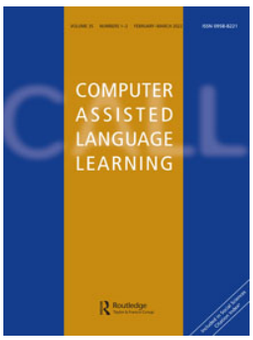

Computer Assisted Language Learning

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ncal20

Learners' perceived AI presences in AI-supported language learning: a study of AI as a humanized agent from community of inquiry

Xinghua Wang, Hui Pang, Matthew P. Wallace, Qiyun Wang & Wenli Chen

To cite this article: Xinghua Wang, Hui Pang, Matthew P. Wallace, Qiyun Wang & Wenli Chen (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, DOI: 10.1080/09588221.2022.2056203

To link to this article: https://doi.org/10.1080/09588221.2022.2056203

COMPUTER ASSISTED LANGUAGE LEARNING
https://doi.org/10.1080/09588221.2022.2056203

Routledge
Taylor & Francis Group

Learners' perceived AI presences in AI-supported language learning: a study of AI as a humanized agent from community of inquiry

Xinghua Wang[a], Hui Pang[a], Matthew P. Wallace[b], Qiyun Wang[c] and Wenli Chen[c]

[a]Normal College, Qingdao University, China; [b]Faculty of Arts and Humanities, University of Macau, China; [c]National Institute of Education, Nanyang Technological University, Singapore

ABSTRACT

This study investigated the application of an artificial intelligence (AI) coach for second language (L2) learning in a primary school involving 327 participants. In line with Community of Inquiry, learners were expected to perceive social, cognitive, and teaching presences when interacting with the AI coach, which was considered a humanized agent. To examine how learners' perceived AI presences were related to their language learning, this study drew on AI usage data, actual learning outcomes, and attitudinal data. Results from hierarchical regression analyses suggest that cognitive presence and learners' affection for AI's appearance were significant predictors of L2 enjoyment, which also positively predicted learning outcomes. The score of English shadowing (representing the quality of AI usage) positively predicted learning outcomes. Contrary to intuition, teaching presence was found to negatively predict learning outcomes. Based on cluster analysis and subsequent MANOVA results, this study indicates that the learners perceiving higher social and cognitive presences via interacting with AI and showing greater affection for AI's appearance tended to use the AI coach more frequently, demonstrate higher L2 enjoyment, and achieve higher learning outcomes. The present study contributes to the limited but increasing knowledge of human-AI interaction in educational settings and carries implications for future efforts on the use of AI for L2 learning.

KEYWORDS

Artificial intelligence; language learning; community of inquiry; human-AI interaction

1. Introduction

The use of artificial intelligence (AI) in the field of education, albeit limited, has been increasing in recent years, particularly for second language (L2) learning (Dizon, 2020; Moussalli & Cardoso, 2020). This is probably due to the fact that L2 classrooms face challenges of limited

CONTACT Hui Pang hopehuipang@gmail.com Normal College, Qingdao University, China
Supplemental data for this article is available online at https://doi.org/10.1080/09588221.2022.2056203.

classroom time and insufficient listening and speaking practice in a stress-free environment (de Vries et al., 2015; Tai & Chen, 2020). And teachers find it nearly impossible to give feedback to every student in a large class with more than 20 students (Luo, 2016). The AI technology presents a potential solution to these challenges by providing learners with endless opportunities for practice without much pressure from teachers and peers and with personalized and just-in-time feedback (Dizon, 2017; Moussalli & Cardoso, 2020). In particular, voice-driven AI such as Amazon's Alexa and Apple's Siri seem to be apt for language learning due to their lifelike interactional skills (Underwood, 2017).

Previous studies on AI-supported language learning have reported a variety of benefits of AI for learners' L2 acquisition and affect (e.g., Leeuwestein et al., 2020; Pennington & Rogerson-Revell, 2019; Underwood, 2017). Nonetheless, many of these studies were short-term case studies lasting a few hours or sessions and were characterized by small sample sizes (Dizon, 2020; Randall, 2019). They were mostly preoccupied with justifying the usefulness of AI for perceived language gains. Few have attempted to examine how AI could support actual learning improvement. Additionally, most research findings were solely obtained from attitudinal data such as surveys and interviews (e.g., Dizon, 2017; Tai & Chen, 2020), which have a discrepancy from actual AI usage data and actual learning outcomes (Randall, 2019), consequently leading to questions about the credibility of the research findings. Furthermore, although previous studies have given AI applications a variety of humanized roles in language learning such as tutors, assistants, teachers, and peers for learners (Engwall & Lopes, 2020; Randall, 2019), to our latest knowledge, no studies have ever practically treated AI applications as humanized agents and investigated the human-AI interactions from the framework of Community of Inquiry (CoI; Garrison & Arbaugh, 2007), specifically, how learners' perceived AI presences (social, cognitive, and teaching presences) affect their learning.

Therefore, this study seeks to bridge these gaps by investigating the application of an AI coach, which was specifically developed for English as a Foreign Language (EFL) learning, in a primary school involving a comparatively larger sample (327 participants). The AI coach was a voice-driven AI application and was humanized as a female teacher often with a smiling face. Its voice is intelligently derived. It was installed on mobile devices such as smartphones and tablets, and could provide unlimited opportunities for English practice, imitation, and personalized feedback on pronunciation. This study lasted two and a half months and combined different data sources, including AI usage data, attitudinal data, and actual learning outcomes, to examine how AI could support language learning following the CoI framework.

Though gender has been considered an important factor affecting social, cognitive, and teaching presences in conventional digital learning environments (C. Lee et al., 2016), little is known regarding its effect in AI-supported learning settings. Moreover, AI's appearance is a non-trivial issue as even a subtle variation of it can affect learners' perceptions of AI (Randall, 2019). However, it is still uncertain whether and how learners' perceptions of AI's appearance can influence their L2 learning. In addition, L2 enjoyment, which refers to the positive emotion in L2 learning that is related to intellectual focus and enhanced attention, is considered one of the most prevailing and salient positive emotions in L2 learning across different contexts and can make students more persistent in the face of challenges and better process target languages (Jiang & Dewaele, 2019; J. S. Lee & Lee, 2021; Pavelescu & Petric, 2018). As such, this study incorporates the factors of gender, learners' affection for AI's appearance, and L2 enjoyment to investigate AI-supported language learning alongside the CoI framework. Overall, the present study aims to address the following research questions:

RQ1: How do learners' perceived AI presences, affection for AI's appearance, and AI usage data predict L2 enjoyment?

RQ2: How do learners' perceived AI presences, L2 enjoyment, affection for AI's appearance, and AI usage data predict actual learning outcomes?

