Computers & Education 194 (2023) 104703

Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

What matters in AI-supported learning: A study of human-AI interactions in language learning using cluster analysis and epistemic network analysis

Xinghua Wang[a,*,1], Qian Liu[a,1], Hui Pang[a,**], Seng Chee Tan[b], Jun Lei[c], Matthew P. Wallace[d], Linlin Li[a]

[a] Normal College, Qingdao University, Qingda 1st Road 16, China
[b] National Institute of Education, Nanyang Technological University, 1 Nanyang Walk, 637616, Singapore
[c] Faculty of Foreign Languages, Ningbo University, 818 Fenghua Road, China
[d] Faculty of Arts and Humanities, University of Macau, A, Avenida da Universidade, Taipa, Macau, China

A R T I C L E I N F O

Keywords:
Elementary education
Applications in subject areas
Learning communities
Human-computer interface

A B S T R A C T

This study investigates how students interact with artificial intelligence (AI) for English as a Foreign Language (EFL) learning and what matters in AI-supported EFL learning. It was conducted in naturalistic learning settings, involving sixteen primary school students and lasting approximately three months. The students' usage data of an AI agent and their reflection essays about the interactions with the AI agent were analyzed using cluster analysis and epistemic network analysis based on the frameworks of community of inquiry and students' approaches to learning. The results suggest four clusters of students, each with its distinct way of interacting with AI for language learning. More importantly, the comparisons of the four clusters of students reveal that even in AI-supported learning, not everyone can benefit from the potential promised by AI. The deep approach to AI-supported learning may amplify the benefits of AI's personalized guidance and strengthen the sense of the human-AI learning community. Passively or mechanically following AI's instruction, albeit with high levels of participation, may decrease the sense of the human-AI learning community and eventually lead to low performance. This study contributes to and has implications for the educational implementation of AI, as well as the facilitation and graphical representation of learner-AI interactions in educational settings.

1. Introduction

Studies exploring the use of affordable AI agents such as chatbots and conversational agents for second language (L2) learning have been increasing in recent years (Jeon, 2022; Lin & Mubarok, 2021; Underwood, 2017). AI agents provide students with ample opportunities to practice target languages in relatively stress-free learning environments, and particularly, offer personalized instruction for individual students in large classes (Dizon, 2020; Moussalli & Cardoso, 2020). Students have mostly reported that they enjoy and

* Corresponding author.
** Corresponding author.
E-mail addresses: wangxinghua379@gmail.com (X. Wang), lucy.5267@163.com (Q. Liu), hopehuipang@gmail.com (H. Pang), sengchee.tan@nie.edu.sg (S.C. Tan), rayjun.lei@outlook.com (J. Lei), mpwallace@um.edu.mo (M.P. Wallace), linlin908@qdu.edu.cn (L. Li).
[1] Xinghua Wang and Qian Liu contribute equally to this study and share first authorship.

https://doi.org/10.1016/j.compedu.2022.104703
Received 22 July 2022; Received in revised form 4 December 2022; Accepted 6 December 2022
Available online 7 December 2022
0360-1315/

feel more relaxed about speaking to AI in L2 than to real humans (Wang et al., 2022; Tai & Chen, 2020). These benefits associated with AI agents offer promising solutions to intricate problems encountered in conventional L2 classrooms, such as inadequate speaking practice opportunities and limited class time (e.g., de Vries et al., 2015), unwillingness to speak (e.g., Tai & Chen, 2020), and lack of personalized feedback to every student (e.g., Luo, 2016).

Despite the benefits associated with AI-supported language learning, there is a noteworthy dearth of knowledge regarding how students interact with AI agents for language learning and what differences may exist among distinct types of students in human-AI interactions. Without such knowledge, the mechanism of AI-supported language learning remains a black box, that is, while students' initial language competence at one end and their final competence at the other have been examined, what happens in between is largely unknown. Consequently, few insights can be gained to improve the design and implementation of AI agents in language learning as well as the pedagogy of AI-supported language learning.

To address this gap, this study investigated the human-AI interactions in an English as a Foreign Language (EFL) class involving 16 sixth-grade students in a primary school. An AI agent called the "AI coach" was used in this study. The AI coach was a virtual AI agent installed on mobile devices and was embodied as a female teacher. It was specifically created for EFL learning, supporting unlimited English practice and providing personalized feedback on students' speaking.

Considering that prior studies in AI-supported language learning have anthropomorphized AI applications as peers, teachers, tutors, and teaching assistants for learners (Engwall & Lopes, 2020; Randall, 2019), students and AI may form a learning community where the social, cognitive, and teaching presences students perceive in interacting with AI can forge meaningful learning experiences and facilitate L2 acquisition (Jeon, 2022; Yu & Li, 2022). As such, theoretically, community of inquiry (CoI; Garrison et al., 2010; Garrison & Arbaugh, 2007) serves as a feasible framework for analyzing human-AI interactions to reveal how learning happens in AI-supported language learning (Wang et al., 2022; Yu & Li, 2022). Nonetheless, young learners may not necessarily be competent to use digital technologies, including AI, effectively for learning, though they are often perceived to be tech-savvy (Qi, 2019; Thompson, 2013). Substantial research (e.g., Niu et al., 2022; Ellis & Bliuc, 2019; Lindblom-Ylänne et al., 2019) has shown that the ways in which students approach online or offline learning play a crucial role in determining their learning experiences and performance. Students' approaches to learning (SAL) framework has been found to effectively explain why some students are more successful than others in school (Ellis & Bliuc, 2019). Though SAL have been widely applied in a variety of learning contexts (e.g., Biggs & Tang, 2011; Ellis & Bliuc, 2019; Hailikari & Parpala, 2014), it is still unknown how students' approaches to AI-supported learning function in a human-AI community. Therefore, building on the combined frameworks of CoI and SAL, this study employed cluster analysis and epistemic network analysis (ENA) to provide insights into how distinct types of students interacted with the AI coach for EFL learning. Overall, this study sought to address the following two research questions:

(RQ1) How many distinct types of learners can be identified with respect to the AI usage?
(RQ2) How do distinct types of learners interact with AI for language learning and what are their differences in human-AI interactions, if any?

2. Literature review

2.1. AI agents for language learning

AI agents in this study refer to entities that are capable of performing tasks related to intelligent beings through reasoning, learning, and expressing themselves, to a certain extent (Wang et al., 2022; Berendt et al., 2020). They are created in different forms (Randall, 2019), for instance, anthropomorphic (with human-like appearances), cartoon-like (with exaggerated cartoon features), and mechanomorphic (with machine-like features). Based on the visibility of AI agents, they can also be clustered into two types: embodied and disembodied (Lee & Jeon, 2022). Embodied AI agents include physical robots and virtual agents (e.g., Randall, 2019; van den Berghe et al., 2019), while disembodied agents are often invisible and delivered through smart devices such as Google Assistants and Amazon's Alexa (e.g., Dizon, 2020; Moussalli & Cardoso, 2020).

AI agents (e.g., robots, virtual assistants) are increasingly applied in educational fields to support the learning of various disciplines, in particular, L2 acquisition, leveraging the advancement of technologies such as speech recognition and natural language processing techniques (Engwall & Lopes, 2020; Natale & Cooke, 2020; Tai & Chen, 2020). As the pedagogy about the use of AI agents in language learning is still nascent (Lee & Jeon, 2022), some studies started by examining learners' perceptions of AI so as to inform pedagogical development thereof. For instance, Chen et al. (2020) investigated how L2 learners with different English proficiency levels perceived the potential of Google Assistant for language learning. The survey and interview results based on students' short usage of Google Assistant showed that the learners found interactions with it enjoyable and perceived it to be useful in improving speaking and listening competence, thus indicating Google Assistant's educational affordances. To gain a better understanding of what benefits AI agents may have for education, Jeon (2022) investigated the affordances of AI chatbots for EFL learning based on primary school students' interviews and the log data of their interactions with the chatbot. The qualitative analysis of the data revealed the chatbot's pedagogical affordances (interactional opportunities for students to acquire EFL), technological affordances (technological functionalities supporting or hindering EFL learning, e.g., speech recognition techniques, online dictionaries, and search engines), and social affordances (AI′ ability to create a stress-free or anxiety-loaded experience).

As for the benefits associated with various forms of AI agents, much of the previous research (e.g., Lee & Jeon, 2022; van den Berghe et al., 2019) has discovered that they are useful and valuable for improving students' linguistic as well as non-linguistic performance. The linguistic performance involves areas such as vocabulary acquisition as well as L2 listening, speaking, and reading competence (Randall, 2019). With regard to non-linguistic performance, AI agents have been reported to help increase learners' willingness to communicate, L2 enjoyment, and L2 confidence as well as decrease anxiety in learning (e.g., Alemi et al., 2015; Tai & Chen, 2020; Wang et al., 2022). For instance, Awada (2022) examined the integration of robotics and weblog into the EFL curriculum of middle and secondary schools and showed that such integration improved students' EFL academic scores. Underwood (2017) investigated the effectiveness of including voice-driven AI assistants such as Amazon's Alexa and Apple's Siri in an EFL class for elementary school students. Students asked questions and gave commands to AI in the L2 and constantly reformulated their questions when the AI could not recognize their utterances. The results showed that the students were highly engaged and enjoyed when interacting with the AI assistants. Compared with speaking to humans, they spoke more English when using the AI assistants in group activities and demonstrated more persistence in speaking to the AI assistants in English to let them follow their orders.

On the whole, AI agents have reportedly been conducive to increasing learners' overall L2 competence. However, much of the prior research involved a short intervention, for instance, using language robots for only ten to 15 min over two days (Engwall & Lopes, 2020) or using only an hour of Google Assistant (Chen et al., 2020), rendering the research findings questionable due to possible novelty effects (van den Berghe et al., 2019). Many studies (e.g., Chen et al., 2020; Lee & Jeon, 2022) solely relied on attitudinal data from surveys or interviews as their data sources, thereby subjecting the research findings to potential respondents' biases. Particularly, most studies (e.g., Dizon, 2020; Tai & Chen, 2020) aimed at examining the effectiveness of AI agents in improving students' L2 competence. Little is known about how learning actually happens through human-AI interactions. In addition, there is a keen lack of knowledge regarding whether there are differences in human-AI interactions among distinct types of students. Without such knowledge, the research on AI-supported learning would remain superficial, making it difficult to inform pedagogical development in the use of AI and the design of AI for education.

