-set qualitative comparative analysis (fsQCA) was used to examine which configurations of CoI presences were associated with high GLU and which configurations of GLU, LSS, and SE were associated with high MRT.

3 Methods

The study was situated in informal GenAI-mediated L2 learning settings in which learners were encouraged to engage with GenAI chatbots (e.g., Doubao, Kimi) for English-learning activities, including conversational practice, information gathering, and language development. Data were collected between January and September 2025, during which participants were encouraged to integrate GenAI-mediated L2 learning activities alongside their regular College English coursework.

3.1 Participants

Participants were recruited from seven universities across Southwest China through established collaboration with English teachers at each institution. They included non-English majors in College English courses and English majors, from science and engineering, the social sciences, and the humanities.

Two criteria were applied to screen the questionnaires: failure of an instructed-response attention check and identical ratings across all substantive items. An instructed-response item was embedded in the questionnaire to identify inattentive responding (Meade and Craig, 2012): "This is an attention check. Please select 7 (strongly agree)." According to the study records, 755 of the 2,335 submissions (32.3%) failed this check and were excluded; their individual records were unavailable for re-examination. Of the 1,580 available records, 420 contained identical ratings across all 46 substantive items and were excluded. The final analytical sample comprised 1,160 respondents (49.7% of the initial submissions). Completion time, Mahalanobis distance, and a GenAI usage threshold were not exclusion criteria.

To assess potential selection bias, we compared the 1,160 retained responses with the 420 available excluded responses. Small differences were observed in gender, residential background, and English examination categories (Cramér's $V = 0.065–0.103$). Construct-level comparisons also showed differences (Hedges' $g = -0.273$ to -0.553; Appendix Table E1). Given that these excluded responses were identified through a uniform-response criterion, the construct-score differences should be interpreted in relation to the screening procedure.

Participation was voluntary, anonymous, and unrelated to course grades. Students could withdraw at any time, and teachers did not see individual responses. All participants were adults and gave informed consent online. The study was approved by the Academic Committee of the School of Languages and Communication Studies, Chongqing

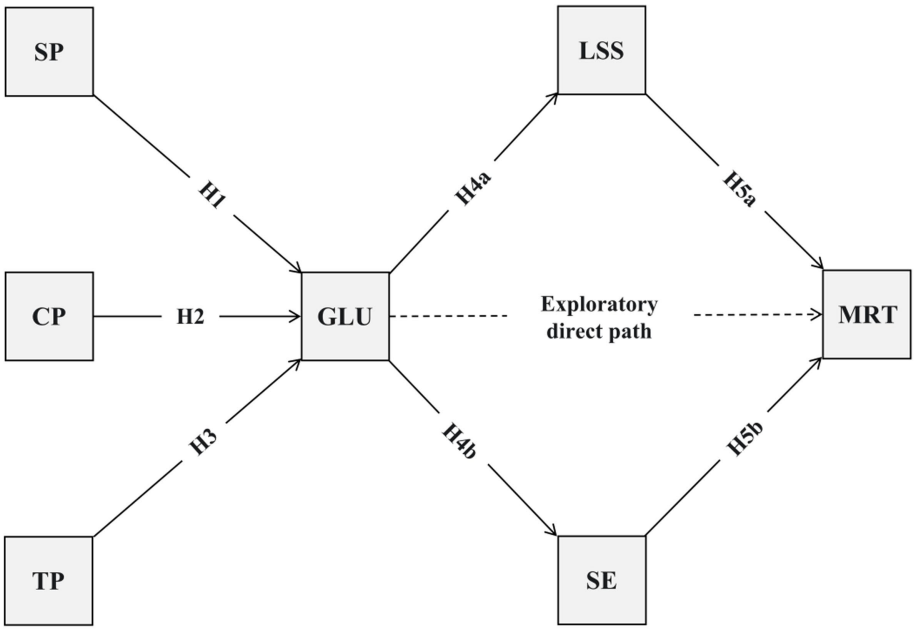

FIGURE 1
Conceptual model. SP, Social Presence; CP, Cognitive Presence; TP, Teaching Presence; GLU, GenAI use for L2 learning; LSS, L2 Speaking Self-efficacy; SE, Student Engagement; MRT, Metacognitive Regulation Tendency. Solid arrows represent hypothesized direct associations (H1–H5b). H6 concerns the statistical indirect associations of GLU with MRT through LSS and SE. The dashed GLU → MRT arrow represents an exploratory direct association.

University of Technology (approval no. CQUTSLCS 2025/001), which conducts ethics review for humanities and social science research with adult participants at this university participants profile in Table 1.

3.2 Instruments

The questionnaire was developed in English and administered in Chinese after translation and back-translation. All items used a 7-point scale (1 = strongly disagree, 7 = strongly agree). Of the original 46 items, 10 were excluded from the primary analysis on content grounds, while four problem-solving-related beliefs items were reserved for a supplementary outcome specification (Appendix A). Item selection was guided by conceptual alignment with the intended constructs and item-level measurement diagnostics. Items that reflected group collaboration, affective reactions, satisfaction, or conceptual content outside the target constructs were excluded from the corresponding scales. The resulting measurement model comprised 32 items across seven factors. Four items measuring problem-solving-related beliefs were examined separately as a supplementary outcome. To assess the sensitivity of the structural associations to item selection, an alternative 37-item specification was estimated using the same sample (Appendix Table C2). The complete item wording and selection decisions are provided in Appendix A.

SP (3 items), CP (4 items), and TP (3 items) were adapted from Law et al. (2019). SP captures opportunities to express opinions and exchange ideas with the chatbot. CP captures identifying problems, connecting information, reflecting, and integrating ideas. TP captures guidance, adaptive task allocation, and the chatbot's support for learning. GLU (6 items) was adapted from the Actual Use scale of Liu et al. (2025) and asks whether learners use GenAI chatbots for purposes such as practicing communication, improving listening, and gathering information. LSS was adopted from the Communicative Self-Efficacy Questionnaire developed by Harris (2025) and assessed learners' confidence in performing English tasks. SE (4 items) was adapted from Li et al. (2025) and covers behavioral and cognitive engagement, such as listening carefully, expressing opinions, and relating content to one's own experience. MRT (4 items) was adapted from Hwang et al. (2018), following the monitoring and evaluation statements of the Metacognitive Awareness Inventory (Schraw and Dennison, 1994), for example, "I ask myself periodically if I am meeting my goals." These MRT items assess domain-general self-reported regulation tendencies; their wording is not specific to L2 learning or GenAI use.

TABLE 1 Participants profile.