RQ3: How are learners clustered based on their perceived AI presences, gender, and affection for AI's appearance, and how do they differ regarding AI usage data, L2 enjoyment, and actual learning outcomes?

The rest of this paper is organized in the following manner. The theoretical foundation is first given to allow readers an understanding of the application of AI in L2 learning and how the CoI framework helps inform AI-supported language learning. Next, the methodology is presented, followed by the reporting of the results. Then, the research results are discussed. Contributions and implications for theories and practice are highlighted. Several limitations and directions for future research are given.

2. Theoretical framework

2.1. Reviewing the application of AI in L2 learning

Even though giving a widely accepted definition to AI is challenging, for the purpose of this study, AI is broadly defined as the capability of a computer or computer-controlled robot to reason, to learn, and to express themselves in a human-like manner (Berendt et al., 2020;

Williamson & Eynon, 2020). The AI for language learning usually involves such technologies as natural language processing, voice recognition, and speech synthesis (Engwall & Lopes, 2020; Natale & Cooke, 2021; Tai & Chen, 2020). AI technologies have demonstrated great potentials for personalized language learning and precision education (Lin & Mubarok, 2021; Moussalli & Cardoso, 2020). However, the research on AI-supported language learning has so far remained nascent (Chen et al., 2021).

Different types of AI applications have been implemented in L2 learning, for instance, voice-driven and text-driven AI applications (Bibauw et al., 2019). Voice-driven AI applications such as Amazon's Alexa, social robots, and AI chatbots are capable of automatic speech recognition and oral communication, while text-driven AI applications can recognize errors in writing automatically and provide instantaneous feedback (Bibauw et al., 2019). Few researchers have also started to experiment with multimodal AI technologies by connecting AI with other types of technologies. For instance, Divekar* et al. (2021) combined AI and extended reality (XR) in a pilot study involving 10 university students to create an intelligent and immersive language learning environment for learners of Chinese-as-a-foreign language (CFL). In their study, the AI agents engaged with learners in multimodal conversations by hearing, seeing, and understanding them in a virtual world powered by XR, significantly increasing their CFL vocabulary and improving listening and speaking skills.

In general, AI has been applied in L2 learning mainly for two purposes: language acquisition and affect in language learning. Regarding language acquisition, AI-supported language learning has been found to be useful for L2 listening comprehension, speaking, and reading. For instance, Dizon (2020) investigated the influence of Amazon's Alexa on promoting learners' EFL development through a quasi-experimental study. By comparing the performance of an experimental group with that of a control group involving 28 EFL university learners in total, the research findings suggested that the experimental group made more gains in L2 speaking proficiency. As for affective benefits, AI applications have been shown to consistently produce positive effects on learners' affect as they provide a less threatening environment and authentic experience for meaningful communications where learners feel motivated and less anxious. For example, Tai and Chen (2020) investigated the effect of Google Assistant on developing adolescent EFL learners' willingness to communicate (WTC) and their views of Google Assistant for English learning. Through the use of questionnaires, the results indicated that Google Assistant enhanced the learners' WTC, increased communicative confidence, and alleviated speaking anxiety.

Overall, in AI-supported language learning, human-AI interactions help reduce affective barriers, offers immediate feedback for timely self-correct, and provides ample opportunities for repeated practice, eventually enhancing learners' language comprehension and acquisition (Moussalli & Cardoso, 2020; Tai & Chen, 2020; Underwood, 2017). Nevertheless, some research findings were derived from short studies (e.g., 10-15 minutes for two days; Engwall & Lopes, 2020) involving a small number of participants (e.g, 11 participants; Moussalli & Cardoso, 2020), thus, making them subject to novelty effects and difficult to be generalized (van den Berghe et al., 2019). Even though AI applications have been seen as carrying great benefits for improving L2 learning, learners' actual learning outcomes were seldom examined in many prior studies (Zhang & Zou, 2020). Moreover, little is known regarding how L2 learners are cognitively and socially engaged in AI-supported language learning and how L2 learners' experience with AI is related to their L2 enjoyment and actual learning outcomes.

2.2. Community of inquiry in AI-supported language learning

CoI is a widely used theoretical framework for quality online education and focuses on meaningful learning through interacting with others in a community (Garrison & Arbaugh, 2007; Lomicka, 2020). CoI consists of three components: social presence, cognitive presence, and teaching presence, which together forge effective educational experience (Garrison et al., 2010; Garrison & Arbaugh, 2007). The three different presences have been positively related to student learning processes and outcomes. Social presence is important to generate positive learning experiences and to enhance learning engagement (Smidt et al., 2021; Wu et al., 2017). Cognitive presence can facilitate exploration, mastering learning materials, and problem-solving (Garrison & Arbaugh, 2007). Teaching presence enables students to achieve meaningful and worthwhile learning performance (Garrison, 2017; Wu et al., 2017).

The interaction between humans and AI, particularly, the voice-driven AI, is characterized by verbal communications. This is sharply different from the interaction between humans and conventional technologies, which is often in non-verbal forms such as textual communications and mouse clicking. The AI as a humanized agent coupled with verbal communications makes CoI a potentially appropriate framework that facilitates our understanding of AI-supported language learning.

CoI has been applied in many digital language learning settings. For instance, Lomicka (2020) successfully used the CoI framework to create

and sustain virtual language learning communities in responding to the school closedown due to the Covid-19. Smidt et al. (2021) explored the Global Englishes users' experience of social, cognitive, and teaching presences and knowledge construction based on the three presences in asynchronous discussion boards.

In the present study, the AI coach acted as a virtual intelligent teacher. In line with CoI, we argued that learners and the AI coach formed a learning community in which the learners perceived social, cognitive, and teaching presences via interacting with the AI coach. According to Anderson et al. (2001), Garrison and Arbaugh (2007), and Smidt et al. (2021) while taking into consideration the context of this study, social presence is defined as the learners' abilities to identify with the human-AI learning community and communicate purposefully in a stress-free environment. In human-AI interactions, learners first get themselves acquainted with the AI coach, then understand its expectations, and gradually feel comfortable and secure in communicating with it. Through sustained interactions, the learners can develop social bonds with the AI coach.

Cognitive presence refers to the learners' abilities to construct L2 knowledge and build up L2 skills through discourse with the AI coach. During human-AI interactions, learners identify problems of pronunciation and utterances with the help of AI, address the problems, and eventually apply their L2 knowledge in real-life settings. However, cognitive presence is considered the most challenging to develop in digital learning environments (Garrison & Arbaugh, 2007).

Teaching presence refers to the instructional design, facilitating discourse, and direct instruction provided by the AI coach, which can help the learners achieve worthwhile learning outcomes both personally and educationally. The AI coach designed for language learning can automatically diagnose problems of L2 learning and provide feedback on their performance. According to Garrison and Arbaugh (2007) and Smidt et al. (2021), teaching presence significantly determines student satisfaction and perceived learning. However, its effect on actual learning performance is lesser known.