2.2. Community of inquiry for AI-supported language learning

The CoI framework refers to a process model for digital instructional environments with three essential constructs: social, teaching, and cognitive presences (Garrison et al., 2010). It emphasizes meaningful learning through community interactions and has been extensively used to guide the development of quality online education (Garrison & Arbaugh, 2007; Lomicka, 2020; Smidt et al., 2021). Nonetheless, prior studies using the CoI framework were primarily carried out to examine human interactions through conventional technologies (e.g., online discussion boards and courses) which are unable to perform human-like processing and logic (Wang et al., 2022). In view of the growing use of AI as a humanized agent in education (Randall, 2019), new perspectives on AI-supported learning can be gained by applying the CoI theory to human-AI interactions.

The CoI theory has been successfully applied in L2 learning settings and holds implications for facilitating technology-supported language learning. For example, Smidt et al. (2021) used the CoI framework to investigate how Malaysian students collaboratively constructed English knowledge concerning their experience with social, cognitive, and teaching presences in asynchronous online learning. Their research findings indicated that the participants had high perceptions of teaching presence, followed by cognitive and social presence, which facilitated the collaborative construction of knowledge. Wu et al. (2017) explored the influence of an online learning community in a flipped classroom on EFL learners' oral proficiency and perceptions. The research results suggested that the online learning community facilitated productive collaboration, increased learning engagement, and improved learners' oral proficiency. Yu and Li (2022), in a bibliometric analysis of the use of CoI over the past decade, opined that social, cognitive, and teaching presences reside in students' interactions with diverse digital technologies, including artificial intelligence and robots. There are, however, surprisingly few empirical studies using CoI for analyzing human-AI interactions.

In the context of the current study, social presence is concerned with students' identification and perceived connection with the AI coach as well as other learners in AI-supported learning. According to Swan et al. (2009), social presence consists of three components: affective expression (sharing emotions and feelings), open communication (forging mutual recognition and awareness), and group cohesion (maintaining group commitment). Social presence has been associated with students' learning performance and their satisfaction with digital learning activities (Watson et al., 2016).

Teaching presence is related to the perceived instructional design embedded in the AI coach that facilitates students' language acquisition and enhances their learning experience. It has been shown to significantly predict students' satisfaction, sense of community, and learning performance (Garrison et al., 2010; Turula, 2017). In this study, teaching presence involves the design and organization of L2 learning and the direct instruction provided by the AI coach.

Cognitive presence refers to the process of students making sense of and acquiring the L2 through sustained reflection and problem-solving during interactions with the AI coach. With the assistance of the AI coach, students identify problems in EFL speaking, work with the AI coach to fix those problems, improve their language competence, and apply what they have learned to real-life situations. Cognitive presence is strongly linked to both perceived and actual learning (Akyol et al., 2011), but is also reportedly difficult to develop (Garrison & Arbaugh, 2007).

The CoI framework and the SAL framework, which will be explained in the next section, will be used as the theoretical framework for this study.

2.3. Students' approaches to learning in AI-supported settings

The theory of SAL has been an established body of research attracting an increasing amount of attention (Biggs, 1993; Ellis et al., 2008; Gibbs & Coffey, 2004; Umapathy et al., 2020). SAL are concerned with distinct ways and processes of students' engagement in learning activities (Ellis et al., 2008; Thompson, 2013). They are significantly related to students' academic performance and provide an effective framework for understanding the quality of student learning (Biggs & Tang, 2011; Ellis & Bliuc, 2019). Overall, three types of SAL have been identified: deep, surface, and organized approaches (Biggs & Tang, 2011; Herrmann et al., 2017; Rozgonjuk et al., 2018). The three approaches of SAL have been validated and refined in substantial numbers of studies involving different school levels from primary schools (e.g., Cheng, 2017) to institutes of higher education (e.g., Tsai et al., 2017), and across diverse disciplines such as language learning (e.g., Yu, 2019), mathematics (e.g., Cai et al., 2019), and engineering (e.g., Ellis et al., 2008).

A deep approach concerns active learning and sense-making of learning resources. Deep learners focus on systematic and thorough comprehension of content knowledge by drawing on higher-order thinking strategies (e.g., critical thinking and problem-solving). In a surface approach to learning, students tend to fear failure and focus on finishing required tasks instead of constructing and acquiring new knowledge. They spend the least amount of effort to meet the minimum learning requirements. An organized approach to learning is characterized by a focus on attaining high performance through studying in a systematic and well-organized way. The organized approach to learning differs from the deep or surface approaches in that the former is related to effort management while the latter two are about cognitive processing (Ellis & Bliuc, 2019; Herrmann et al., 2017). Students can adopt any one of three approaches they deem appropriate and effective based on their motivation and understanding of external environments. They can also shift from one approach to another when responding to varying course requirements (Ellis & Bliuc, 2019; Garrison & Cleveland-Innes, 2005).

Students' approaches to EFL learning largely reflect their conceptions and knowledge regarding EFL (Mak & Chik, 2011; Yu, 2019). Though EFL learning has been viewed by many learners as rote learning (a noted surface approach), emphasizing substantially repetitive memorization (Li & Cutting, 2011), prior studies have also identified the importance of the deep approach to EFL learning and the negative influence associated with the surface approach. For instance, Yu (2019) investigated how SAL predicted EFL learners' learning outcomes. From the survey data, they found that a deep approach to learning significantly predicted the learning outcomes of the students from mainland China. Niu et al., 2022 found in technology-enhanced learning contexts that the surface approach to EFL learning featuring mechanical repetition and memorization was equivalent to a maladaptive approach to language learning, eventually resulting in a negative learning experience (e.g., stress and burnout) and decreased learning performance. Heift (2002) examined the impact of learner control, which is related to learners' agentic power for cognitive processing and essentially reflects different approaches to learning, on students' usage patterns of an intelligent language tutoring system for German grammar practice. Four types of learners with distinct usage patterns were identified: browsers who only surfed exercises without providing any answers; frequent peekers who frequently stole a look at correct answers provided by the system; sporadic peekers who corrected their errors more often than they peeked at the correct answers; and adamants who seldom corrected their errors and peeked at the correct answers. Those with low language skills tended to take a surface approach by frequently peeking at the system-provided answers.

Given the increased use of AI in education, it is essential that we gain a better understanding of the role of students' approaches to AI-supported learning in the human-AI community. The application of the SAL framework to the human-AI community can unravel the intents and strategies underlying human-AI interactions, thereby holding promises for facilitating the use of AI and the development of pedagogy for AI-supported learning.

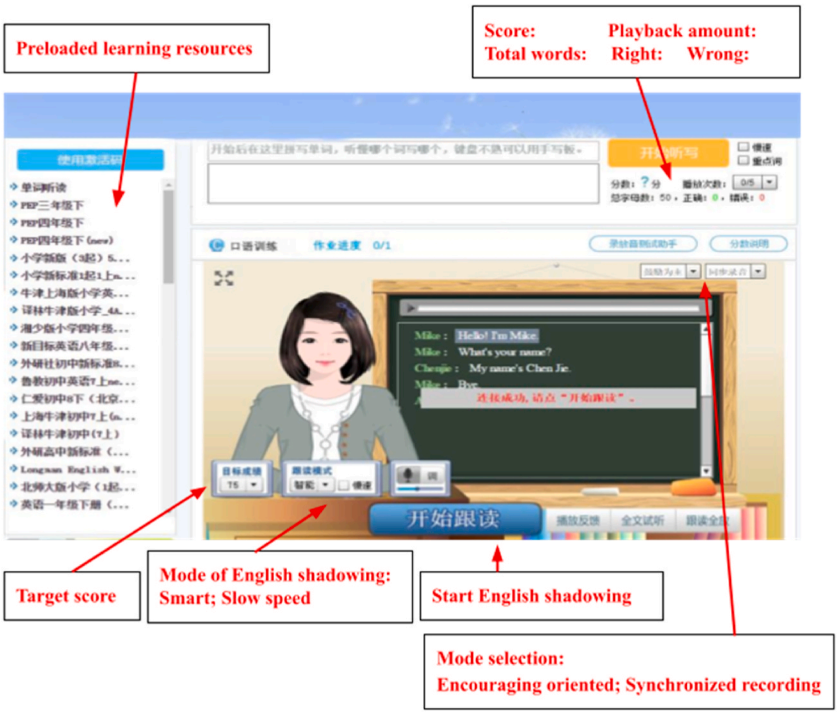

Fig. 1. A screenshot of the AI coach (Wang et al., 2022). *Note.* Copyright belongs to kouyu100.com.

According to the literature related to the associations among CoI, SAL, and learning motivation (Akyol et al., 2011; Ellis & Bliuc, 2019; Garrison & Cleveland-Innes, 2005; Herrmann et al., 2017), students who join technology-enhanced learning activities with intrinsic motivation, adopt deep and organized approaches, and perceive social, cognitive, and teaching presences tend to demonstrate high engagement and to excel in learning performance. Overall, considering that the interplay among social, teaching, and cognitive presences could lead to deep learning in digital settings (Akyol et al., 2011), the present study applied the CoI and SAL frameworks to the human-AI learning community with the goal of discovering how the interplay between the three presences and SAL shaped students' AI-supported learning. The CoI and SAL frameworks guided the coding of the students' reflection essays on their interactions with the AI coach, which served as one of the main data sources for this study and were subsequently processed by epistemic network analysis.

3. Methodology

3.1. Participants and research context

The present study was conducted in a 6th grade EFL class in an urban primary school in China, lasting approximately three months. The research context in this study was similar to Wang et al., 2022. The students started to learn English in Year 1 of primary school. Nonetheless, their English proficiency was normally at very basic levels due to limited exposure to authentic L2 environments and exam-oriented education. The AI coach (kouyu100.com) was purchased by the school to cope with the lack of native English-speaking teachers and to assist the students in learning English anywhere and anytime. In the current study, the AI coach was used as a part of the EFL course requirements and to supplement the face-to-face class in terms of speaking and listening practice and vocabulary learning. The students could seek support from their teacher who had over ten years' experience teaching EFL and from their parents who mostly had college degrees. This study collected data from sixteen students who used the AI coach (refer to Fig. 1 as an illustration) regularly, who generated intact usage data in the AI system and completed the reflection essay at the end of this study. Informed consent was obtained from the students as well as their parents.