Characteristic	Category	*n*	%
Gender	Male	434	37.4
	Female	726	62.6
Residency	Urban	678	58.4
	Rural	482	41.6
English examination category	Below CET-4	433	37.3
	Passed CET-4	672	57.9
	Passed CET-6/TEM-4	55	4.7

n = 1,160. Of 1,580 retained source records, 420 gave the same rating on all 46 original substantive items and were excluded for uniform responding. CET, College English Test; TEM, Test for English Majors. Language categories are examination categories.

3.3 Data collection

The study examined GenAI-mediated L2 learning as it occurred alongside students' regular English coursework, rather than introducing a standardized instructional intervention. Data were collected in four independent survey waves in January 2025, March 2025, June 2025, and September 2025. During each data-collection period, collaborating English teachers encouraged students to engage in informal English practice with GenAI chatbots outside class. Following Du and Reynolds (2025a), students were advised to have at least one 15-min English conversation with a GenAI chatbot each week.

Since these activities were intended to extend opportunities for language practice beyond the classroom, they were open-ended and learner-directed. Students were encouraged to explore themes related to their coursework or topics of personal interest. They could initiate conversations, ask follow-up questions, seek clarification or language feedback, and explore ideas through exchanges with GenAI. Chatbots such as Doubao and Kimi were among the available tools, and no single platform was assigned to all participants.

At the end of each corresponding learning period, the online questionnaire was distributed via Wenjuanxing.[1] The four survey waves recruited separate course cohorts and were analyzed as pooled cross-sectional responses. As no participant-level identifiers were retained, repeat participation across waves could not be independently verified.

3.4 Data analysis

To examine the associations among perceived CoI presences, GLU, and MRT, we combined covariance-based structural equation modeling (CB-SEM) with fsQCA. CB-SEM was used to estimate the hypothesized structural relationships and indirect associations among latent constructs, whereas fsQCA examined combinations of conditions associated with high membership in the GLU and MRT sets. All analyses were performed in R 4.5.3, primarily using the lavaan, psych, semTools, and QCA packages.

3.4.1 Covariance-based structural equation modeling

The measurement and structural models were estimated using robust maximum likelihood (MLR) in lavaan (Rosseel, 2012). Model fit was evaluated using the scaled χ^2, robust CFI, TLI, RMSEA with its 90% confidence interval, and SRMR. Reliability and convergent validity were assessed using standardized factor loadings, Cronbach's α, composite reliability (CR), and average variance extracted (AVE). Discriminant validity was examined through latent correlations, the Fornell–Larcker criterion, and HTMT (Henseler et al., 2015), with 95% percentile confidence intervals obtained from 2,000 bootstrap resamples. The seven-factor model was also compared with alternative measurement models combining CP with TP or TP with SE. Potential common method variance was examined using an unrotated principal component analysis, a single-factor CFA, and an unmeasured latent method factor with equal loadings specified as orthogonal to the substantive factors (Podsakoff et al., 2024).

Measurement invariance across gender and English examination categories was evaluated using multigroup CFA. Configural, metric,

1 www.wjx.cn

and scalar invariance were examined sequentially by imposing equality constraints on factor loadings and item intercepts. Model comparisons considered scaled χ^2 difference tests alongside changes in robust CFI, RMSEA, and SRMR (Chen, 2007).

Indirect associations were examined using 5,000 ordinary case-bootstrap resamples under maximum likelihood estimation of the hypothesized structural model. Completely standardized coefficients and indirect products were recalculated within each resample, and percentile 95% confidence intervals were obtained.

Further sensitivity analyses constrained TP → GLU to zero, incorporated demographic covariates, constrained the exploratory GLU–MRT path, examined an alternative structural ordering, and applied WLSMV estimation treating the indicators as ordered categorical variables. Nested MLR specifications were compared using the Satorra and Bentler (2001) scaled difference test. Additional models specified a second-order CoI factor, included the equal-loading orthogonal method factor in the full SEM, and added six direct paths from SP/CP/TP to SE and MRT. A content-based 27-item sensitivity removed SE4, CP4, MRT2, GLU1, and GLU6, retaining at least three indicators per factor. New indirect associations were evaluated using 5,000 ordinary case ML bootstrap resamples, with standardization recomputed within each draw; method-factor intervals were treated as descriptive because resampling was unstable. Other specifications retained the same 32-item framework unless item exclusions were stated. We also re-estimated the primary model in all 1,580 attention-check-passing records and substituted problem-solving-related beliefs (PSB) for MRT (Appendix C).

3.4.2 fsQCA for configurational analysis

fsQCA was employed to examine configurations of conditions associated with high GLU and MRT (Ragin, 2008). Model A examined SP, CP, and TP as conditions for high GLU, while Model B examined GLU, LSS, and SE as conditions for high MRT. Construct scores (item means) were calibrated using the direct method with substantive anchors, following the calibration principles of Ragin (2008) and Schneider and Wagemann (2012). Anchors of 3 (full non-membership), 5 (crossover), and 6 (full membership) defined membership in high-level sets: endorsement at or below 3 was treated as low, scores above 5 as increasingly high, and 6 as strong endorsement. This operational choice followed Du et al. (2026a). The outer anchors mapped to 0.05/0.95 membership. Cases with membership scores exactly equal to 0.50 were retained in consistency and coverage calculations but excluded from crisp truth-table row assignment.

Necessity was assessed using a consistency threshold of 0.90, together with coverage and relevance of necessity (Schneider and Wagemann, 2012). Truth tables were constructed using a frequency threshold of 24 cases (approximately 2% of the sample), a consistency threshold of 0.80, and a proportional reduction in inconsistency (PRI) threshold of 0.70. Complex, parsimonious, and intermediate solutions were derived, with the intermediate solutions used for interpretation and core conditions identified through comparison with parsimonious solutions (Fiss, 2011). Directional expectations specified the presence of SP and CP for high GLU, leaving TP unspecified, and the presence of GLU, LSS, and SE for high MRT. All equally minimal solutions were retained. The negated outcomes were examined using corresponding absence expectations. Finally, 15 sensitivity specifications varied calibration anchors, frequency and consistency thresholds, PRI thresholds, and crossover handling one dimension at a time.

4 Results

4.1 Measurement model

The seven-factor CFA showed good fit (scaled $\chi^2(443) = 1159.82$, $p < 0.001$; robust CFI = 0.970, TLI = 0.967, RMSEA = 0.048, 90% CI [0.045, 0.052]; SRMR = 0.029). Standardized factor loadings ranged from 0.746 to 0.910, and Cronbach's α, CR, and AVE met conventional criteria (Table 2; Appendix Table B1). The three CoI presences, however, were highly correlated (r = 0.893–0.929). HTMT exceeded 0.90 for CP–TP (0.931) and TP–SE (0.932), indicating potential construct overlap (Appendix Table B2). The seven-factor model fit better than alternative models merging CP with TP or TP with SE (ΔBIC = 193.77 and 228.59, respectively; Appendix Table B3). However, the Fornell–Larcker criterion was not satisfied for five construct pairs (Appendix Table B2). Although the seven-factor model fit better than the merged alternatives, substantial overlap remained among several constructs, particularly CP–TP and TP–SE.