In the present study, we aimed to apply the CoI framework to examine how the perceived social, cognitive, and teaching presences of AI predicted learners' L2 enjoyment and learning outcomes. Furthermore, we sought to identify how learners with different perceived AI presences would differ in AI usage, L2 enjoyment, and learning outcomes. Throughout this study, we also took into account learners' demographic information (i.e., gender) and affection for the AI coach's look, which have not often been examined but may affect AI-supported language learning (Randall, 2019).

3. Methodology

3.1. Participants

The participants were recruited from a public primary school (the target school hereafter) in China. There were initially 581 students enrolled for this study involving 12 classes. They were all Grade one students who just began their formal schooling at the time of this study. During the whole study, the students could discontinue using the AI coach at any time at their discretion. After two and a half months, we obtained valid responses from 327 students, who consistently used the AI coach for English learning, filled out the survey with complete personal information and no items left empty, and attended the midterm English test altogether. Among them were 171 female and156 male students aged between 6 and 8.

3.2. Research context

The AI coach was introduced to the target school for two purposes: (a) to provide the students with a tool that supported them to learn authentic English anytime and anywhere; (b) to cope with the shortage of native English-speaking teachers who could not work on-site due to the pandemic. The school teachers used the AI coach as an extension of the face-to-face language classroom for training English pronunciation so as to facilitate the development of listening, speaking, and vocabulary. This is because mastering correct pronunciation is critical for L2 communication as mispronunciation can impede the intelligibility and comprehensibility of one's speech, thereby causing communication breakdowns in human-human interactions (Moussalli & Cardoso, 2020). For first-grade primary school students who are right in the middle of a critical period where acquisition of 'native-like' phonetic features can happen, receiving authentic and intensified English pronunciation training was highly important.

Normally, the students used the AI coach both as homework and as part of L2 learning in class. The average usage time was at least 15 minutes every day. The workflow of the AI coach could unfold in the following manner. The AI coach read a sentence from a textbook stored in the system for the students to follow and recorded the students' utterances while the students were reading the sentence. The AI coach then identified problems, if there were any, in the students' pronunciation, gave scores for the utterances, and provided standard American- or British-accented utterances for the students to repeat. The higher the score, the closer the students' pronunciation to the standard American or British pronunciation. The students could repeat the utterances

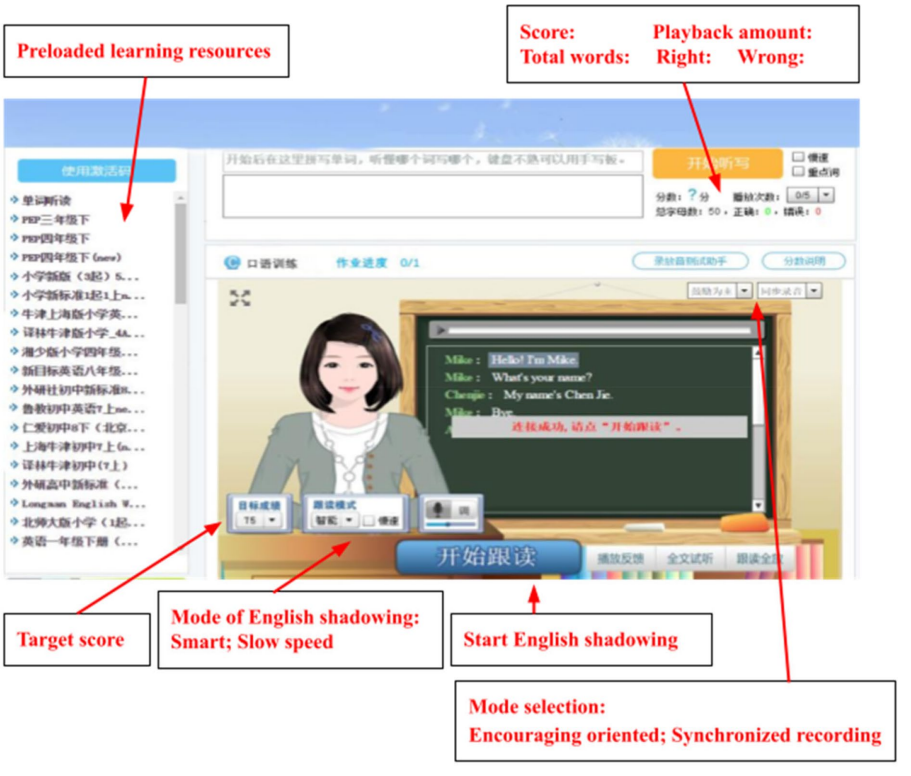

Figure 1. A screenshot of the AI coach.*Note.* Copyright belongs to kouyu100.com.

generated by the AI coach as many times as possible. This teaching technique is called English shadowing, which can help students speak English out loud right from the beginning and is suitable for beginner learners as the students in this study to improve their English speaking (Hamada, 2016). The AI coach generated two datasets for English shadowing based on a series of algorithms: total frequencies and averaged scores of English shadowing for each student.

To facilitate the understanding of the workflow of the AI coach, an example is given here with a screenshot of the AI coach (see Figure 1). In an English shadowing practice, the AI coach first speaks a sentence 'Hello, I am Mike'. Then the students follow the AI coach and read the sentence. If the students read it wrongly, for instance, 'Hello, I am me', the AI coach will give them 63 points, for a total score of 100. Meanwhile, the AI coach will make a comment such as 'You don't speak the word correctly. It should be Mike. But what you said is me… (The AI coach will play the recording of the students' previous statement for the students to recognize their errors) … Please say it again.' If the students can pronounce correctly this time, the AI coach will give encouraging statements such as 'Well done' and 'I am so proud of you'.

The AI coach also posted the students' scores to a bulletin board which automatically ranked all students' performance. The students could visit one another's virtual space, check their performance, and send flowers to one another. Through this social interaction and comparison, the students could develop an awareness of the differences between their performance and others and subsequently increase their motivation for further improvement.

During the learner-AI interaction, the AI coach would produce various encouraging or supportive statements based on automatic analyses of the students' utterances, such as 'I am so proud of you' and 'You have a lot of room for improvement', and reminder comments, such as 'Your voice is too low' and 'Your utterance is not fluent'. Through the social interactions with the AI coach, improvement of L2 knowledge, and the pedagogical design of the AI coach, the students might perceive social, cognitive, and teaching presences in the AI-supported language learning.

3.3. Data collection

To obtain a more objective and comprehensive understanding of how AI could support L2 learning, we collected three data sources: actual learning outcomes, AI usage data, and attitudinal data.