The AI coach employed speech recognition and natural language processing algorithms to parse students' language output and diagnose their output by comparing it with the AI's linguistic models, and then used its recommendation system and speech synthesis techniques to generate personalized feedback for the students (Engwall & Lopes, 2020; Heift & Schulze, 2007; Settles et al., 2020). The students used the AI coach in and out of the naturalistic classroom for an average of 15 min every school day as a daily assignment to address pronunciation problems, develop vocabulary, and improve listening and conversational skills. To ensure students' healthy and proper use of the AI coach (Felix, 2020), the teacher checked the students' usage data through the administrator's access portal but only reminded those with excessive usage, such as students who used it at home for an hour at a time, far exceeding the stipulated 15 min. In the learner-AI interactions, the AI coach would automatically analyze the students' utterances based on grammar and pronunciation, present the students' utterances on screen and highlight any errors made, and then provide feedback by offering encouraging comments (e.g., "Your English is very authentic") or indirect error correction (e.g., "Please note the rhythm of your utterances") accordingly.

Take pronunciation learning, which used the English shadowing (ES) method (Hamada, 2016), for example. The learner-AI interactions started with the AI coach speaking a sentence for the students to follow. As the students spoke, the AI coach presented what they said on screen and immediately highlighted the incorrect or mispronounced words or phrases for the students to attend to. Meanwhile, it recorded every utterance made by the students for them to compare with the AI coach as well as other students' utterances, thereby giving them the opportunity to reflect on their performance, recognize their errors, and subsequently initiate actions for further improvement. In addition, the AI coach would give a score for every utterance made by the students on a scale of 100 points based on its algorithm settings. The students could also practice as many times as possible until they obtained desirable scores.

3.2. Data sources

The data for this study came from two sources: (a) the students' usage data of the AI system and (b) their reflection essays. The students' usage data comprised four components: average frequencies of English shadowing, listening practice, and vocabulary learning, as well as average scores of English shadowing. The average frequencies refer to the number of times the students completed a task on average. The scores of the listening practice and vocabulary learning practice were not included as they were all full scores (100 points), which the students had to obtain through a trial-and-error process once they clicked into the two types of practice. In the listening practice, the students had to answer questions according to their understanding of the materials read by the AI coach. In the vocabulary learning practice, the students needed to do Chinese-English translations based on the pictures or words they saw. The listening and vocabulary learning practices had one correct response per question, whereas the English shadowing practice was flexible and had no standard answers. The AI coach scored each of the students' utterances out of 100 points according to its natural language processing algorithms. Based on the students' language proficiency, they could obtain varied scores in the English shadowing practice.

The average frequencies of the English shadowing practice, listening practice, and vocabulary learning practice together were used to represent students' engagement in AI-supported language learning, while the scores of English shadowing practice were used to indicate students' learning performance. Though the English shadowing scores may not reflect students' overall language competence, they were given by the AI coach without possible human bias, thus representing a reasonable proxy of learning performance in AI-supported language learning settings.

At the end of the semester, the students were required to write an essay reflecting on their interactions with the AI coach and how such interactions worked or did not work for their EFL learning. There was no word limit. The writing was in Chinese as the students' English competence was not yet adequate for clear expression.

3.3. Coding of reflection essays

Based on the literature on CoI and SAL (Ellis & Bliuc, 2019; Garrison & Arbaugh, 2007; Richardson et al., 2012) while taking into account the research data, the research team collectively created an initial coding framework, including three components with seven sub-components for CoI: social presence (emotional expression and affinity), teaching presence (instructional design, guidance & facilitation, and feedback), and cognitive presence (agentic exploration and problem-solving). There were three components for SAL: deep, surface, and organized approaches to AI-supported language learning. However, other variables (L2 learning enjoyment and motivation) were progressively included in the coding framework while the research team applied the CoI and SAL theories to the data, yielding a final coding framework as shown in Table 1.

Two researchers independently coded all reflection essays with a natural sentence (punctuated with a period) as a unit of analysis. Discussions were conducted daily to address disagreements arising in the coding process. Eventually, 87 percent intercoder reliability was obtained.

3.4. Data analysis

Cluster analysis was used alongside epistemic network analysis in this study. A two-step cluster analysis was performed on students' system usage data generated from their interactions with the AI coach, with the aim of identifying distinct groups of students based on differences in engagement and performance, respectively. The two-step cluster analysis is capable of automatically identifying the optimal cluster solutions by comparing a group of clustering criteria across different cluster solutions instead of using arbitrary decisions (Benassi et al., 2020). It is seen as one of the most reliable clustering techniques (Kent et al., 2014).

Subsequently, ENA (epistemicnetwork.org/) was conducted on the coding of the reflection essays of distinct clusters of students to reveal how different types of students interacted with the AI coach. ENA has been used to examine students' social-cognitive

Table 1
Finalized coding framework.

Components	Sub-components	Explanations and examples
Social presence	Emotional expression	Students clearly express their positive feelings toward the AI coach. e.g., "It (the AI coach) is an excellent learning platform." "… I deeply appreciate the help of the AI coach."
	Affinity	Students perceive the affinity with the AI coach. e.g., "It (the AI coach) is like an authentic British person." "It (the AI coach) always encourages us …"
Teaching presence	Instructional design	Students perceive that the AI learning platform has a systematic instructional design supporting their English learning. e.g., "It (the AI coach) has multiple functions, supporting pronunciation, vocabulary learning, …, giving you authentic learning experience, …., encouraging you in different ways …"
	Guidance & Facilitation	Students perceive the guidance and facilitation provided by the AI coach which increase their learning engagement. e.g., " …. what surprises me is that it (the AI coach) automatically categorizes my dubbing of movie clips for me to retrieve and review them easily." "It (the AI coach) facilitates us to challenge other learners …."
	Feedback	Students perceive that the AI coach provides timely and clear feedback for them so that they can know their achievements and problems. e.g., "It (the AI coach) will give you very specific feedback on your English speaking right after you read a sentence." "I will get a score for each practice so that I will know the difference between me and other learners who are doing the same practice."
Cognitive presence	Agentic exploration	During the student-AI interaction, students constantly reflect on their progress and take agentic actions to engage with the AI coach. e.g., "Ever since I have the AI coach, I rarely need my parents or teacher to remind me to learn English." "I often reflect on my progress aided by the AI coach."
	Problem-solving	During the infinite students-AI interaction, students can solve problems arising in the learning process by themselves. E.g., "I am making progress every day by tackling problems related to grammar and intonation."
Approaches to learning	Deep approach	Students use higher-order thinking skills in English learning. e.g., "I practice English with a clear goal." "… I consciously compare my pronunciation with the AI coach who also uses different colors to highlight the differences in pronunciation between my English and the more authentic one."
	Surface approach	Students emphasize memorization and reproduction in English learning. e.g., "… I often mimic movie clips or memorize words and sentences with the help of the AI coach."
	Organized approach	Students have an organized schedule in the use of the AI coach. e.g., "… I use it (the AI coach) every day and make progress every day."
L2 learning enjoyment		Learning English with the AI coach is a pleasant experience for the students. e.g., "I am enjoying the learning process with the AI coach." "I cannot help but express my joy when I finally overcome the problems of pronunciation and intonation"
Intrinsic motivation		Students like the mode of learning English with the AI coach and have a strong interest in this mode. e.g., "I like to learn English with it (the AI coach) as it helps improve my English competence"
Extrinsic motivation		Students are extrinsically motivated to learn English. e.g., "You can get (virtual) flowers and awards if you practice English with the AI coach every day and achieve good performance."

engagement in digital learning settings (e.g., Ouyang et al., 2021; Tan et al., 2022) as it is useful in revealing the complexity of students' learning engagement. It combines quantitative and qualitative analyses and closes the interpretative loop by enabling researchers to triangulate the network results using original data (Shaffer & Ruis, 2017).

ENA visualizes network models of students' interactions with the AI coach based on the co-occurrence of the elements developed from the CoI and SAL frameworks and the original data. "Elements" in the ENA literature refer to codes in coding frameworks, which are visualized as nodes in network graphs (Shaffer & Ruis, 2017). The co-occurrences of the elements represent the interconnection among different learning behaviors that make up the socio-cognitive structures in AI-supported learning. For instance, if the elements of "feedback" and "deep approach" co-occurred frequently, there would be a strong interrelationship between them. In ENA, each utterance can be assigned more than one code when it contains more than one layer of meaning. The codes are converted to binary adjacency matrices, which are then added up to cumulative adjacency matrices that show the weighted patterns of co-occurrences among different codes. Subsequently, each student's cumulative adjacency matrix is visualized in a high-dimensional space (Shaffer et al., 2016; Shaffer & Ruis, 2017). To facilitate interpretation, ENA typically generates networks using two dimensions through the techniques of normalization and dimension reduction, among which researchers select the two-dimensional space that can best explain the variance in the research data (Shaffer et al., 2016). In the current study, this is 30.60%. The networks are weighted with thicker lines indicating stronger connections and thinner lines representing weaker connections between two elements.

In addition, ENA allows for direct comparisons between individuals or groups, as the elements appear in the same places across the networks of each individual or group, thereby creating invariance in the placement of elements (Shaffer & Ruis, 2017). Each individual's or group's network is represented by a centroid in a two-dimensional space. The centroid of a network is equivalent to the center of an object's mass, summarizing the network of nodes as a single point in the two-dimensional space. The locations of the centroids are determined by minimizing the distance between the centroids and the nodes forming the network. Thus, each centroid can be interpreted by its location in the four quadrants of the two-dimensional space, in which each quadrant is characterized by the nodes in it and again characterizes the centroids placed in it.

4. Results

In this section, the results related to the cluster analysis were first reported, followed by the ENA results of different clusters of students' retrospective interactions with the AI coach.