The first unrotated principal component accounted for 55.0% of the total item variance, while a single common factor explained 53.7%. The single-factor CFA showed poor fit (robust CFI = 0.664, RMSEA = 0.158). Adding an equal-loading latent method factor to the seven-factor model improved CFI slightly from 0.970 to 0.972, although SRMR increased from 0.029 to 0.035 (Appendix Table B3). These results suggest that shared method variance may contribute to the observed associations.

Multigroup CFA supported metric invariance across gender, scaled $\Delta\chi^2(25) = 37.21$, $p = 0.055$. The scalar constraints yielded a significant difference test ($p < 0.001$), although changes in approximate fit indices were small (ΔCFI = −0.0013). Across examination categories, metric and scalar comparisons were nonsignificant, but the smallest group ($n = 55$) introduced rank deficiency in the robust parameter covariance matrix. Sensitivity analyses involving the two larger examination groups provided additional evidence of measurement comparability (Appendix Table E2).

4.2 Structural associations

The hypothesized model showed acceptable fit (scaled $\chi^2(453) = 1680.06$, $p < 0.001$; robust CFI = 0.949, TLI = 0.944, RMSEA = 0.062, 90% CI [0.059, 0.066]; SRMR = 0.079). SP ($\beta = 0.519$) and CP ($\beta = 0.284$) were positively associated with GLU (both $p < 0.001$), consistent with H1 and H2 within this specification.

TABLE 2 Descriptive statistics and measurement quality ($N = 1,160$).

Construct	*k*	*M*	*SD*	*α*	CR	AVE
SP	3	5.44	1.09	0.918	0.918	0.790
CP	4	5.37	1.04	0.929	0.930	0.769
TP	3	5.44	1.06	0.900	0.901	0.753
GLU	6	5.33	1.16	0.958	0.958	0.791
LSS	8	5.16	1.09	0.957	0.957	0.739
SE	4	5.49	1.03	0.929	0.930	0.769
MRT	4	5.33	1.01	0.905	0.905	0.705

SP, social presence; CP, cognitive presence; TP, teaching presence; GLU, GenAI use for L2 learning; LSS, L2 speaking self-efficacy; SE, student engagement; MRT, metacognitive regulation tendency.

TP was not significantly associated with GLU ($\beta = 0.124$, $p = 0.137$), so H3 was not supported within this specification. GLU was positively associated with LSS ($\beta = 0.405$) and SE ($\beta = 0.796$), which were in turn associated with MRT ($\beta = 0.267$ and 0.600, respectively; all $p < 0.001$), supporting H4a–H5b within the hypothesized model. The exploratory direct GLU–MRT association was not significant ($\beta = 0.050$, $p = 0.341$). These conditional coefficients do not establish stable differences between the three CoI contributions. The model explained 80.9% of the variance in GLU, 16.4% in LSS, 63.3% in SE, and 59.6% in MRT (Table 3; Figure 2).

The TP–GLU path could be constrained to zero without significantly worsening fit under the primary MLR specification (scaled $\Delta\chi^2(1) = 2.22$, $p = 0.136$; Appendix Table C1). However, ordered-item WLSMV estimation yielded a positive TP–GLU coefficient ($\beta = 0.555$, $p < 0.001$), indicating estimator-dependent inference (Appendix Table C3).

The substantial overlap among the three CoI presences motivated a second-order specification, in which the overall CoI factor was strongly associated with GLU ($\beta = 0.909$). Although this model fit slightly worse than the primary specification, it highlighted the substantial shared variance among the presences. Excluding five potentially overlapping items likewise did not resolve concerns about construct distinctiveness (Appendix Tables C6–C8).

An alternative model allowing direct paths from SP, CP, and TP to SE and MRT significantly improved fit (scaled $\Delta\chi^2(6) = 294.90$, $p < 0.001$). Notably, the GLU–SE association reversed from positive ($\beta = 0.796$) to negative ($\beta = -0.116$), while TP was positively associated with both SE and MRT. Introducing a latent method factor also altered the hypothesized associations, although poor residual fit limited the interpretation of this specification. Together, these analyses indicate that the structural associations, particularly those involving GLU and engagement, were sensitive to model specification (Appendix Tables C9–C11).

4.3 Indirect associations

In the hypothesized model, GLU was indirectly associated with MRT through LSS [$\beta = 0.108$, 95% CI (0.069, 0.151)] and SE [$\beta = 0.477$, 95% CI (0.370, 0.569)]. The total indirect association was

TABLE 3 Structural associations.

Hypothesis	Path	*β*	*SE*	95% CI	*p*
H1	SP → GLU	0.519	0.083	[0.357, 0.681]	< 0.001
H2	CP → GLU	0.284	0.084	[0.120, 0.448]	< 0.001
H3	TP → GLU	0.124	0.083	[−0.039, 0.288]	0.137
H4a	GLU → LSS	0.405	0.039	[0.330, 0.481]	< 0.001
H4b	GLU → SE	0.796	0.029	[0.739, 0.853]	< 0.001
H5a	LSS → MRT	0.267	0.048	[0.173, 0.362]	< 0.001
H5b	SE → MRT	0.600	0.066	[0.471, 0.729]	< 0.001
Exploratory direct	GLU → MRT	0.050	0.053	[−0.053, 0.154]	0.341

Model fit: scaled $\chi^2(453) = 1680.06$, p < 0.001; robust CFI = 0.949, TLI = 0.944, RMSEA = 0.062, 90% CI [0.059, 0.066]; SRMR = 0.079. Latent R^2: GLU = 0.809; LSS = 0.164; SE = 0.633; MRT = 0.596. GLU → MRT is the exploratory direct path. H6 concerns statistical indirect associations in Table 4.