3.3.1. Actual learning outcomes

We used midterm English test scores to represent the students' learning outcomes. Although course test scores are a narrow measure of learning, they remain an essential indicator of language performance for schools and students (Shea et al., 2012). After two and a half months' usage, the students attended the schoolwide midterm English test which assessed their listening comprehension and vocabulary pronunciation and spelling and was scored on 0–100 marks. However, the students who did not attend the test, regardless of having valid AI usage data and attitudinal data, were not included in the data analysis.

3.3.2. AI usage data

The AI usage data were extracted from the AI coach, which contained frequencies of English shadowing and scores of English shadowing given by the AI coach. However, the students who had missing values in the two types of data, albeit with valid midterm English test scores and attitudinal data, were not included in the data analyses.

3.3.3. Attitudinal data

To measure students' perceived AI presences, L2 enjoyment, and affection for AI's appearance, a survey was created based on prior studies.

Besides the items collecting demographic information, there were 19 items in the survey (see Appendix A), which were rated on a five-point Likert scale, with 1 indicating strongly disagree/dislike and 5 strongly agree/like.

Social presence (three items), cognitive presence (three items), and teaching presence (three items) were adapted from Law et al. (2019) and had Cronbach's alpha values of 0.76, 0.90, and 0.70, respectively, in this study. The seven items measuring L2 enjoyment were adapted from J. S. Lee and Lee (2021). Its Cronbach's alpha value was 0.79. Affection for AI's appearance was developed based on Chen et al. (2020) and Randall (2019) and contained three items, with a Cronbach's alpha value of 0.83. A pilot study was conducted with three students to test their understanding of each item, which was refined subsequently to deliver clearer information. The survey was administered to the students' parents whose informed consent was obtained and who assisted the students to fill out the survey. However, the students who had missing values in their responses to the survey were kept out of the study.

Overall, after combining the three data sources, we only kept the data of the students who satisfied the three criteria at the same time:

- Completing the midterm English tests;
- Having valid AI usage data;
- Having valid survey responses.

We then integrated the three data sources into one single dataset, which was used for subsequent data analyses.

3.4. Data analysis

The AI coach aimed to complement the students' L2 learning, rather than dominating it. Thus, to examine the usefulness of the AI coach for L2 learning, the perceived social, cognitive, and teaching presences in AI-supported L2 learning and the AI usage were investigated to see how these factors could predict students' general L2 learning outcomes and L2 enjoyment. For this purpose, hierarchical regression analyses were performed. In addition, to examine how the students with different perceived AI presences would differ in their AI usage, L2 enjoyment, and learning outcomes with the objective of getting an insight into the impact of the human-AI interaction on L2 learning, cluster analyses were conducted first to group the students into different clusters based on their perceived AI presences. Two-step cluster analysis was used as gender information was also included.

4. Results

In this section, we responded to each research question one by one. Two hierarchical regression analyses were first conducted to investigate how the perceived AI presences, together with other factors, predict L2 enjoyment and learning outcomes. Then, cluster analyses were performed to group the students in different clusters based on their perceived AI presences. Subsequently, possible differences between different clusters in AI usage, L2 enjoyment, and learning outcomes were examined.

Assumptions for the hierarchical regression analyses were first examined. Durbin-Watson statistics for L2 enjoyment and learning outcomes were 2.02 and 2.00, respectively, implying no autocorrelation in the data. Variance inflation factor (VIF) values for all the predictors examined in this study ranged from 1.00 to 1.43, which were substantially below 10, suggesting no signs of multicollinearity in the hierarchical regression analyses. In addition, correlation analyses were conducted among the factors examined in this study, and the results were illustrated in Figure 2. As shown in Figure 2, there were significant correlations among social, cognitive, and teaching presences, ranging from 0.52, $p < 0.000$ to 0.76, $p < 0.000$. L2 enjoyment was significantly correlated with learning outcomes, affection for AI's appearance, social, cognitive, and teaching presences, with correlation coefficients varying from 0.19, $p < 0.01$ to

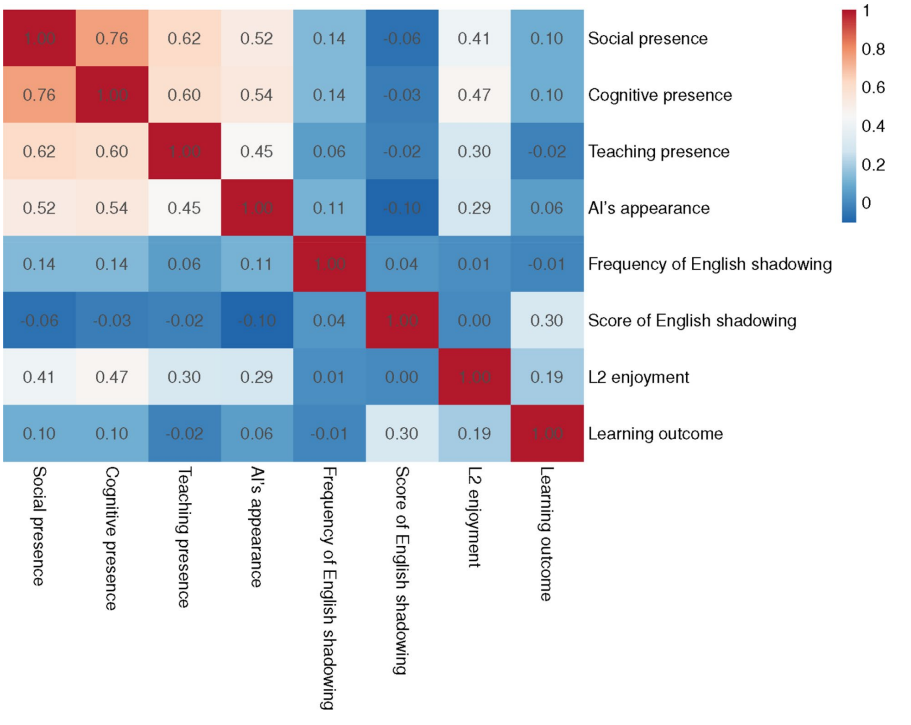

Figure 2. Heatmap of the correlations among different factors.

Table 1. Hierarchical regression analysis with L2 enjoyment as the dependent variable.[a]

	Model 1			Model 2			Model 3		
Variables	*B*	*SE*	*β*	*B*	*SE*	*β*	*B*	*SE*	*β*
Constant	4.253	0.348		3.343	0.374		2.494	0.359	
Frequency of English shadowing	0.000	0.001	0.008	0.000	0.000	−0.025	−0.001	0.000	−0.065
Score of English shadowing	0.000	0.004	0.001	0.002	0.004	0.032	0.002	0.003	0.027
AI's appearance				0.197	0.036	0.293***	0.021	0.041	0.031
Social presence							0.091	0.057	0.129
Cognitive presence							0.292	0.062	0.375***
Teaching presence							−0.007	0.044	−0.010
R^2		0.000			0.084			0.235	
Adjusted R^2		−0.006			0.075			0.221	
F		0.011			9.855***			16.397***	
ΔR^2		0.000			0.084			0.151	
ΔF		0.011			29.542***			21.099***	

Note: $^{**}p < .01$; $^{***}p < .001$; [a]=Three decimal places were kept in this table for displaying parameters with small values.