4.1. Results of cluster analysis

The two-step cluster analysis was conducted to identify clusters of students in terms of their interactions with the AI coach as well as their learning performance. Multiple rounds of analyses were performed with the log-likelihood distance measure to determine the probabilities of cluster memberships, yielding two-to five-cluster solutions (see Appendix C for details). Silhouette analysis was used to examine the overall goodness-of-fit of these cluster solutions, with silhouette coefficients ranging from −1 to +1 and a higher coefficient indicating a model with more coherent clusters. The silhouette coefficients of the two-, three-, four-, and five-cluster solutions were 0.54, 0.52, 0.56, and 0.61, respectively, implying reasonable cluster structure (Kaufman & Rousseeuw, 2009). Though the five-cluster solution had the highest overall goodness-of-fit, it was not selected because its smallest cluster contained only one student. Both the two- and three-cluster solutions had a smaller overall goodness-of-fit and lower interpretability than the four-cluster solution. Given the cluster sizes in each cluster solution, the overall goodness-of-fit, and the interpretability based on students' system usage data and the research context, the four-cluster solution (see Table 2 and Fig. 2) was chosen as the most meaningful and interpretable solution.

Considering the small sample size in each cluster, no ANOVA tests were conducted to identify statistical differences in each variable among the four clusters. The interpretation of each cluster was based on the arithmetic values of each variable. As mentioned in the Methodology section, students' engagement in AI-supported learning was represented by the average frequencies of English shadowing, listening practice, and vocabulary learning.

The comparisons of the four clusters of students (see the final row in Table 2) indicate that Cluster 1 (C1, $N = 4$) was characterized by moderate engagement and the highest score, thus, being named the most effective learner. Cluster 2 (C2, $N = 2$) featured the lowest

Table 2
Summary of the four-cluster solution.

Clusters		Frequency of ES		Frequency of LP		Frequency of VL		Scores of ES	
N (%)		*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
C1	4 (25%)	51.75	1.26	22.50	3.11	16.00	7.44	92.30	1.023
C2	2 (12.50%)	51.50	6.37	27.50	6.36	6.00	0.00	83.05	1.48
C3	7 (43.80%)	50.571	3.46	31.86	2.27	10.43	6.13	90.53	1.30
C4	3 (18.80%)	55.67	4.04	24.67	5.13	57.00	9.54	88.73	3.02
Comparisons		[$C1_{sum} = 90.25$; $C2_{sum} = 85.00$; $C3_{sum} = 92.86$; $C4_{sum} = 137.33$][a] C4 > C3 > C1 > C2						C1 > C3 > C4 > C2	

Note. ES = English shadowing; LP = Listening practice; VL = Vocabulary learning.
[a] Comparisons were calculated based on summed values of the three engagement variables.

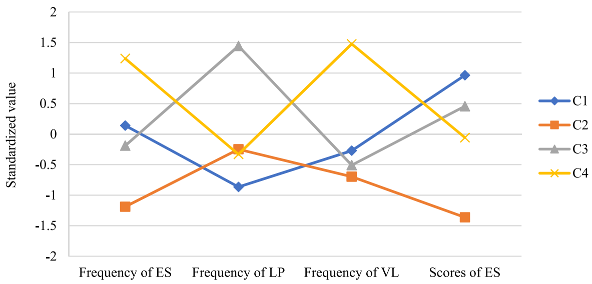

Fig. 2. Distribution of the four clusters across the four variables. *Note.* C1 ($N = 4$); C2 ($N = 2$); C3 ($N = 7$); C4 ($N = 3$).

engagement and the lowest score and was labeled as passive learners. Cluster 3 (C3, $N = 7$) had high engagement as well as a high score and was thus named as well-balanced learners. And Cluster 4 (C4, $N = 3$) was characterized by nearly the highest engagement but a moderate score, therefore, being labeled as inefficient learners. As a result, the cluster analysis overall suggests that the four clusters of students likely follow the descending order of C1, C3, C4, and C2, in terms of the quality of their engagement in AI-supported language learning. Fig. 2 also illustrates the distribution of the four clusters across the four variables to enable a visual understanding of the four clusters.

4.2. *Alignment of the cluster analysis with epistemic network analysis*

Fig. 3 illustrates the distribution of the four clusters' centroids across the four quadrants in the network space. To further determine the quality of each quadrant, we can refer to any one of the four clusters' networks for specific information. C1's network in Fig. 4 was used here for illustration purposes, noting that the placement of all nodes across the four clusters' networks remained unchanged. The first quadrant (top right) has two elements related to social and cognitive presences: *emotional expression and problem-solving*. The second quadrant (top left) contains *the surface approach to learning* and *learning enjoyment*. The third quadrant (bottom left) has three elements related to teaching presence (*instructional design*), approaches to learning (*organized approach*), and learning motivation (*intrinsic motivation*). The fourth quadrant (bottom right) involves five elements related to all three presences (social presence: *affinity*; cognitive presence: *agentic exploration*; teaching presences: *feedback* and *guidance & facilitation*), learning motivation (*extrinsic motivation*), and particularly, approaches to learning (*deep approach*).

The elements in each quadrant collectively characterize the very quadrant and therefore the centroids located in the quadrant (Shaffer & Ruis, 2017). A quadrant with more desirable elements (e.g., the three presences, intrinsic motivation, and deep as well as organized approaches to learning) tends to be indicative of the centroids inside it demonstrating a higher quality of learner-AI

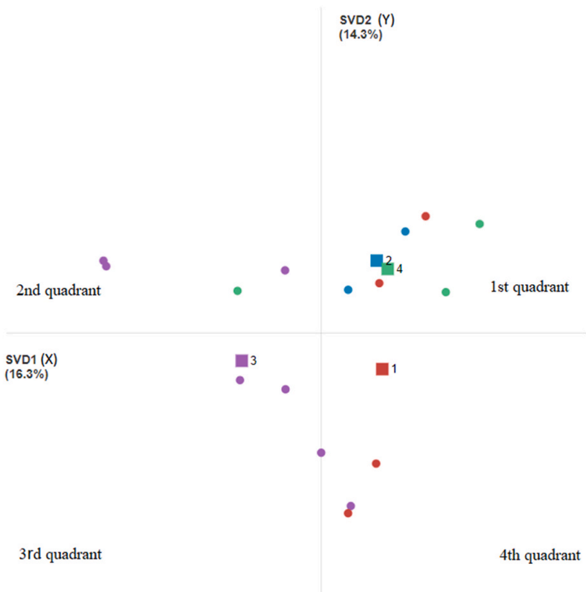

Fig. 3. Distribution of the four clusters' centroids. *Note.* The rectangle represents the centroid of each cluster, which is randomly assigned a color; The dot represents the participant in each cluster with the cluster's color; SVD = Singular Value Decomposition; ENA chooses two dimensions based on SVD as the X and Y axes that best account for the variance in the data. . (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

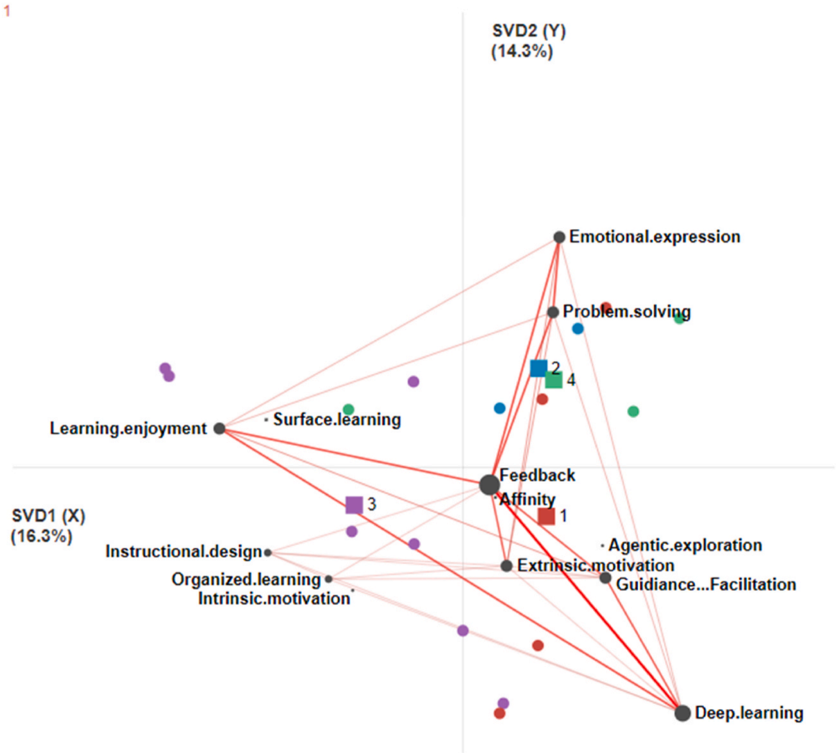

Fig. 4. C1 students' network. *Note*. C1's centroid and network are in red. The edge width represents the frequency of co-occurrences between two codes. . (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

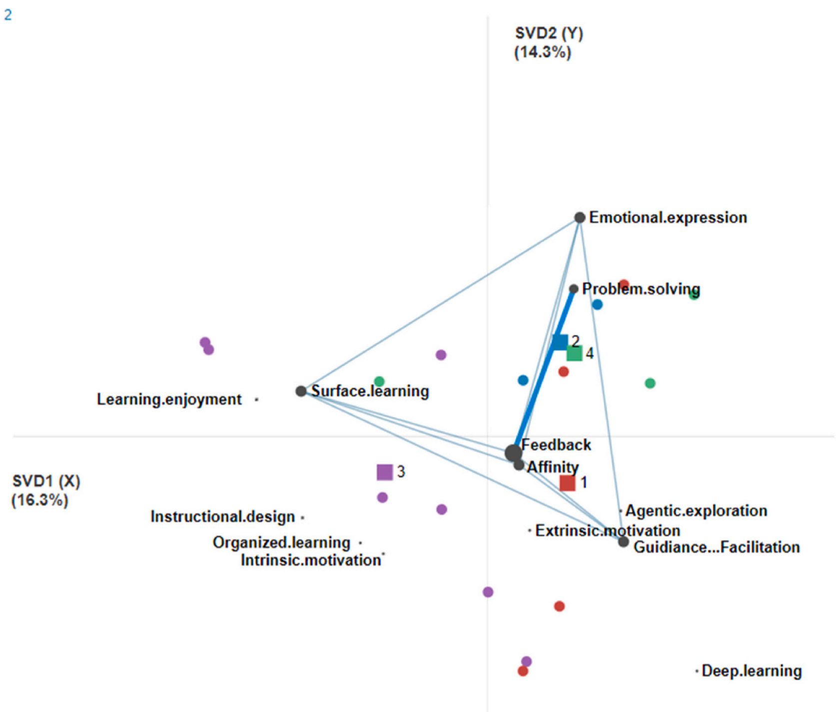

Fig. 5. C2 students' network. *Note*. C2's centroid and network are in blue. The edge width represents the frequency of co-occurrences between two codes. . (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

interactions.