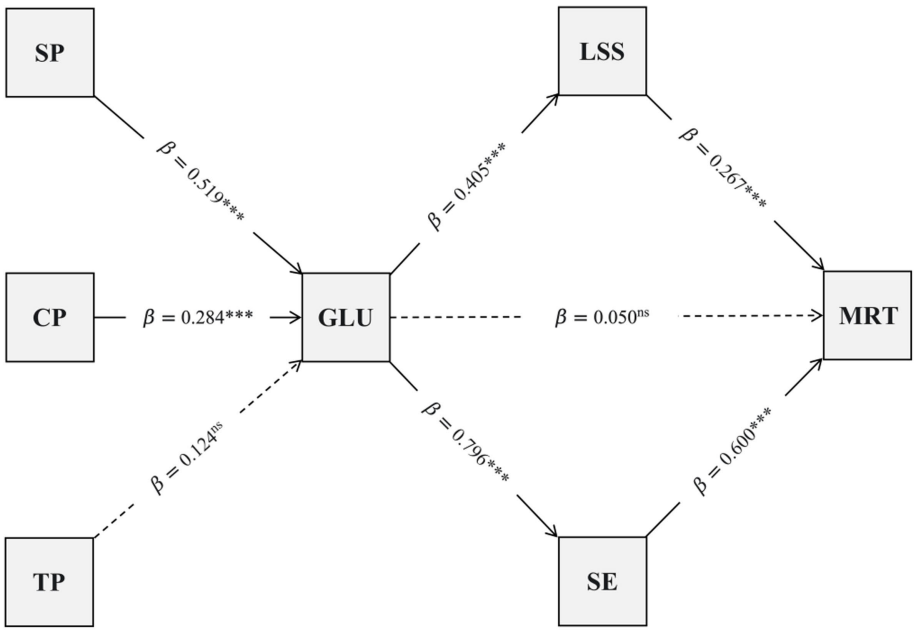

FIGURE 2
Path coefficients of structural model. *β*, standardized path coefficients; ***$p < 0.001$; = not significant. Dashed lines indicate NS = non-significant paths.

0.586, and the total association was 0.636 [95% CI (0.558, 0.708)]. These results were consistent with H6 under the hypothesized specification (Table 4), but the positive SE indirect association did not persist across the additional structural and method-factor models (Appendix Table C11). In the screening sensitivity analysis, all eight paths retained their direction and significance status (N = 1,580 vs. 1,160), although GLU–LSS (β = 0.636 vs. 0.405) and GLU–SE (β = 0.878 vs. 0.796) associations were stronger in the larger sample (Appendix Table C4). In the supplementary PSB model, both LSS (β = 0.146) and SE (β = 0.770) were positively associated with PSB (both $p < 0.001$), whereas the direct GLU–PSB path was not significant (β = 0.004, p = 0.913; Appendix Table C5).

4.4 Configurational analysis

Necessity analyses indicated that SP, CP, and TP each exceeded the 0.90 consistency threshold for high GLU (consistency = 0.914–0.927), while SE met this threshold for high MRT (consistency = 0.926; Appendix Table D2).

For high GLU, the intermediate solution comprised a single configuration combining all three CoI presences, with consistency of 0.943 and coverage of 0.853 (Table 5). Three equally minimal parsimonious solutions gave different core assignments to this same configuration, so no unique core condition could be identified (Appendix Table D4). Among cases assignable to truth-table corners, 558 belonged to the combination in which all three presences were high, and 164 to the combination in which all three were low. Each mixed-presence combination contained fewer than 24 cases and therefore fell below the frequency threshold (Appendix Table D3), limiting differentiation beyond the strong CoI correlations.

For high MRT, two configurations were identified: GLU and SE (consistency = 0.887, raw coverage = 0.829) and LSS and SE (consistency = 0.920, raw coverage = 0.776). SE appeared as a core condition in both configurations, with overall solution consistency of 0.876 and coverage of 0.895 (Table 6). A descriptive examination identified 35 cases combining low GLU membership with high LSS, SE, and MRT membership. The negated-outcome solutions were ~SP ~ CP ~ TP for non-high GLU and ~GLU ~ LSS ~ SE for non-high MRT (Appendix Table D4).

Across the 15 specifications, the high-GLU intermediate solution retained the same expression throughout. The complete high-MRT solution was retained in nine specifications, with alternative calibration and threshold settings producing modified or more restrictive expressions (Appendix Table D5).

5 Discussion

5.1 CoI presences and GenAI use for L2 learning

The findings suggest that learners' reported GenAI use for L2 learning is closely associated with shared perceptions of social connection, cognitive inquiry, and instructional support. The strong second-order CoI–GLU association is consistent with the interrelated learning

TABLE 4 Direct, indirect, and total associations with MRT.

Association	Standardized effect	Bootstrap SE	95% CI
GLU → MRT	0.050	0.053	[−0.045, 0.163]
GLU → LSS → MRT	0.108	0.021	[0.069, 0.151]
GLU → SE → MRT	0.477	0.051	[0.370, 0.569]
Total indirect association	0.586	0.045	[0.493, 0.667]
Total association	0.636	0.038	[0.558, 0.708]

The same complete model was estimated with ML for 5,000 ordinary case-bootstrap resamples; 5,000 were converged and admissible. Percentile 95% intervals. Standardized coefficients and products were recomputed in each resample. Total = direct + both specific indirect associations. Estimates describe model-based cross-sectional associations.

TABLE 5 Configurations associated with high GLU.

Configuration	SP	CP	TP	Consistency	Raw coverage	Unique coverage
SP*CP*TP	●	●	●	0.943	0.853	0.853

Intermediate solution; direct anchors = 3/5/6; n.cut = 24, consistency ≥ 0.80, PRI ≥ 0.70. ● = presence. One configuration is reported; core status is indeterminate because equally minimal parsimonious solutions assign different core conditions. Solution consistency = 0.943, coverage = 0.853, PRI = 0.923. Coverage refers to fuzzy-set membership. Exact 0.50 memberships are retained for fuzzy metrics and unassigned to crisp corners.

TABLE 6 Configurations associated with high MRT.

Configuration	GLU	LSS	SE	Consistency	Raw coverage	Unique coverage
GLU*SE	●		●	0.887	0.829	0.119
LSS*SE		●	●	0.920	0.776	0.066

Intermediate solution; direct anchors = 3/5/6; n.cut = 24, consistency ≥ 0.80, PRI ≥ 0.70. ● = core presence; blank = unspecified. Core status is determined from the paired parsimonious solution. Solution consistency = 0.876, coverage = 0.895, PRI = 0.829. Coverage refers to fuzzy-set membership. Exact 0.50 memberships are retained for fuzzy metrics and unassigned to crisp corners.

experiences described by the CoI framework (Garrison et al., 1999; Garrison and Arbaugh, 2007). Wang X. et al. (2024) similarly reported that stronger social and cognitive presences accompanied more frequent interaction with an AI English coach. GenAI can offer opportunities to express opinions, explore unfamiliar ideas, and seek clarification, sometimes through the same exchanges that provide cognitive prompts and instructional guidance. This convergence may help explain why the three presences were so closely aligned in learners' reported experiences. Such opportunities may make GenAI particularly attractive for language-learning activities requiring frequent practice, information exploration, and reflection (Ebadi and Amini, 2024; Wang X. et al., 2024).