0.47, $p < 0.000$. Learning outcomes was significantly correlated with L2 enjoyment ($r = 0.19$, $p < 0.01$) and scores of English shadowing ($r = 0.30$, $p < 0.000$).

4.1. RQ1: How do learners' perceived AI presences, affection for AI's appearance, and AI usage data predict L2 enjoyment?

In Model 1, the AI usage data (frequency and score of English shadowing) were entered with L2 enjoyment as the dependent variable (see Table 1). The model was not significant. Neither frequency of English shadowing nor score of English shadowing predicted L2 enjoyment.

The students' affection for AI's appearance was included in Model 2, which was statistically significant (adjusted $R^2 = 0.075$, $F = 9.855$, $p < 0.000$). The F value increased significantly ($\Delta F = 29.542$, $p < 0.000$) because of the addition of affection for AI's appearance. Affection for AI's appearance ($\beta = 0.293$, $p < 0.000$) positively predicted L2 enjoyment after controlling the frequency and score of English shadowing.

Model 3 included social, cognitive, and teaching presences and was found to be statistically significant, adjusted $R^2 = 0.221$, $F = 16.397$, $p < 0.000$. The inclusion of the three presences resulted in a significant increase in the F value of 21.099, $p < 0.000$. Among the three presences, cognitive presence ($\beta = 0.375$, $p < 0.000$) was found to be a significant predictor of L2 enjoyment.

4.2. RQ2: How do learners' perceived AI presences, L2 enjoyment, affection for AI's appearance, and AI usage data predict actual learning outcomes?

Model 1 included frequency and score of English shadowing with learning outcomes as the dependent variable (see Table 2). The model was

Table 2. Hierarchical regression analysis with learning outcomes as the dependent variable.[a]

	Model 1			Model 2			Model 3			Model 4		
Variables	*B*	*SE*	*β*	*B*	*SE*	*β*	*B*	*SE*	*β*	*B*	*SE*	*β*
Constant	64.511	5.176		60.162	5.773		57.108	5.980		50.670	6.346	
Frequency of English shadowing	−0.004	0.007	−0.027	−0.005	0.008	−0.037	−0.008	0.007	−0.055	−0.006	0.007	−0.044
Score of English shadowing	0.334	0.058	0.303***	0.344	0.058	0.312***	0.350	0.058	0.318***	0.345	0.057	0.314***
AI's appearance				0.942	0.560	0.090	0.522	0.676	0.050	0.469	0.670	0.045
Social presence							1.585	0.945	0.143	1.349	0.939	0.122
Cognitive presence							1.139	1.028	0.094	0.385	1.052	0.032
Teaching presence							−1.924	0.731	−0.183**	−1.906	0.724	−0.181**
L2 enjoyment										2.582	0.921	0.166**
R^2		0.092			0.100			0.128			0.149	
Adjusted R^2		0.086			0.091			0.111			0.130	
F		16.388***			11.930***			7.803***			7.955***	
ΔR^2		0.092			0.008			0.028			0.021	
ΔF		16.388***			2.829			3.409*			7.862**	

Note: $^{*}p < .05$; $^{***}p < .001$; [a]=Three decimal places were kept in this table for displaying parameters with small values.

Table 3. Full information of the clusters generated by the two-step cluster analysis.

Number of Clusters	Schwarz's Bayesian Criterion (BIC)	BIC Change[a]	Ratio of BIC Changes[b]	Ratio of Distance Measures[c]
1	1409.375			
2	1113.296	−296.079	1.000	1.066
3	838.714	−274.582	.927	3.317
4	792.323	−46.390	.157	1.540
5	780.454	−11.869	.040	1.029
6	770.399	−10.055	.034	1.632
7	784.429	14.030	-.047	1.045
8	800.082	15.654	-.053	1.312
9	824.402	24.320	-.082	1.229
10	853.895	29.493	-.100	1.026
11	883.953	30.057	-.102	1.149
12	916.876	32.923	-.111	1.065
13	950.966	34.090	-.115	1.026
14	985.513	34.547	-.117	1.342
15	1024.533	39.020	-.132	1.087

Note.
[a]The changes are from the previous number of clusters in the table.
[b]The ratios of changes are relative to the change for the two-cluster solution.
[c]The ratios of distance measures are based on the current number of clusters against the previous number of clusters.

statistically significant (adjusted $R^2 = 0.086$, $F = 16.388$, $p < 0.000$). Scores of English shadowing ($\beta = 0.303$, $p < 0.000$) were found to be a significant predictor of learning outcomes.

In Model 2, affection for AI's appearance was added. Model 2 was statistically significant (adjusted $R^2 = 0.091$, $F = 11.930$, $p < 0.000$). But, the addition of this factor did not significantly increase the F value ($\Delta F = 2.829$, $p < 0.094$).

Model 3 included social, cognitive, and teaching presences and was found to be statistically significant (adjusted $R^2 = 0.111$, $F = 7.803$, $p < 0.000$). There was a significant increase in the F value of 3.409, $p < 0.050$. However, among the three presences, teaching presence ($\beta = -0.183$, $p < 0.010$) was found to negatively predict learning outcomes after controlling other factors.

In Model 4, L2 enjoyment was added and the model was statistically significant (adjusted $R^2 = 0.130$, $F = 7.955$, $p < 0.000$). The inclusion of L2 enjoyment led to a significant increase in the F value of 7.862, $p < 0.010$. L2 enjoyment ($\beta = 0.166$, $p < 0.010$) was found to positively predict learning outcomes with the AI usage data and the three presences controlled for.

4.3. RQ3: How are learners clustered based on their perceived AI presences, gender, and affection for AI's appearance, and how do they differ regarding AI usage data, L2 enjoyment, and actual learning outcomes?

The social, cognitive, and teaching presences were included in the two-step cluster analysis together with the students' gender information and their affections for the AI's appearance. The Bayesian Information Criterion

Table 4. Cluster analysis outcomes for the three clusters of students.