Following this line of reasoning, as shown in Fig. 4, the distribution of the nodes in the four quadrants indicates that the fourth quadrant exhibits the highest quality of learner-AI interactions, followed by the third quadrant. The second quadrant may represent the lowest quality of learner-AI interactions among the four quadrants as it only involves a surface approach to learning and L2 learning enjoyment. The first quadrant ends up in third place. Given that each quadrant characterizes the centroids within it, therefore, C1 in the fourth quadrant may demonstrate the highest quality of learner-AI interactions, followed by C3 located in the third quadrant, which again is followed by C2 and C4 located in the first quadrant. Nonetheless, more information is needed to further determine the ranking of C2 and C4.

As centroids with similar locations in the network space stand for networks with similar patterns (Shaffer et al., 2016), it can be seen from Figs. 5 and 7 in subsequent sections that C2's network is similar to that of C4. In addition, along the y-axis, the lower dimension contains greater and more desirable elements (e.g., agentic exploration and a deep approach to learning) than the upper dimension. This implies that the lower the centroid is located, the higher the quality of learner-AI interactions the centroid represents. As shown in Table 3, C4 is lower than C2 along the y-axis, and C1 is located at the lowest place. Therefore, C4 may exhibit a better quality of learner-AI interactions than C2.

On the whole, based on the analyses of the four quadrants and the four clusters' centroids within them, it can be concluded that the quality of learner-AI interactions of the four clusters of students follows the descending order of C1, C3, C4, and C2. Therefore, the results of the cluster analysis of students' system usage data (see Section 4.1 above) are generally aligned with the ENA results concerning the students' retrospective reflection on their interactions with the AI coach.

4.3. Analyses of the networks of the four clusters of students

4.3.1. C1: effective learners

Fig. 4 shows that most elements were interconnected in C1's network. In particular, one significant and obvious difference between C1 and the other three clusters (see Figs. 5–7 in the subsequent sections) is that the element of the *surface approach to learning* was not connected to any other elements in C1's network, and that the *deep approach to learning* had frequent co-occurrences with the *feedback* given by the AI coach, *emotional expression*, and *L2 learning enjoyment* students experienced in AI-supported language learning.

By returning to the raw data of C1 students' reflection essays, C1's network pattern can be triangulated. Overall, C1 students tended to adopt the deep approach to language learning when interacting with the AI coach. They cherished the feedback from the AI coach but did not follow it mechanically. Instead, as revealed by one student of C1 in his reflection essay, "I tried to link what I learned from the AI coach to my textbook and daily life, enhancing my understanding of the words and sentences." And C1 students all indicated that they enjoyed the learning process. However, C1's network also suggests that C1 students made no connections to *agentic exploration* and *intrinsic motivation*, implying that the students may attend AI-supported learning more out of the requests from their parents or teachers than out of their personal interest. As a result, their engagement in the learning activities was not the highest.

Therefore, C1's network partially explains why C1 students achieved the highest scores but demonstrated moderate engagement in AI-supported learning, as evidenced by the cluster analysis of the students' system usage data. Nonetheless, C1 students can also be interpreted as the most effective learners based on both the network analysis and the cluster analysis across all students.

4.3.2. C2: Passive learners

As illustrated in Fig. 5, C2 students made stronger connections to *feedback* and *problem-solving*. There were also interconnections among *surface approach to learning*, *guidance & facilitation*, and *emotional expression*. However, they made no connections to more desirable elements such as *motivation*, *learning enjoyment*, and *deep approach to learning*.

A closer analysis of the students' reflection essays suggests that C2 students tended to passively interact with the AI coach, which was perceived by them as more of a tool than a social agent. They were preoccupied with describing the functions of the AI coach, emphasizing its dominant role in EFL learning. For instance, one student indicated that "… it (the AI coach) improves our oral English through several ways, such as English shadowing, mimicking picture books, and memorizing vocabulary …" Another student wrote, "it (the AI coach) contains many resources linked to our textbooks, picture books, and movie clips for budding practice." C2 students' learning approaches seemed to be limited to memorizing and reproducing vocabulary and sentences based on the AI coach's feedback. Though they held positive attitudes toward the AI coach, appreciating its usefulness in improving their English competence, they did not perceive much enjoyment in AI-supported learning, at least as reported in their reflection essays.

Consequently, such a passive learning mode likely led to C2 students inactively interacting with the AI coach for language learning, resulting in the lowest engagement and lowest learning performance among the four clusters of students. In addition, C2's network has

Table 3
The coordinates of the four clusters of students' centroids.

Cluster	Axes	
	X	Y
C1	0.45	−0.26
C2	0.41	0.53
C3	−0.58	−0.2
C4	0.49	0.47

the simplest structure of connections among the four clusters of students when comparing Fig. 5 with Fig. 4 for C1, Fig. 6 for C3, and Fig. 7 for C4.

Overall, C2's network pattern and its cluster analysis outcome suggest that passively following the AI coach while experiencing little enjoyment may be linked to a reduced sense of learner-AI learning community, leading to low engagement as well as low performance.

4.3.3. C3: well-balanced learners

Fig. 6 shows that C3's network demonstrates more interconnections among all nodes overall, as compared with the other three clusters of students. This echoes the well-balanced characteristics discovered in the cluster analysis of C3 students (see Table 2) who showed relatively active engagement in AI-supported language learning and achieved relatively high performance.

The well-connected network pattern of C3 students was fleshed out by their reflection essays. They all indicated their obsession with the AI coach. For instance, one student wrote, "I like to learn English with the AI coach …. I practice English every day." Another student pointed out that "I have used the AI coach for almost half a year [more than three months actually]. I am very happy that my English teacher and my mom praise the progress I have made in English learning." In tackling the challenges that emerged, one student indicated that "during the learning process, I sometimes felt discouraged as I could not pronounce correctly the words and phrases even when I practiced twenty or thirty times, …but I persisted and was super happy when I finally made it."

Given that C3 and C1 demonstrated similar network patterns as shown in Figs. 6 and 4, the two networks were superimposed to identify possible differences between them, meanwhile hopefully providing insights into their differences in the cluster analysis reported in Section 4.1. As illustrated in the superimposed network in Appendix A, one easily discernible difference is that C1 made stronger connections to the *deep approach to learning* than C3, as indicated by the thickness of the edges linking different nodes.

On the whole, the well-connected network pattern of C3 students implies that they could engage in the learner-AI learning community with motivation and perceive the interactions among teaching, social, and cognitive presences. They likely approached their L2 learning in different ways and were capable of enjoying the process. However, possibly due to relatively weaker use of the deep approach to AI-supported learning by C3 students than C1 students, their performance (see Table 2) may not be as high as that of C1 students.

4.3.4. C4: Inefficient learners

As illustrated in Fig. 7, C4 students made more and stronger connections to *feedback, guidance and facilitation*, and *emotional expression*. Similar to C2's network, the *deep approach to learning* was not connected to any nodes, and the *surface approach to learning*

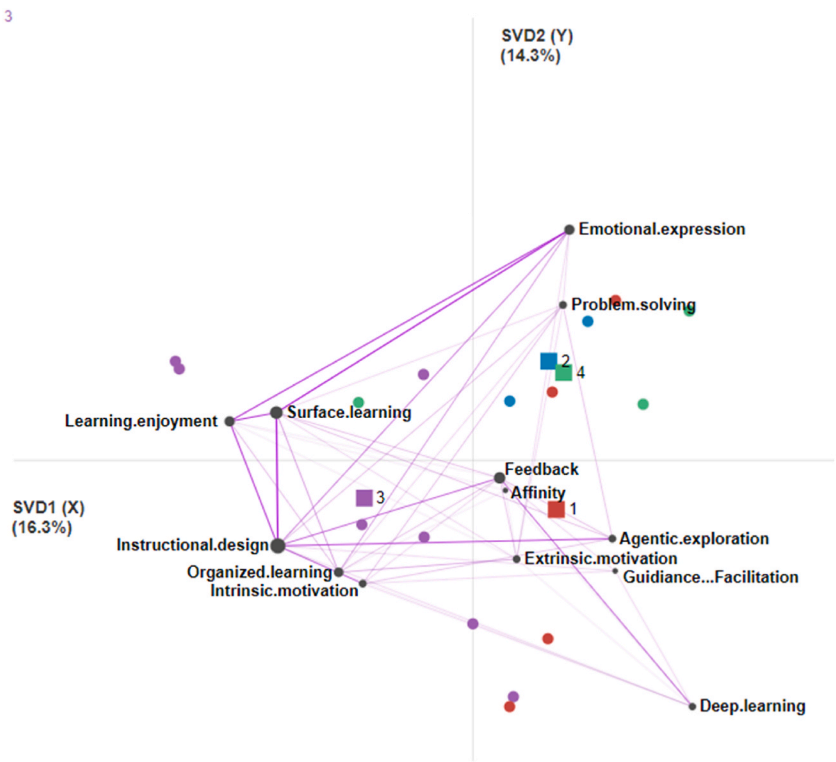

Fig. 6. C3 students' network. *Note*. C3's centroid and network are in purple. The edge width represents the frequency of co-occurrences between two codes. . (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Fig. 7. C4 students' network. *Note*. C4's centroid and network are in green. The edge width represents the frequency of co-occurrences between two codes. . (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

was linked to the node of *feedback*.