GenAI chatbots allow learners to pursue questions, connect information, and reconsider their understanding through responsive exchanges. These opportunities align with the interactionist emphasis on meaningful exchanges and feedback in language learning (Gass and Mackey, 2014; Long, 1996). In such exchanges, the boundary between socially responsive dialogue and cognitive support may become less visible to learners, even though the two remain theoretically distinct functions.

TP, by contrast, showed no statistically significant independent association with GLU in the primary SEM analysis. This is an interesting departure from much CoI research, where instructional organization and facilitation are associated with learning experiences and satisfaction (Caskurlu et al., 2020). In GenAI-mediated environments, guidance can be embedded in chatbot responses, task suggestions, and feedback rather than delivered solely through a recognizable teacher role. Learners may consequently experience instructional support as closely intertwined with opportunities for inquiry and engagement. The strong correlations among TP, CP, and SE, together with the sensitivity of the TP–GLU estimate to model specification, suggest that instructional guidance may be experienced as part of a broader constellation of inquiry and engagement rather than as a clearly separable learning condition. Moreover, the alternative structural model revealed positive associations of TP with SE and MRT, suggesting that instructional guidance may relate to learners' engagement and metacognitive regulation even when its independent association with reported GenAI use is limited. These findings invite a more differentiated understanding of TP in environments where instructional functions are distributed between human teachers and GenAI systems, particularly when instructional guidance is woven into ongoing dialogue rather than experienced as a separate teaching activity.

5.2 Psychological associations with metacognitive regulation tendency

The SEM findings also highlight the relationships among GenAI use, LSS, SE, and MRT. Within the hypothesized model, GLU was positively associated with both LSS and SE, which were in turn associated with MRT. These indirect associations connect reported GenAI use with two distinct aspects of learners' psychological experiences: their beliefs about speaking capability and their active engagement in learning. These relationships bring together two areas of L2 research that have often been considered separately: the psychological resources associated with language use and learners' regulation of their own learning.

The association involving LSS is consistent with Bandura's (1997) account of self-efficacy as a belief relevant to effort, persistence, and responses to difficulty. GenAI chatbots provide opportunities to practice English, revisit unfamiliar expressions, and obtain feedback without the immediate social pressures of classroom communication (Fathi et al., 2024; Wang C. et al., 2024). Learners who report using GenAI for these purposes may also feel more capable of undertaking English-speaking tasks. Such confidence may be particularly relevant when learners encounter communicative difficulties, encouraging persistence and a willingness to reconsider their approaches. The positive relationship between LSS and MRT is consistent with previous research connecting self-efficacy with strategic learning and self-regulatory processes (Mills, 2014; Raoofi et al., 2012).

The association between SE and MRT is particularly interesting given the behavioral and cognitive aspects of engagement examined in this study. Attending carefully, expressing opinions, connecting information with prior experience, and generating examples all involve learners in actively processing learning content. Prior research similarly emphasizes the importance of cognitive engagement for reflective learning and the development of self-regulatory practices (Fredricks et al., 2016; Mercer and Dörnyei, 2020). In GenAI-mediated environments, readily available responses and feedback can provide opportunities for engagement, but learners still need to direct attention and effort toward making sense of what they encounter. However, the alternative structural model complicates a sequential interpretation of these relationships: allowing direct associations from CoI presences to SE and MRT reversed the GLU–SE association and its indirect association with MRT through SE. This finding suggests that learners' engagement may be closely intertwined with their perceptions of the broader learning environment, rather than simply following from GenAI use.

An important distinction concerns the nature of the outcome measured here. MRT reflects domain-general self-reported tendencies to monitor and regulate learning, rather than demonstrated thinking performance or observed regulation during GenAI-mediated activities. The present findings therefore connect GenAI-related learning experiences with learners' perceived orientation toward metacognitive regulation, adding a psychological perspective to studies examining GenAI and higher-order learning (Han et al., 2025; Hwang et al., 2018). They also suggest a useful direction for future research: examining how learners' engagement and self-efficacy relate to observable regulation practices during actual GenAI-supported tasks.

5.3 Configurational patterns associated with high GLU and MRT

The fsQCA findings complement the SEM analysis by showing how learning-related conditions occur together in cases characterized by high GLU and MRT. For high GLU, the intermediate solution identified a conjunction of SP, CP, and TP. This pattern is consistent with the CoI framework's emphasis on the interrelated nature of social connection, cognitive inquiry, and instructional facilitation (Garrison and Arbaugh, 2007). The concentration of cases in the combinations where all three presences were either high or low further suggests that learners tend to experience these conditions together, with little differentiation among alternative high-GLU configurations.

A different pattern emerged for high MRT, where two configurations were identified: GLU combined with SE, and LSS combined with SE. Both included high student engagement as a core condition. This convergence with the positive SE–MRT association in SEM places engagement at the center of the observed psychological relationships. The configurations also show that high MRT membership can occur alongside different combinations of GenAI use and speaking self-efficacy. In the LSS–SE configuration, GLU is unspecified, indicating that the combination of self-efficacy and engagement is associated with high MRT across varying levels of GLU membership. Given their

substantial overlap, these configurations suggest related patterns centered on engagement rather than clearly separate psychological pathways.

These results are consistent with research emphasizing learner engagement as an important aspect of productive technology-supported learning (Han et al., 2025; Mercer and Dörnyei, 2020). They also extend the discussion beyond individual statistical associations by examining how psychological conditions combine within learners' reported experiences. The structural and configurational findings collectively suggest that perceived learning conditions are closely aligned with reported GenAI use, while engagement consistently accompanies learners' metacognitive regulation tendencies. This perspective shifts attention from the contribution of isolated factors toward how learning conditions and psychological resources come together in GenAI-mediated L2 learning.

6 Implications

GenAI-mediated L2 learning should not be treated as a production line in which more use guarantees deeper learning. Our findings suggest more practical consequences. GenAI-mediated L2 learning need not be optimized for speed alone—and that can be its virtue. Pauses, hesitations, reformulations, and false starts can create space for clarification and reflection (Mercer, 2002; Gass and Mackey, 2014). In this context, such breathing room can be easier to sustain because chatbots are available on demand and may offer a less judgmental environment for language practice (Chuah and Kabilan, 2021; Dizon, 2026; Wang C. et al., 2024). The close association between CoI presences and GLU suggests that opportunities to express ideas, pursue questions, and receive instructional guidance deserve particular attention. Designing GenAI activities for L2 learning, then, should focus less on accelerating output and more on encouraging explanation, comparison, and reflection. For example, learners could be encouraged to question AI-generated responses, compare alternative explanations, and articulate why they revise their initial answers. The role of TP deserves similar recalibration, with feedback and task guidance designed to support learners' inquiry and judgment. The challenge is not to replace the teacher with the chatbot, but to redefine their complementarity. Teachers remain essential—not simply as distributors of instruction time, but as architects of inquiry who design and facilitate meaningful dialogue (Mercer et al., 2019). If GenAI supplies conversational space, teachers supply direction and intellectual stance.