Variables[a]	Cluster 1	Cluster 2	Cluster 3
Sizes	*N* = 127 (38.80%)	*N* = 72 (22.00%)	*N* = 128 (39.10%)
Cognitive presence	4.79 (0.40)[b]	2.99 (0.72)	4.69 (0.50)
Gender distribution (*N*)	Female (127)	Male (28); Female (44)	Male (128)
Social presence	4.43 (0.66)	2.72 (0.75)	4.32 (0.73)
AI's appearance	4.28 (0.85)	2.80 (0.70)	4.10 (0.96)
Teaching presence	3.83 (0.93)	2.50 (0.70)	3.85 (0.92)

Note.
[a]=The variables are listed following the decreasing levels of importance in forming clusters.
[b]=Mean (SD).

(BIC) was employed to determine the reasonable numbers of clusters, with smaller BIC values indicating better models (Vrieze, 2012). To obtain an optimal solution, it is necessary to consider not only the BIC values but also the changes in BIC values and the distance measures between different cluster solutions. As the BIC values decrease, the number of clusters increases, thereby complicating the cluster solutions. Following the principle of parsimony, it is critical to maintaining a balance between the BIC values and the complexity of cluster solutions (Benassi et al., 2020). Table 3 shows that the three-cluster classification was the optimal solution, with relatively lower BIC values (838.714), a larger ratio of BIC changes (0.927), and the biggest ratio of distance measures (3.317).

Table 4 presents the composition of each cluster of students. Cluster 1 comprised 127 female students, making up 38.80% of the total participants. Cluster 2 was composed of 72 participants (22%), involving 28 male students and 44 female students. Cluster 3 comprised 128 female students (39.10%). Specifically, Cluster 1 students showed the greatest affection for the AI's appearance and experienced the highest cognitive and social presences when interacting with AI. Contrastingly, Cluster 2 students demonstrated the least affection for the AI's appearance and perceived the lowest cognitive, social, and teaching presences.

As gender accounted for an important factor in creating the three clusters, particularly for Cluster 1, a subsequent independent *t*-test was performed to explore potential differences in the constructs examined in this study. As revealed in Table 5, female and male students did not have significant differences in AI usage, perceived AI presences, affection for AI's appearance, and L2 enjoyment. However, they was a significant difference in learning outcomes with female students ($M = 94.96$, $SD = 7.65$) being statistically higher than males ($M = 90.76$, $SD = 13.32$), $t\ (325) = -3.53$, $p = 0.000$.

To further investigate how the three clusters of students would differ in learning outcomes, enjoyment in learning English, and AI usage, a one-way multivariate analysis of variance (MANOVA) was conducted subsequently. There was a statistically significant difference among the three clusters of students, $F\ (8, 642) = 10.69$, $p < 0.000$; Wilk's $\Lambda = 0.75$, partial $\eta^2 = 0.12$.

Table 5. Mean comparisons of different variables between different genders.

	Gender	*N*	*M*	*SD*	
Frequency of English shadowing	Male	156	148.82	78.09	t (325) = 0.50, p = 0.62
	Female	171	144.58	76.55	
Score of English shadowing	Male	156	87.40	5.44	t (325) = 0.79, p = 0.43
	Female	171	86.54	12.71	
Learning outcome***	Male	156	90.76	13.32	t (325) = −3.53, p = 0.000
	Female	171	94.96	7.65	
Social presence	Male	156	4.05	0.92	t (325) = 0.76, p = 0.45
	Female	171	3.97	1.04	
Cognitive presence	Male	156	4.38	0.85	t (325) = 0.41, p = 0.68
	Female	171	4.34	0.94	
Teaching presence	Male	156	3.63	0.99	t (325) = 1.41, p = 0.16
	Female	171	3.47	1.08	
AI's appearance	Male	156	3.87	1.02	t (325) = −0.15, p = 0.88
	Female	171	3.89	1.06	
L2 enjoyment	Male	156	4.20	0.74	t (325) = −1.64, p = 0.10
	Female	171	4.33	0.66	

Note: ***p < .001.

Table 6. MANOVA outcomes for the three clusters of students.

Variables	Clusters	Mean (SD)	F	*df*	partial η^2	Pairwise comparisons[a]	95% CI for difference
Learning outcome	C1	95.35 (6.38)	5.14**	(2, 324)	0.03	C1 > C2*	[0.29, 7.78]
	C2	91.32 (14.35)				C1 > C3**	[0.67, 7.04]
	C3	91.50 (11.89)					
L2 enjoyment	C1	4.48 (0.54)	33.91***	(2, 324)	0.17	C1 > C2***	[0.53, 0.97]
	C2	3.73 (0.84)				C3 > C2***	[0.41, 0.85]
	C3	4.36 (0.60)					
Frequency of English shadowing	C1	155.61 (79.13)	3.30*	(2, 324)	0.02	C1 > C2*	[2.08, 55.33]
	C2	126.90 (68.81)					
	C3	148.75 (78.32)					
Score of English shadowing	C1	86.16 (14.41)	0.72	(2, 324)	0.004		
	C2	87.10 (6.47)					
	C3	87.65 (4.73)					

Note: [a]= Only statistically significant comparisons were reported; *p < .05; **p < .01; ***p < .001.

A series of post-hoc analyses were performed to investigate the mean difference comparisons across the three clusters. Table 6 shows that the three clusters of students significantly differed on learning outcomes (F (2, 324) = 5.14, p < 0.006; partial η^2 = 0.03), L2 enjoyment (F (2, 324) = 33.91, p < 0.000; partial η^2 = 0.17), and frequency of English shadowing (F (2, 324) = 3.30, p < 0.04; partial η^2 = 0.02). However, there was no significant difference in the score of English shadowing (F (2, 324) = 0.72, p < 0.49; partial η^2 = 0.004).

As indicated in Table 6, Cluster 1 (M = 95.35, SD = 6.38) obtained significantly higher learning outcomes than Cluster 2 (M = 91.32, SD = 14.35; p = 0.03) and Cluster 3 (M = 91.50, SD = 11.89; p = 0.01). As for L2 enjoyment, both Cluster 1 (M = 4.48, SD = 0.54; p = 0.000) and Cluster 3 (M = 4.36, SD = 0.60; p = 0.0000) were significantly greater than Cluster 2 (M = 3.73, SD = 0.84). In addition, Cluster 1 (M = 155.61, SD = 79.13) had significantly higher frequencies of English shadowing than Cluster 2 (M = 126.90, SD = 68.81; p = 0.03).