A closer look at the reflection essays of C4 students indicates that C4 students quite enjoyed interacting with the AI coach. They were responsive to the feedback provided by the AI coach and closely followed its guidance, thus perceiving a strong teaching presence. However, they mostly stressed the importance of mimicking and memorization techniques in interacting with the AI coach. Their reflection essays connote a dominant role played by the AI coach and a lack of agentic actions from the students. Therefore, even though C4 students demonstrated high engagement as shown by their system usage data, they tended to adopt the surface approach to AI-supported language learning, eventually failing to achieve very high performance, as revealed in the cluster analysis.

At first glance, the network patterns of C4 and C2 are similar. However, after superimposing the networks of C2 and C4 (see Appendix B), it can be seen that the main differences between C2 and C4 lie in the fact that C4 students tended to experience enjoyment and motivation when interacting with the AI coach, thus demonstrating higher engagement in AI-supported learning than C2 students.

Overall, C4's network suggests that high engagement in AI-supported learning but mechanically following AI instruction does not likely guarantee high learning performance.

5. Discussion

Following the CoI and SAL frameworks, this study combined cluster analysis and epistemic network analysis to investigate how students of distinct types interacted with the AI coach for L2 learning and to reveal what matters in AI-supported language learning. Overall, three key findings can be derived.

5.1. Not everyone can benefit from the potential promised by AI

As revealed in this study, the cluster analysis yielded four distinct clusters of students based on their engagement with the AI coach and learning scores. The subsequent network analysis of their reflection essays on their interaction experience with the AI coach unveiled distinct types of human-AI interactions. Overall, C1 students seemed to be the most effective learners, featuring strong connections to the deep approach to learning and no connections to the surface approach to learning, while experiencing social, cognitive, and teaching presences, as well as L2 learning enjoyment when interacting with the AI coach. In contrast, C2 students were characterized by passive interactions with the AI coach, using the surface approach to learning and experiencing limited social, cognitive, and teaching presences as well as limited enjoyment in AI-supported learning, eventually demonstrating the lowest performance.

C3 students tended to be well-balanced learners drawing on different learning approaches when appropriate, following AI's personalized instruction while maintaining individual agentic actions, as well as experiencing enjoyment in interacting with the AI coach. C4 students seemed to be inefficient learners, characterized by high engagement in AI-supported learning but moderate performance, mainly relying on the surface approach to learning despite enjoying interacting with the AI coach.

Garrison and Cleveland-Innes (2005) pointed out that the student-instructor interaction had to be of high quality, involving critical discourse, specific goals, and social presence, so as to facilitate students' achieving higher levels of learning. Nonetheless, students of different competence levels may have distinct interpretations of the same learning activity, regardless of how optimally designed the activity is, and therefore adopt distinct approaches to learning. This may apply to the student-AI interaction setting, considering that humans' interactions with AI agents are similar to human-human interactions, as our perceptions of AI agents normally conform to the similar evaluations we make of other people (Reeves et al., 2020). In a similar vein, the students' abilities to benefit from AI-supported learning may vary.

Even though AI technologies hold the promise of transforming student learning and instructional design (Xu & Ouyang, 2022), the results of the four patterns of AI-supported language learning demonstrated by the four clusters of students imply that not every student can actually benefit from the potential promised by the AI coach.

5.2. The deep approach to learning may amplify the potential benefits of AI's personalized guidance and strengthen the sense of the human-AI learning community

The comparisons among the four clusters' networks (see Figs. 4–7) indicate that the deep approach to learning may play a more important role in both the networks of C1 and C3 than those of C2 and C4, where no connections were made to the *deep approach to learning*. Teaching presence (*instructional design*, *feedback*, and *guidance & facilitation*), in particular, as well as *social and cognitive presences*, were more densely interrelated and related to other elements in the networks of C1 and C3 than those of C2 and C4. Furthermore, the comparisons of the four clusters' learning performance as shown in the cluster analysis (see Table 2) reveal that C1 and C3 achieved higher scores than C2 and C4.

Based on the comparisons above and the interpretations of the four clusters' networks in Section 4.3, while considering the critical importance of the deep approach to learning in either conventional (e.g., Ellis & Bliuc, 2019; Garrison & Cleveland-Innes, 2005) or AI-supported settings (e.g., Wang et al., 2022), it can be tentatively concluded that the deep approach to learning may enhance the benefits of the personalized feedback from the AI coach and strengthen the students' sense of human-AI learning community, thereby improving the quality of student learning and enhancing their enjoyment experience in interacting with the AI coach.

These results corroborate the conceptual argument that in AI-supported learning, students need to take agency and initiative in their learning endeavors to achieve high performance (Xu & Ouyang, 2022). The essential role of the deep approach to learning in the human-AI learning community also echoes Akyol et al. (2011) in conventional digital learning, who investigated the relationship between the SAL and CoI frameworks and argued that deep and meaningful learning is at the core of the CoI model. The emergence of and interplay among social, cognitive, and teaching presences depend on students engaging in learning activities through higher-order learning such as critical thinking and sustained reflection.

In addition, the four network graphs suggest that the teaching presence students perceived in interacting with the AI coach, particularly *feedback* and *guidance & facilitation*, are widely linked to other elements in the networks. This pattern is consistent with Turula (2017), who noted that in the context of human-human interactions, teaching presence is critically important for sustained communication and meaning construction in digital learning settings. It serves as a binding element in developing a community of inquiry for language learning (Garrison et al., 2010; Turula, 2017).

5.3. Passively or mechanically following AI's instruction, albeit with high levels of participation, may decrease the sense of the human-AI learning community and cause low performance

Though C4 students demonstrated higher levels of participation in AI-supported learning than C2 students as suggested by the cluster analysis, they had similar networks (see Figs. 7 and 5), particularly in the co-occurrences between the *surface approach to learning* and *teaching presence* when interacting with the AI coach. By triangulating with their reflection essays, C4 students seemed to be active in exploring the different functions of the AI coach, nonetheless, with limited knowledge regarding how to effectively use them to acquire EFL knowledge. They were inclined to emphasize the usage details and to mechanically follow the AI's instruction. As for C2 students, they passively interacted with the AI coach, limiting their learning to memorization and reproduction of vocabulary and sentences.

According to Li and Cutting (2011) and Mak and Chik (2011), EFL learning resembles rote learning for some students, marked by memorization and repetition, which typically characterize a surface approach to learning. This misbelief particularly appears among many Chinese EFL learners who prefer rote learning (Li & Cutting, 2011), as reflected by what C2 and C4 students did in the present study. Passively or mechanically accessing and memorizing information through interacting with AI makes it difficult for the students to integrate new knowledge into pre-existing mental structures.

Such a surface approach to learning, though facilitated by AI technologies, may underplay the role of humans while overplaying the role of AI, causing a dichotomy of humans and AI and severing the human-AI learning community (Mervich, 2020), and eventually leading to relatively low learning performance.

These research results are largely consistent with those of prior studies (e.g., Ellis & Bliuc, 2019; Rao et al., 2007) on SAL in conventional digital and non-digital learning settings, which found that the surface approach such as paraphrasing, re-reading, and rote memorization was mostly adopted by less successful students. Even in AI-supported learning settings, if students adopt surface processing strategies, they tend to be easily distracted by inessential information, eventually overlooking the purpose of using AI for L2 learning (Niu et al., 2022; Thompson, 2013).

6. Contributions and implications

This study contributes to AI-supported language learning in several ways. First, the present study advances the state-of-the-art understanding of how learners interact with AI technologies for language learning. Though AI has been considered useful and effective in improving learners' L2 competence and forging positive learning experiences, little is known regarding how human-AI interactions actually function for L2 learning. Though few studies (e.g., Lee & Jeon, 2022; Moussalli & Cardoso, 2020) have tried to explore learner-AI interactions by investigating their perceptions of AI, it would be difficult to gain any real insight thereof without the support of actual human-AI interaction data. This study triangulated students' actual usage data of the AI with their reflection essays, thereby standing in a better position to reveal human-AI interactions in language learning.

Second, this study contributes to the CoI theory by expanding its application to human-AI interactions and revealing the role of students' approaches to learning in building a sense of a human-AI learning community. AI agents used in education are not simply instructional tools, as they can engage in social interactions with learners and establish and sustain social relationships with them (Jeon, 2022; Reeves et al., 2020), thereby creating a learner-AI community. Applying the CoI and SAL frameworks, which have been predominantly used in human-human interactions, to human-AI interactions can provide us with an innovative angle to examine learners as well as AI. In AI-supported learning settings, the students adopting a deep approach to learning rather than a surface approach, can best appreciate the humanized dimension of and develop social bonds with AI, thereby being more capable of enjoying the benefits of AI. As such, the empirical evidence derived from this study casts doubt on the misbelief of EFL learning as a surface approach to learning characterized by rote learning (Larsen-Freeman, 2012).

Third, cluster analysis and ENA jointly provide a viable solution to identifying fine-grained differences between learners in AI-supported learning by first categorizing students into different clusters of performance and then using ENA to visually analyze each cluster's learning experiences with AI as well as to perform pairwise group comparisons. Prior studies often rely on conventional qualitative and/or qualitative context analysis (e.g., Hwang et al., 2018; Sun et al., 2020) and even machine learning techniques (e.g., Fan, Jiang, Liu, & Zhou, 2021; Peng & Xu, 2020) for learning behavior analyses, which, albeit achieved useful results, tended to treat learning behaviors as isolated from one another and normally fell short of presenting a visualized representation of them. Contrastingly, the joint analysis approach employed in this study allows for a graphical and in-depth understanding of the complexity of students' social-cognitive interactions with AI for L2 learning and how learning behaviors are interrelated to foster different learning experiences.

In addition, this study holds implications for the field of AI-supported learning. First, researchers and practitioners of AI-supported learning are advised to take a critical and realistic view on AI technologies. People normally hold a presumption that AI technologies are beneficial for education, and few have ever questioned such a presumption. Though some researchers (e.g., Xu & Ouyang, 2022) have conceptually argued that AI technologies cannot ensure quality instruction and high learning outcomes, few have empirically contemplated such an argument. The benefits each student can gain from AI technologies may depend on how students learn, as suggested by this study. They may also depend on the fit between AI technologies' manifestations (e.g., social roles and forms) and students' learning approaches and expectations. For different groups of students, AI agents with anthropomorphic or cartoon-like forms acting as tutors, peers, or assistants could work differently (Randall, 2019).