The broader implication is for how GenAI-mediated learning is evaluated. The recurring role of engagement in the high-MRT configurations suggests that educators should attend not only to opportunities for GenAI use but also to how actively learners process, question, and evaluate information. Such engagement could be made more visible by asking learners to explain how they assess AI feedback, select among alternative responses, or adjust their learning strategies. These activities would also provide opportunities to examine whether learners' self-reported metacognitive regulation tendencies correspond to their actual learning practices. Rather than assuming a uniform relationship between GenAI use and learning, GenAI-mediated L2 learning can bring together opportunities for inquiry, speaking confidence, sustained engagement, and deliberate self-monitoring in ways responsive to learners' needs.

7 Conclusion

This study examined how CoI presences are associated with GLU and how such use relates to MRT through LSS and SE. The SEM findings revealed a strong shared association between CoI presences and reported GenAI use, although their distinct contributions were difficult to establish. Within the hypothesized model, GLU was indirectly associated with MRT through LSS and SE, but these associations were sensitive to alternative model specifications. Complementing these findings, fsQCA identified the joint presence of all three CoI conditions as a configuration associated with high GLU, while two configurations associated with high MRT shared student engagement as a core condition.

Several limitations should be acknowledged. First, the study relied on self-reported data from four independent cross-sectional survey waves. The findings therefore concern perceived learning experiences and statistical associations rather than observed interaction processes, changes over time, or causal effects. Individual records of GenAI use, including actual frequency, duration, platform, and interaction modality, were unavailable. Second, participant-level university, class, and survey-wave identifiers were not retained, preventing analyses of clustering and measurement invariance across institutions or survey rounds, as well as verification of possible repeated participation. The earlier attention-check exclusions could not be independently examined. Finally, construct overlap, common method influences, and sensitivity to alternative structural specifications limit the interpretation of distinct psychological contributions and indirect associations.

Future research could address these limitations by combining longitudinal or experimental designs with interaction logs, behavioral measures of metacognitive regulation, and more detailed institutional and participant-level records. Such evidence would help clarify how perceived learning conditions and engagement relate to learners' actual practices in GenAI-mediated L2 learning. More broadly, this study draws attention to a central educational challenge: expanding opportunities for human–AI interaction is only part of the task; understanding how learners actively engage with these opportunities and regulate their learning is equally important.

Data availability statement

The anonymized data supporting the conclusions of this article are available from the corresponding author upon reasonable request.

Ethics statement

The studies involving humans were approved by Academic Committee of the School of Languages and Communication Studies, Chongqing University of Technology. The studies were conducted in accordance with the local legislation and institutional requirements.

The participants provided their written informed consent to participate in this study.

Author contributions

YD: Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Writing – original draft, Writing – review & editing. BR: Methodology, Supervision, Writing – review & editing. JL: Formal analysis, Methodology, Writing – review & editing. CG: Formal analysis, Writing – review & editing. JX: Data curation, Writing – review & editing.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This research was supported by the Chongqing Municipal Education Commission (Grant No. 26SKGH202) and the Chongqing Municipal Social Science Planning Fund (Grant No. 2025NDYB139). Both funders provided financial support for the research.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that Generative AI was used in the creation of this manuscript. At the revision and proofreading stages, the authors used ChatGPT only for polishing the language. After using this tool, the authors rigorously reviewed and edited the content as needed and take full responsibility for the content of the publication.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2026.1999726/full#supplementary-material

References

Anderson, J. E., Nguyen, C. A., and Moreira, G. (2025). Generative AI-driven personalization of the Community of Inquiry model: enhancing individualized learning experiences in digital classrooms. *Int. J. Inf. Learn. Technol.* 42, 296–310. doi: 10.1108/IJILT-10-2024-0240

Anderson, T., Rourke, L., Garrison, D. R., and Archer, W. (2001). Assessing teaching presence in a computer conferencing context. *JALN* 5, 1–17. doi: 10.24059/olj.v5i2.1875

Arbaugh, J. B. (2007). An empirical verification of the community of inquiry framework. *JALN* 11, 73–85. doi: 10.24059/olj.v11i1.1738

Bandura, A. (1997). *Self-Efficacy: The Exercise of Control*. New York, NY: W. H. Freeman.

Cai, Q., Lin, Y., and Yu, Z. (2024). Factors influencing learner attitudes towards ChatGPT-assisted language learning in higher education. *Int. J. Hum. Comput. Interact.* 40, 7112–7126. doi: 10.1080/10447318.2023.2261725

Caskurlu, S., Maeda, Y., Richardson, J. C., and Lv, J. (2020). A meta-analysis addressing the relationship between teaching presence and students' satisfaction and learning. *Comput. Educ.* 157:103966. doi: 10.1016/j.compedu.2020.103966

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Struct. Equ. Model.* 14, 464–504. doi: 10.1080/10705510701301834

Chen, J., and Alibakhshi, G. (2026). AI-powered applications' effects on English language learners' cognitive, metacognitive, and resource management strategies, and language achievement. *J. Comput. Assist. Learn.* 42:e70171. doi: 10.1002/jcal.70171

Chen, Y., Ke, N., Huang, L., and Luo, R. (2025). The role of GenAI in EFL speaking: effects on oral proficiency, anxiety and risk-taking. *RELC J.* doi: 10.1177/00336882251341072

Chuah, K. M., and Kabilan, M. (2021). Teachers' views on the use of chatbots to support English language teaching in a mobile environment. *Int. J. Emerg. Technol. Learn.* 16, 223–237. doi: 10.3991/ijet.v16i20.24917

Dizon, G. (2026). ChatGPT as a tool for self-directed foreign language learning. *Innov. Lang. Learn. Teach.* 20, 334–350. doi: 10.1080/17501229.2024.2413406

Du, Y., and Gao, H. (2022). Determinants affecting teachers' adoption of AI-based applications in EFL context: an analysis of analytic hierarchy process. *Educ. Inf. Technol.* 27, 9357–9384. doi: 10.1007/s10639-022-11001-y

Du, Y., and Reynolds, B. L. (2025a). Exploring interaction process, L2 motivation and intended effort in chatbot-mediated interactions: evidence from latent Dirichlet allocation (LDA) topic modeling. *Comput. Assist. Lang. Learn.* 1–26. doi: 10.1080/09588221.2025.2552116

Du, Y., and Reynolds, B. L. (2025b). Rooted in and beyond interaction: a systematic review of interactive affordances of chatbots for language learning amidst the rise of large language models. *Acta Psychol.* 259:105307. doi: 10.1016/j.actpsy.2025.105307