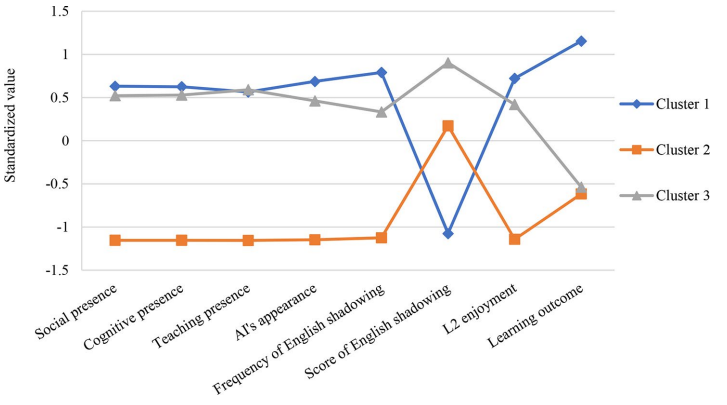

Figure 3. Line graph comparing clusters of students on the variables in the format of standardized values.*Note.* Cluster 1 (N = 127); Cluster 2 (N = 72); Cluster 3 (N = 128).

Subsequently, Clusters 1-3 were projected to Figure 3 in which the three presences, AI usage data, L2 enjoyment, learning outcomes, and affection for AI's appearance were standardized. As illustrated in Figure 3, specifically, Cluster 1 perceived the highest cognitive and social presences. They were most fond of AI's appearance. Their interactions with AI were the most frequent among the three clusters. And they perceived the most enjoyment in learning English and obtained the highest score in the midterm English test. By comparison, Cluster 2 experienced the lowest cognitive, social, and teaching presences. They were least fond of AI's appearance. They did not interact with AI frequently. They did not seem to enjoy studying English. Their scores of midterm English tests were almost the lowest.

5. Discussion

This study examined the implementation of an AI coach, which was conceptualized as a humanized agent, for L2 learning in a primary school. Drawing on multiple data sources and following the CoI framework, the current study sought to investigate the influence of students' perceived AI presences on L2 learning. In what follows, we answered each research question by focusing on discussing important research findings.

5.1. RQ1: How do learners' perceived AI presences, affection for AI's appearance, and AI usage data predict L2 enjoyment?

Cognitive presence was found to be a significant predictor of L2 enjoyment. This may be due to the fact that the students were learning English in a supportive and personalized environment formed by AI. Through communicating with the AI coach, the students got feedback on their language pronunciation and utterances, continued to refine

their speaking, and received verbal encouragement from the AI coach. As a result, the students gradually improved their English proficiency and developed or reinforced a positive feeling towards English. This finding is also in line with J. S. Lee and Lee (2021) who indicated that sustained interactions with technologies for L2 learning outside the classroom, such as talking to others in English and listening to English songs through digital tools, were significantly correlated with learners' L2 enjoyment. However, J. S. Lee and Lee (2021) emphasized the role of out-of-class settings, which seemed to be less stressful than in-class settings for digital L2 learning. The present study indicated that the AI coach produced dual effects on L2 learning: generating a safe learning environment in-and-out-of-class while providing personalized guidance.

Social presence did not significantly predict L2 enjoyment in AI-supported language learning. This may be because that even though the AI coach was expected to support humanized interactions, its capacity to do so may be limited as other types of digital technologies (Yung, 2015; Yung & Chiu, 2020). Compared with human teachers, the learning community supported by the AI coach may not yet be able to provide sufficient sensory stimulation and direct human interactions, thereby falling short of increasing the students' L2 enjoyment (Yung & Chiu, 2020).

Affection for AI's appearance was found to positively predict L2 enjoyment. Literature (e.g., Lee et al., 2011; Wu et al., 2015) has already indicated that AI should have pleasant appearances, which may, in one way or another, affect student learning. In particular, the appearances of AI tend to have a higher influence on young children than other age groups (Wu et al., 2015). Unpleasant appearances may cause anxiety and displeasure in L2 learning, while likable looks can reduce such negative affective experience (Randall, 2019; Wu et al., 2015). Although the AI coach in this study was in the form of a teacher and thus, possibly involved a certain degree of implicit authority, its female appearance with the smiley face may have decreased the pressure and anxious feelings in L2 learning and increased learning enjoyment. This research finding also provides new evidence about how AI's appearances may influence L2 acquisition, a gap as noted in Randall (2019).

AI usage did not predict L2 enjoyment. This result may be caused by possibly unpleasant experiences in AI-supported language learning. As with most other voice-driven AI tools (Chen et al., 2020; Randall, 2019), the speech recognition module of the AI coach was not perfect. When there were background noises, the students needed to repeat themselves several times before the AI coach could understand them. There were also times when the students might have been frustrated by the scores given by the AI coach, which did not satisfy their expectations.

5.2. RQ2: How do learners' perceived AI presences, L2 enjoyment, affection for AI's appearance, and AI usage data predict actual learning outcomes?

Among the three presences, teaching presence was found to negatively predicate learning outcomes. This finding is at odds with many previous studies (e.g., Garrison & Arbaugh, 2007; Ke, 2010) that argued for the benefits associated with teaching presence in technology-supported learning. As suggested by Randall (2019), AI often exhibits one of the following roles in language learning: (a) teacher, (b) assistant, and (c) peer/tutor. On a continuum from authority to connectedness, the teacher role is high in authority while the peer/tutor role is high in connectedness. High authority is often perceived as intimidating (Park et al., 2011). In the current study, the AI coach was in the form of a female teacher, even though with a pleasant look. Moreover, it was also emphasized by the target school to play a teacher's role in improving the students' English listening and speaking through the approach of English shadowing. This might unintentionally strengthen the teaching presence associated with the AI coach. When learners perceived a high teacher role in AI, they might find interactions with AI stressful, eventually leading to a negative influence on L2 learning.

L2 enjoyment was found to be a positive predictor of learning outcomes. This is in line with Dewaele et al. (2018) and Jiang and Dewaele (2019) who postulated that positive emotions such as L2 enjoyment could enable students to become more perceptive of language input and increase their intellectual focus, thereby facilitating the acquisition of the target language. Their studies were done in conventional classrooms with real teachers. However, based on multiple data sources, this study provided evidence-based support to their postulation in an AI-supported L2 learning setting.

As an indicator of the quality of AI usage, the English shadowing score was found to positively predict learning outcomes. This finding corroborated Dizon (2017) and Tai and Chen (2020) who found that the pronunciation feedback provided by AI applications (e.g., Amazon's Alexa) enabled learners to acquire more language knowledge. However, these studies were either very short with small samples or solely relied on self-reported data. In this regard, the present study contributed more robust evidence for the relationship between AI usage and learning outcomes. The AI applications could provide undivided attention and individualized support to every learner thus enhance the quality of human-AI interaction, eventually leading to strengthened effectiveness and increased performance in language learning.

However, in the two hierarchical regression analyses, social presence did not predict L2 enjoyment and learning outcomes. This may have been caused by the characteristic of social presence. According to Annand (2011) and Garrison and Arbaugh (2007), social presence was not a fixed construct. Instead, it may decrease over time as the learners' use of the AI coach become more focused on academic purposes rather than on social bonds with the AI coach. Hence, the predictive power of social presence for L2 enjoyment and learning outcomes tends to be volatile.