Second, while many claims have been made about the potential impact of AI on education, more attention should be devoted to identifying how students should interact with AI technologies to optimize their learning gains and what pedagogy can be developed to better use AI for language learning as well as other disciplines. Third, as the deep approach to learning is crucial to learner-AI interactions, human instructors should facilitate students' adoption of it, instead of simply leaving them to AI. As was discovered in this study, there was still a group of learners who could not find effective ways to interact with AI for high performance. Therefore, human instructors can intervene when appropriate and facilitate students to consciously reflect on and assess their learning progress based on AI's feedback and to apply new knowledge acquired from learner-AI interactions to new settings.

And finally, the AI agents for education are different from those in general social settings where much attention has been paid to the design of AI to make them appear more attractive or sociable (Reeves et al., 2020). In view of the significant role played by teaching presence, particularly personalized feedback provided by AI, in linking together most other elements (e.g., SAL and motivation) in EFL learning, AI researchers in the field of education may consider devoting more attention to improving the accuracy and power of the algorithms used for personalized feedback so as to optimize student learning.

7. Limitations and future research

The findings of this study should be interpreted with the following limitations in mind. First, this study involved a small number of participants studying EFL and was conducted in a primary school. Future research seeking to generalize the research findings can be conducted with more participants from diverse disciplines and school levels. Second, the present study mainly relied on two data sources: students' actual usage data and their reflection essays on their using experiences. To obtain potentially more useful and valuable findings, future research is advised to adopt more data collection techniques, such as follow-up interviews and questionnaires specifically assessing the CoI presences and SAL. Third, the four clusters of students generated by the cluster analysis had unequal sizes, with C2 having only two students. Future studies may examine whether the (a)symmetry of student clusters may have an influence on research results by experimenting with different cluster analysis techniques and cluster sizes. Fourth, the AI coach used in this study was an anthropomorphic virtual agent, thereby possibly limiting the research findings to such a type of AI agent. As such, future studies are advised to investigate more types of AI agents, such as mechanomorphic and cartoon-like agents, so as to unveil possible differences between them in terms of human-AI interactions in educational settings, which can inform the design of AI agents for education and pedagogy related to them. Fifth, even though the CoI and SAL frameworks empowered us to better understand how the students interacted with the AI coach for language learning, the results of this study are also, in turn, confined by the combined frameworks. Therefore, researchers in future studies may explore and experiment with alternative promising theories to gain a more comprehensive understanding of human-AI interactions in the field of education in hopes of best facilitating learners' use of AI technologies for a joyful learning experience, as well as high performance.

Credit author statement

Xinghua Wang: Conceptualization, Investigation, Methodology, Software, Formal analysis, Writing-original draft; **Qian Liu**: Conceptualization, Investigation, Formal analysis, Writing-original draft; **Hui Pang**: Resources, Data curation, Methodology, Writing-original draft, Writing-review & editing; **Seng Chee Tan**: Conceptualization, Methodology, Writing-review & editing. **Jun Lei:** Conceptualization, Methodology, Writing-review & editing. **Matthew P. Wallace**: Conceptualization, Methodology, Writing-review & editing. **Linlin Li**: Data curation, Writing-review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data availability

The data that has been used is confidential.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compedu.2022.104703.

References

Akyol, Z., Vaughan, N., & Garrison, D. R. (2011). The impact of course duration on the development of a community of inquiry. *Interactive Learning Environments, 19*(3), 231–246. https://doi.org/10.1080/10494820902809147

Alemi, M., Meghdari, A., & Ghazisaedy, M. (2015). The impact of social robotics on L2 learners' anxiety and attitude in English vocabulary acquisition. *International Journal of Social Robotics, 7*(4), 523–535. https://doi.org/10.1007/s12369-015-0286-y

Awada, G. (2022). Robotics and Weblog integration into EFL curriculum of middle and secondary schools: Exploratory study. *Language Teaching Research*. https://doi.org/10.1177/13621688221091611, 13621688221091611.

Benassi, M., Garofalo, S., Ambrosini, F., Sant'Angelo, R. P., Raggini, R., De Paoli, G., … Piraccini, G. (2020). Using two-step cluster analysis and latent class cluster analysis to classify the cognitive heterogeneity of cross-diagnostic psychiatric inpatients. *Frontiers in Psychology, 11*, 1085. https://doi.org/10.3389/fpsyg.2020.01085

Berendt, B., Littlejohn, A., & Blakemore, M. (2020). AI in education: Learner choice and fundamental rights. *Learning, Media and Technology, 45*(3), 312–324. https://doi.org/10.1080/17439884.2020.1786399

van den Berghe, R., Verhagen, J., Oudgenoeg-Paz, O., Van der Ven, S., & Leseman, P. (2019). Social robots for language learning: A review. *Review of Educational Research, 89*(2), 259–295. https://doi.org/10.3102/0034654318821286

Biggs, J. (1993). What do inventories of students' learning processes really measure? A theoretical review and clarification. *British Journal of Educational Psychology, 63*(1), 3–19. https://doi.org/10.1111/j.2044-8279.1993.tb01038.x

Biggs, J., & Tang, C. (2011). *Teaching for quality learning at university*. UK): McGraw-hill education.

Cai, S., Liu, E., Yang, Y., & Liang, J. C. (2019). Tablet-based AR technology: Impacts on students' conceptions and approaches to learning mathematics according to their self-efficacy. *British Journal of Educational Technology, 50*(1), 248–263. https://doi.org/10.1111/bjet.12718

Cheng, K. H. (2017). Exploring parents' conceptions of augmented reality learning and approaches to learning by augmented reality with their children. *Journal of Educational Computing Research, 55*(6), 820–843. https://doi.org/10.1177/0735633116686082

Chen, H. H.-J., Yang, C. T.-Y., & Lai, K. K.-W. (2020). Investigating college EFL learners' perceptions toward the use of Google Assistant for foreign language learning. *Interactive Learning Environments*, 1–16. https://doi.org/10.1080/10494820.2020.1833043

Dizon, G. (2020). Evaluating intelligent personal assistants for L2 listening and speaking development. *Language, Learning and Technology, 24*(1), 16–26.

Ellis, R. A., & Bliuc, A.-M. (2019). Exploring new elements of the student approaches to learning framework: The role of online learning technologies in student learning. *Active Learning in Higher Education, 20*(1), 11–24. https://doi.org/10.1177/1469787417721384

Ellis, R. A., Goodyear, P., Calvo, R. A., & Prosser, M. (2008). Engineering students' conceptions of and approaches to learning through discussions in face-to-face and online contexts. *Learning and Instruction, 18*(3), 267–282. https://doi.org/10.1016/j.learninstruc.2007.06.001

Engwall, O., & Lopes, J. (2020). Interaction and collaboration in robot-assisted language learning for adults. *Computer Assisted Language Learning*, 1–37. https://doi.org/10.1080/09588221.2020.1799821

Fan, J., Jiang, Y., Liu, Y., & Zhou, Y. (2021). Interpretable MOOC recommendation: A multi-attention network for personalized learning behavior analysis. *Internet Research, 32*(2), 588–605. https://doi.org/10.1108/INTR-08-2020-0477