Du, Y., Zeng, Y., Guo, C., Zhou, H., and Xiong, J. (2026a). Sure of themselves, not yet beholden: motivation, trust, and generative AI reliance in higher education. *Acta Psychol.* 270:107880. doi: 10.1016/j.actpsy.2026.107880

Du, Y., Zhao, H., Zhu, X., Zhang, H., and Liu, X. (2026b). When disciplines diverge: an explanatory sequential mixed-method study of humanities and STEM college students' acceptance of GenAI. *Int. J. Hum. Comput. Interact.* 42, 16574–16591. doi: 10.1080/10447318.2026.2622584

Ebadi, S., and Amini, A. (2024). Examining the roles of social presence and human-likeness on Iranian EFL learners' motivation using artificial intelligence technology: a case of CSIEC chatbot. *Interact. Learn. Environ.* 32, 655–673. doi: 10.1080/10494820.2022.2096638

Edumadze, J. K. E., and Govender, D. W. (2024). The community of inquiry as a tool for measuring student engagement in blended massive open online courses (MOOCs): a case study of university students in a developing country. *Smart Learn. Environ.* 11:19. doi: 10.1186/s40561-024-00306-9

Elkhodr, M., Gide, E., Wu, R., and Darwish, O. (2023). ICT students' perceptions towards ChatGPT: an experimental reflective lab analysis. *STEM Educ.* 3, 70–88. doi: 10.3934/steme.2023006

Fathi, J., Rahimi, M., and Derakhshan, A. (2024). Improving EFL learners' speaking skills and willingness to communicate via artificial intelligence-mediated interactions. *System* 121:103254. doi: 10.1016/j.system.2024.103254

Fiss, P. C. (2011). Building better causal theories: a fuzzy set approach to typologies in organization research. *Acad. Manag. J.* 54, 393–420. doi: 10.5465/amj.2011.60263120

Fredricks, J. A., Filsecker, M., and Lawson, M. A. (2016). Student engagement, context, and adjustment: addressing definitional, measurement, and methodological issues. *Learn. Instr.* 43, 1–4. doi: 10.1016/j.learninstruc.2016.02.002

Garrison, D. R., Anderson, T., and Archer, W. (1999). Critical inquiry in a text-based environment: computer conferencing in higher education. *Internet High. Educ.* 2, 87–105. doi: 10.1016/S1096-7516(00)00016-6

Garrison, D. R., Anderson, T., and Archer, W. (2001). Critical thinking, cognitive presence, and computer conferencing in distance education. *Am. J. Dist. Educ.* 15, 7–23. doi: 10.1080/08923640109527071

Garrison, D. R., and Arbaugh, J. B. (2007). Researching the community of inquiry framework: review, issues, and future directions. *Internet High. Educ.* 10, 157–172. doi: 10.1016/j.iheduc.2007.04.001

Gass, S. M., and Mackey, A. (2014). "Input, interaction, and output in second language acquisition," in (eds.) B. VanPatten & J. Williams *Theories in Second Language Acquisition*, (Routledge), 194–220. doi: 10.4324/9780203628942

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

Han, Y., Yang, S., Han, S., He, W., Bao, S., and Kong, J. (2025). Exploring the relationship among technology acceptance, learner engagement, and critical thinking in the Chinese college-level EFL context. *Educ. Inf. Technol.* 30, 14761–14784. doi: 10.1007/s10639-025-13375-1

Harris, J. (2025). Measuring listening and speaking self-efficacy in EFL contexts: the development of the communicative SE questionnaire. *Lang. Teach. Res.* 29, 1210–1234. doi: 10.1177/13621688221091608

Henseler, J., Ringle, C. M., and Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *J. Acad. Mark. Sci.* 43, 115–135. doi: 10.1007/s11747-014-0403-8

Hwang, G.-J., Lai, C.-L., Liang, J.-C., Chu, H.-C., and Tsai, C.-C. (2018). A long-term experiment to investigate the relationships between high school students' perceptions of mobile learning and peer interaction and higher-order thinking tendencies. *Educ. Technol. Res. Dev.* 66, 75–93. doi: 10.1007/s11423-017-9540-3

Jeon, J., Lee, S., and Choe, H. (2023). Beyond ChatGPT: a conceptual framework and systematic review of speech-recognition chatbots for language learning. *Comput. Educ.* 206:104898. doi: 10.1016/j.compedu.2023.104898

Kartal, G. (2024). The influence of ChatGPT on thinking skills and creativity of EFL student teachers: a narrative inquiry. *J. Educ. Teach.* 50, 627–642. doi: 10.1080/02607476.2024.2326502

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

Law, K. M. Y., Geng, S., and Li, T. (2019). Student enrollment, motivation and learning performance in a blended learning environment: the mediating effects of social, teaching, and cognitive presence. *Comput. Educ.* 136, 1–12. doi: 10.1016/j.compedu.2019.02.021

Li, L., and Kim, M. (2024). It is like a friend to me: critical usage of automated feedback systems by self-regulating English learners in higher education. *Australas. J. Educ. Technol.* 40, 1–18. doi: 10.14742/ajet.8821

Li, Y., Sadiq, G., Qambar, G., and Zheng, P. (2025). The impact of students' use of ChatGPT on their research skills: the mediating effects of autonomous motivation, engagement, and self-directed learning. *Educ. Inf. Technol.* 30, 4185–4216. doi: 10.1007/s10639-024-12981-9

Lin, Y. C., Li, Y., Chen, X., and Chiu, T. K. F. (2025). A systematic review on generative AI in L2 speaking development from a self-determination theory perspective. International journal of. *Appl. Linguis.* 36, 3229–3241. doi: 10.1111/ijal.70091

Liu, G. L., Darvin, R., and Ma, C. (2025). Exploring AI-mediated informal digital learning of English (AI-IDLE): a mixed-method investigation of Chinese EFL learners' AI adoption and experiences. *Comput. Assist. Lang. Learn.* 38, 1632–1660. doi: 10.1080/09588221.2024.2310288

Long, M. H. (1996). "The role of the linguistic environment in second language acquisition," in *Handbook of Second Language Acquisition*, eds. W. C. Ritchie and T. K. Bhatia (San Diego, CA: Academic Press), 413–468.

Meade, A. W., and Craig, S. B. (2012). Identifying careless responses in survey data. *Psychol. Methods* 17, 437–455. doi: 10.1037/a0028085

Mercer, N. (2002). *Words and Minds: How We Use Language to Think Together*. 1st Edn. London: Routledge.

Mercer, S., and Dörnyei, Z. (2020). *Engaging Language Learners in Contemporary Classrooms*. Cambridge: Cambridge University Press.

Mercer, N., Hennessy, S., and Warwick, P. (2019). Dialogue, thinking together and digital technology in the classroom: some educational implications of a continuing line of inquiry. *Int. J. Educ. Res.* 97, 187–199. doi: 10.1016/j.ijer.2017.08.007

Mills, N. (2014). "Self-efficacy in second language acquisition," in *Multiple Perspectives on the Self in SLA*, eds. S. Mercer and M. Williams (Bristol: Multilingual Matters), 6–22.

Phan, H. P. (2010). Students' academic performance and various cognitive processes of learning: an integrative framework and empirical analysis. *Educ. Psychol.* 30, 297–322. doi: 10.1080/01443410903573297

Podsakoff, P. M., Podsakoff, N. P., Williams, L. J., Huang, C., and Yang, J. (2024). Common method bias: it's bad, it's complex, it's widespread, and it's not easy to fix. *Annu. Rev. Organ. Psych. Organ. Behav.* 11, 17–61. doi: 10.1146/annurev-orgpsych-110721-040030

Ragin, C. C. (2008). "Measurement versus calibration: a set-theoretic approach," in *The Oxford Handbook of Political Methodology*, eds. J. M. Box-Steffensmeier, H. E. Brady and D. Collier (Oxford: Oxford University Press), 174–198.

Rahman, A., and Tomy, P. (2024). Intelligent personal assistant-an interlocutor to mollify foreign language speaking anxiety. *Interact. Learn. Environ.* 32, 4569–4586. doi: 10.1080/10494820.2023.2204324

Raoofi, S., Tan, B. H., and Chan, S. H. (2012). Self-efficacy in second/foreign language learning contexts. *Engl. Lang. Teach.* 5, 60–73. doi: 10.5539/elt.v5n11p60

Reeve, J., and Tseng, C. M. (2011). Agency as a fourth aspect of students' engagement during learning activities. *Contemp. Educ. Psychol.* 36, 257–267. doi: 10.1016/j.cedpsych.2011.05.002

Richardson, J. C., Maeda, Y., Lv, J., and Caskurlu, S. (2017). Social presence in relation to students' satisfaction and learning in the online environment: a meta-analysis. *Comput. Hum. Behav.* 71, 402–417. doi: 10.1016/j.chb.2017.02.001

Rosseel, Y. (2012). Lavaan: an R package for structural equation modeling. *J. Stat. Softw.* 48, 1–36. doi: 10.18637/jss.v048.i02

Satorra, A., and Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika* 66, 507–514. doi: 10.1007/BF02296192

Schneider, C. Q., and Wagemann, C. (2012). *Set-Theoretic Methods for the Social Sciences: a Guide to Qualitative Comparative Analysis*. Cambridge: Cambridge University Press.

Schraw, G., and Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemp. Educ. Psychol.* 19, 460–475. doi: 10.1006/ceps.1994.1033

Shang, J., Huang, Y., Xu, M., Huang, Y., Shen, X., Wang, G., et al. (2025). Competition between human learners and ChatGPT: enhancing university EFL students' reading comprehension and critical thinking through competitive questioning. *Comput. Assist. Lang. Learn.*, 1–35. doi: 10.1080/09588221.2025.2577356

Tai, T. Y., and Chen, H. H. J. (2024). Improving elementary EFL speaking skills with generative AI chatbots: exploring individual and paired interactions. *Comput. Educ.* 220:105112. doi: 10.1016/j.compedu.2024.105112

Timpe-Laughlin, V., and Dombi, J. (2020). Exploring L2 learners' request behavior in a multi-turn conversation with a fully automated agent. *Intercult. Pragmat.* 17, 221–257. doi: 10.1515/ip-2020-0010

Wan, Y., and Moorhouse, B. L. (2025). Using call Annie as a generative artificial intelligence speaking partner for language learners. *RELC J.* 56, 489–498. doi: 10.1177/00336882231224813

Wang, X., Pang, H., Wallace, M. P., Wang, Q., and Chen, W. (2024). Learners' perceived AI presences in AI-supported language learning: a study of AI as a humanized agent from community of inquiry. *Comput. Assist. Lang. Learn.* 37, 814–840. doi: 10.1080/09588221.2022.2056203

Wang, X., and Reynolds, B. L. (2024). Beyond the books: exploring factors shaping Chinese English learners' engagement with large language models for vocabulary learning. *Educ. Sci.* 14:496. doi: 10.3390/educsci14050496

Wang, T., Wu, T., Liu, H., Brown, C., and Chen, Y. (2025). Generative co-learners: enhancing cognitive and social presence of students in asynchronous learning with generative AI. *Proc. ACM Hum. Comput. Interact.* 9, 1–24. doi: 10.1145/3701198

Wang, C., Zou, B., Du, Y., and Wang, Z. (2024). The impact of different conversational generative AI chatbots on EFL learners: an analysis of willingness to communicate, foreign language speaking anxiety, and self-perceived communicative competence. *System* 127:103533. doi: 10.1016/j.system.2024.103533

Wu, W.-C., Hsieh, J. S. C., and Yang, J. C. (2017). Creating an online learning community in a flipped classroom to enhance EFL learners' oral proficiency. *J. Educ. Technol. Soc.*, 20, 142–157. Available online at: https://www.jstor.org/stable/90002170 (Accessed March 7, 2025).

Wu, J., Li, Y., Zhou, J., and Chen, S. (2025). The impact of intelligent personal assistants on mandarin second language learners: interaction process, acquisition of listening and speaking ability. *Comput. Assist. Lang. Learn.* 38, 1769–1794. doi: 10.1080/09588221.2024.2317849

Yang, G., Wang, Y., Zhang, Y., Yang, M., Zeng, Q., and Song, Z. (2025). An empirical study of AI-supported interleaved training strategy to improve EFL students' English impromptu speaking performance, learning engagement, technology acceptance and epistemic network structure. *Asia Pac. Educ. Res.* 34, 1519–1540. doi: 10.1007/s40299-024-00962-y

Zhang, R., Zou, D., and Cheng, G. (2023). Chatbot-based training on logical fallacy in EFL argumentative writing. *Innov. Lang. Learn. Teach.* 17, 932–945. doi: 10.1080/17501229.2023.2197417

Zhang, R., Zou, D., and Cheng, G. (2025). ChatGPT affordance for logic learning strategies and its usefulness for developing knowledge and quality of logic in English argumentative writing. *System* 128:103561. doi: 10.1016/j.system.2024.103561

Zheng, Y. B., Zhou, Y. X., Chen, X. D., and Ye, X. D. (2025). The influence of large language models as collaborative dialogue partners on EFL English oral proficiency and foreign language anxiety. *Comput. Assist. Lang. Learn.* 39, 1092–1118. doi: 10.1080/09588221.2025.2453191