Felix, C. V. (2020). The role of the teacher and AI in education. In E. Sengupta, P. Blessinger, & M. S. Makhanya (Eds.), *International perspectives on the role of technology in humanizing higher education. (Innovations in higher education teaching and learning* (pp. 33–48). Emerald. https://doi.org/10.1108/S2055-364120200000033003.
Garrison, D. R., Anderson, T., & Archer, W. (2010). The first decade of the community of inquiry framework: A retrospective. *The Internet and Higher Education, 13*(1), 5–9.
Garrison, D. R., & Arbaugh, J. B. (2007). Researching the community of inquiry framework: Review, issues, and future directions. *The Internet and Higher Education, 10*(3), 157–172. https://doi.org/10.1016/j.iheduc.2007.04.001
Garrison, D. R., & Cleveland-Innes, M. (2005). Facilitating cognitive presence in online learning: Interaction is not enough. *American Journal of Distance Education, 19*(3), 133–148. https://doi.org/10.1207/s15389286ajde1903_2
Gibbs, G., & Coffey, M. (2004). The impact of training of university teachers on their teaching skills, their approach to teaching and the approach to learning of their students. *Active Learning in Higher Education, 5*(1), 87–100. https://doi.org/10.1177/1469787404040463
Hailikari, T. K., & Parpala, A. (2014). What impedes or enhances my studying? The interrelation between approaches to learning, factors influencing study progress and earned credits. *Teaching in Higher Education, 19*(7), 812–824. https://doi.org/10.1080/13562517.2014.934348
Hamada, Y. (2016). Shadowing: Who benefits and how? Uncovering a booming EFL teaching technique for listening comprehension. *Language Teaching Research, 20*(1), 35–52. https://doi.org/10.1177/1362168815597504.
Heift, T. (2002). Learner control and error correction in ICALL: Browsers, peekers, and adamants. *Calico Journal, 19*(2), 295–313.
Heift, T., & Schulze, M. (2007). *Errors and intelligence in computer-assisted language learning: Parsers and pedagogues*. New York, NY: Routledge.
Herrmann, K. J., Bager-Elsborg, A., & McCune, V. (2017). Investigating the relationships between approaches to learning, learner identities and academic achievement in higher education. *Higher Education, 74*(3), 385–400. https://doi.org/10.1007/s10734-016-9999-6
Hwang, W.-Y., Chen, H.-R., Chen, N.-S., Lin, L.-K., & Chen, J.-W. (2018). Learning behavior analysis of a ubiquitous situated reflective learning system with application to life science and technology teaching. *Educational Technology & Society, 21*(2), 137–149.
Jeon, J. (2022). Exploring AI chatbot affordances in the EFL classroom: Young learners' experiences and perspectives. *Computer Assisted Language Learning*, 1–26. https://doi.org/10.1080/09588221.2021.2021241
Kaufman, L., & Rousseeuw, P. J. (2009). *Finding groups in data: An introduction to cluster analysis*. John Wiley & Sons.
Kent, P., Jensen, R. K., & Kongsted, A. (2014). A comparison of three clustering methods for finding subgroups in MRI, SMS or clinical data: SPSS TwoStep cluster analysis, latent gold and SNOB. *BMC Medical Research Methodology, 14*(1), 1–14. https://doi.org/10.1186/1471-2288-14-113
Larsen-Freeman, D. (2012). On the roles of repetition in language teaching and learning. *Applied Linguistics Review, 3*(2), 195–210. https://doi.org/10.1515/applirev-2012-0009
Lee, S., & Jeon, J. (2022). Visualizing a disembodied agent: Young EFL learners' perceptions of voice-controlled conversational agents as language partners. *Computer Assisted Language Learning*, 1–26. https://doi.org/10.1080/09588221.2022.2067182
Li, X., & Cutting, J. (2011). Rote learning in Chinese culture: Reflecting active confucian-based memory strategies. In L. Jin, & M. Cortazzi (Eds.), *Researching Chinese learners* (pp. 21–42). Palgrave Macmillan. https://doi.org/10.1057/9780230299481_2.
Lindblom-Ylänne, S., Parpala, A., & Postareff, L. (2019). What constitutes the surface approach to learning in the light of new empirical evidence? *Studies in Higher Education, 44*(12), 2183–2195. https://doi.org/10.1080/03075079.2018.1482267
Lin, C.-J., & Mubarok, H. (2021). Learning analytics for investigating the mind map-guided AI Chatbot approach in an EFL flipped speaking classroom. *Educational Technology & Society, 24*(4), 16–35. https://www.jstor.org/stable/48629242.
Lomicka, L. (2020). Creating and sustaining virtual language communities. *Foreign Language Annals, 53*(2), 306–313. https://doi.org/10.1111/flan.12456
Luo, B. (2016). Evaluating a computer-assisted pronunciation training (CAPT) technique for efficient classroom instruction. *Computer Assisted Language Learning, 29*(3), 451–476. https://doi.org/10.1080/09588221.2014.963123
Mak, B., & Chik, P. (2011). Differences in perceived approaches to learning and teaching English in Hong Kong secondary schools. *Educational Review, 63*(2), 195–218. https://doi.org/10.1080/00131911.2010.534769
Mervich, C. (2020). *The human infrastructure of artificial intelligence*. University of Twente. https://doi.org/10.13140/RG.2.2.20431.10401
Moussalli, S., & Cardoso, W. (2020). Intelligent personal assistants: Can they understand and be understood by accented L2 learners? *Computer Assisted Language Learning, 33*(8), 865–890. https://doi.org/10.1080/09588221.2019.1595664
Natale, S., & Cooke, H. (2020). Browsing with Alexa: Interrogating the impact of voice assistants as web interfaces. *Media, Culture & Society, 43*(6), 1000–1016. https://doi.org/10.1177/0163443720983295
Niu, L., Wang, X., Wallace, M. P., Pang, H., & Xu, Y. (2022). Digital learning of English as a foreign language among university students: How are approaches to learning linked to digital competence and technostress? *Journal of Computer Assisted Learning, 38*(5), 1332–1346. https://doi.org/10.1111/jcal.12679.
Ouyang, F., Chen, S., & Li, X. (2021). Effect of three network visualizations on students' social-cognitive engagement in online discussions. *British Journal of Educational Technology, 52*(6), 2242–2262. https://doi.org/10.1111/bjet.13126
Peng, X., & Xu, Q. (2020). Investigating learners' behaviors and discourse content in MOOC course reviews. *Computers & Education, 143*, Article 103673. https://doi.org/10.1016/j.compedu.2019.103673
Qi, C. (2019). A double-edged sword? Exploring the impact of students' academic usage of mobile devices on technostress and academic performance. *Behaviour & Information Technology, 38*(12), 1337–1354. https://doi.org/10.1080/0144929X.2019.1585476
Randall, N. (2019). A survey of robot-assisted language learning (RALL). *ACM Transactions on Human-Robot Interaction (THRI), 9*(1), 1–36. https://doi.org/10.1145/3345506
Rao, Z., Yongqi Gu, P., Jun Zhang, L., Hu, G. (2007). Reading strategies and approaches to learning of bilingual primary school pupils. *Language Awareness, 16*(4), 243–262. https://doi.org/10.2167/la423.0
Reeves, B., Hancock, J., & Liu, X. (2020). Social robots are like real people: First impressions, attributes, and stereotyping of social robots. *Technology, Mind, and Behavior, 1*(1), 1–37. https://doi.org/10.1037/tmb0000018
Richardson, J. C., Arbaugh, J. B., Cleveland-Innes, M., Ice, P., Swan, K. P., & Garrison, D. R. (2012). Using the community of inquiry framework to inform effective instructional design. In L. Moller, & J. Huett (Eds.), *The next generation of distance education* (pp. 97–125). Springer. https://doi.org/10.1007/978-1-4614-1785-9_7.
Rozgonjuk, D., Saal, K., & Täht, K. (2018). Problematic smartphone use, deep and surface approaches to learning, and social media use in lectures. *International Journal of Environmental Research and Public Health, 15*(1), 92. https://doi.org/10.3390/ijerph15010092
Settles, B., T LaFlair, G., & Hagiwara, M. (2020). Machine learning–driven language assessment. *Transactions of the Association for Computational Linguistics, 8*, 247–263. https://doi.org/10.1162/tacl_a_00310
Shaffer, D., Collier, W., & Ruis, A. (2016). A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data. *Journal of Learning Analytics, 3*(3), 9–45. https://doi.org/10.18608/jla.2016.33.3
Shaffer, D., & Ruis, A. (2017). Epistemic network analysis: A worked example of theory-based learning analytics. In C. Lang, G. Siemens, A. Wise, & D. Gasevic (Eds.), *Handbook of learning analytics* (pp. 175–187). SOLAR, Society for Learning Analytics and Research. https://doi.org/10.18608/HLA17.015.
Smidt, E., Chau, M. H., Rinehimer, E., & Leever, P. (2021). Exploring engagement of users of Global Englishes in a community of inquiry. *System, 98*, Article 102477. https://doi.org/10.1016/j.system.2021.102477
Sun, Y., Guo, Y., & Zhao, Y. (2020). Understanding the determinants of learner engagement in MOOCs: An adaptive structuration perspective. *Computers & Education, 157*, Article 103963. https://doi.org/10.1016/j.compedu.2020.103963
Swan, K., Garrison, D., & Richardson, J. C. (2009). A constructivist approach to online learning: The community of inquiry framework. In C. R. Payne (Ed.), *Information technology and constructivism in higher education: Progressive learning frameworks* (pp. 43–57). https://doi.org/10.4018/978-1-60566-654-9.ch004. IGI global.
Tai, T.-Y., & Chen, H. H.-J. (2020). The impact of Google Assistant on adolescent EFL learners' willingness to communicate. *Interactive Learning Environments*, 1–18. https://doi.org/10.1080/10494820.2020.1841801

Tan, S. C., Wang, X., & Li, L. (2022). The development trajectory of shared epistemic agency in online collaborative learning: A study combing network analysis and sequential analysis. *Journal of Educational Computing Research, 59*(8), 1655–1681. https://doi.org/10.1177/07356331211001562.

Thompson, P. (2013). The digital natives as learners: Technology use patterns and approaches to learning. *Computers & Education, 65*, 12–33. https://doi.org/10.1016/j.compedu.2012.12.022

Tsai, P. S., Chai, C. S., Hong, H. Y., & Koh, J. H. L. (2017). Students' conceptions of and approaches to knowledge building and its relationship to learning outcomes. *Interactive Learning Environments, 25*(6), 749–761. https://doi.org/10.1080/10494820.2016.1178653

Turula, A. (2017). Teaching presence in telecollaboration. Keeping an open mind. *System, 64*, 21–33. https://doi.org/10.1016/j.system.2016.12.001

Umapathy, K., Ritzhaupt, A. D., & Xu, Z. (2020). College students' conceptions of learning of and approaches to learning computer science. *Journal of Educational Computing Research, 58*(3), 662–686. https://doi.org/10.1177/0735633119872659

Underwood, J. (2017). *Exploring AI language assistants with primary EFL students. CALL in a climate of change: Adapting to turbulent global conditions-short papers from EUROCALL*.

de Vries, B. P., Cucchiarini, C., Bodnar, S., Strik, H., & van Hout, R. (2015). Spoken grammar practice and feedback in an ASR-based CALL system. *Computer Assisted Language Learning, 28*(6), 550–576. https://doi.org/10.1080/09588221.2014.889713

Wang, X., Pang, H., Wallace, M. P., Wang, Q., & Chen, W. (2022). Learners' perceived AI presences in AI-supported language learning: a study of AI as a humanized agent from community of inquiry. *Computer Assisted Language Learning*, 1–27. https://doi.org/10.1080/09588221.2022.2056203.

Watson, S. L., Watson, W. R., Richardson, J., & Loizzo, J. (2016). Instructor's use of social presence, teaching presence, and attitudinal dissonance: A case study of an attitudinal change MOOC. *International Review of Research in Open and Distance Learning, 17*(3), 54–74. https://doi.org/10.19173/irrodl.v17i3.2379

Wu, W.-C. V., Hsieh, J. S. C., & Yang, J. C. (2017). Creating an online learning community in a flipped classroom to enhance EFL learners' oral proficiency. *Journal of Educational Technology & Society, 20*(2), 142–157.

Xu, W., & Ouyang, F. (2022). A systematic review of AI role in the educational system based on a proposed conceptual framework. *Education and Information Technologies, 27*, 4195–4223. https://doi.org/10.1007/s10639-021-10774-y

Yu, B. (2019). The predicting roles of approaches to learning, L2 learning motivation, L2 learning strategies and L2 proficiency for learning outcomes: A comparison between mainland and Hong Kong Chinese students. *Educational Studies, 45*(4), 520–532. https://doi.org/10.1080/03055698.2018.1509775

Yu, Z., & Li, M. (2022). A bibliometric analysis of Community of Inquiry in online learning contexts over twenty-five years. *Education and Information Technologies*, 1–20. https://doi.org/10.1007/s10639-022-11081-w