5.3. RQ3: How are learners clustered based on their perceived AI presences, gender, and affection for AI's appearance, and how do they differ regarding AI usage data, L2 enjoyment, and actual learning outcomes?

The students were formed into three clusters based on their perceived AI presences, gender, and affections for AI's appearance, with Cluster 1 (with the highest social and cognitive presences and affections for AI's appearance) being in stark contrast with Cluster 2. Based on the cluster analysis and subsequent MANOVA results related to Clusters 1 and 2, in particular, it can be seen that students' affection for the AI's appearance and social and cognitive presences were largely in line with their AI use, enjoyment in learning English, and learning outcomes.

These findings from a different perspective indicate that the CoI framework may still hold in AI-supported learning contexts, albeit with limited predicting powers. The interaction between the students and the AI coach could form a learning community that might increase the students' satisfaction with the AI coach as well as the learning process, promote active knowledge processing, and lead to improved performance (Garrison, 2017; Garrison & Arbaugh, 2007; Wu et al., 2017).

As for gender, female students constituted Cluster 1, which was the best performing cluster. Although there is a gender stereotype in the use of technologies that male students have stronger digital skills and more positive attitudes than female students (Li & Kirkup, 2007), female students achieved significantly higher learning outcomes than male counterparts in this study. This could be because female students tend to have more positive attitudes towards foreign languages, find more enjoyment, and feel more creative than males in L2 learning, eventually developing higher competences in foreign languages (Dewaele et al., 2018).

5.4. Implications

The present study has implications for future research and practice on AI for education in the following ways. First, as AI's appearance is

important for L2 learning, AI developers are suggested to make the appearance of AI as appealing and engaging as possible. They are suggested to find out which forms of AI's appearance (e.g., humanlike or cartoon-like) work for learners of different ages, genders, and personalities (Randall, 2019). Other factors can be taken into consideration, such as colors, sizes, and genders of AI characters (Köse et al., 2015; Randall, 2019).

Second, contrary to intuition, teaching presence associated with AI can be counterproductive to L2 learning. High teaching presence may be intimidating to learners. Thus, AI designers are suggested to take into account this issue when developing the AI characters and modifying learner-AI interactions. This is because when learners perceive AI as more of a friend than an instructor, they may find this role greatly desirable and tend to be more connected to it in L2 learning (Randall, 2019; Young et al., 2010).

And third, long-term studies are crucial for the research on novel technologies such as AI (van den Berghe et al., 2019). This is because the novelty effect can affect learners' behaviors, for instance, increased interaction time and higher likeability during learners' initial interactions with AI, and consequently bias research findings (Engwall & Lopes, 2020; Randall, 2019; van den Berghe et al., 2019). However, with increased exposure to new technologies, the novelty effect decreases (Leeuwestein et al., 2020).

6. Conclusion

As AI technologies are increasingly finding their ways into language learning, our understanding of AI-supported language learning is nonetheless limited. Conceptualizing the AI coach as a humanized agent, this study applied the CoI framework to investigate how the students' perceived social, cognitive, and teaching presences when interacting with the AI coach could support L2 learning, while considering factors related to L2 learners' affect and learning gains. This study identified that the students' cognitive presence and affection for the AI's appearance significantly predicted their L2 enjoyment. Teaching presence was negatively associated with L2 learning outcomes while the students' L2 enjoyment was positively related to them. The quality, instead of the quantity, of AI usage seemed to matter in AI-supported language learning. In addition, the students' social and cognitive presences and their affection for the appearance of the AI coach were generally aligned with their usage of AI, L2 enjoyment, and learning outcomes. Overall, the CoI framework offered a new angle for understanding the human-AI interactions in L2 learning.

This present study contributes to the research and practice in AI-supported language learning in the following ways. First, this study contributes to CoI by extending its usage boundary to human-AI interactions. Previous studies on CoI were predominantly conducted in the context of conventional technologies that are incapable of human-like processing and logic. Given the increasing use of AI as a humanized agent in educational settings, the use of CoI in this study affords us a new perspective in the understanding of human-AI interactions in language learning and associated effects on L2 learners' affect and learning outcomes.

Second, this study provides further evidence that substantiates the importance of AI's appearance in affecting L2 learning by showing that learners' affection for AI's appearance positively predicted L2 enjoyment, which in turn positively predicted their actual language gains.

Third, this study contributes new knowledge to our understanding of teaching presence in technology-supported learning. Previous studies (e.g., Garrison, 2017; Ke, 2010) often emphasized the potentials and benefits associated with teaching presence, such as stimulating cognitive and social presence and increasing perceived learning performance. However, in the scenario of AI-supported language learning, it may be a different case. As different roles perceived in AI tend to invoke different affective reactions in learners, high teaching presence may be stressful and cause the AI to be perceived as less friendly to the learners, eventually leading to a negative influence on learner-AI interactions and L2 gains (Randall, 2019).

Nonetheless, the interpretation of the research findings should be taken with caution. First, the participants of this study were all Year 1 primary school students whose social, cognitive, and affective capacities may not be as fully developed as learners in older age groups. Their perceptions of AI and its role in language learning may be different from those in other age groups. Thus, researchers in future studies are suggested to validate them in different population groups and different settings to examine the generalization of the research findings.

Second, the students' learning gains could be affected by external confounding variables such as their instructors who may unintentionally have varying influences on the students' language learning even if following the same curriculum. Future studies are encouraged to investigate the potential roles these variables play in AI-supported language learning by conducting sub-group analyses thereof. Third, even though this study suggested that students' affection for AI's appearance might have an influence on L2 learning, little is still known regarding what kind of AI can bring the maximal benefit to L2 learners. Future studies are suggested to test different looks of AI (e.g., anthropomorphic, cartoon-like, or mechanomorphic) so as to identify an ideal one, if there is any, for different types of L2 learners.

Fourth, consistent with other AI technologies reported in previous studies, the AI coach in this study could not wholly understand learners' utterances in every interaction, possibly due to the low proficiency levels of the early EFL learners or the limited competence of AI algorithms at the current stage (Chen et al., 2020). Future studies are suggested to take into consideration the accuracy rates of AI's feedback so as to gain a more in-depth understanding of AI's effects on language learning. And fifth, given that the students completed the survey under the guidance of their parents, it is unavoidable that their survey responses could be more or less influenced by their parents. Researchers are suggested to control for such influence in future studies, if possible, by taking measures such as assigning one teacher to explain survey items to all participants.

In spite of these limitations, the current study, involving 327 participants and spanning two and a half months instead of a few hours or sessions, contributes to the limited but increasing knowledge of human-AI interaction in language learning, the research field of which has been characterized by short-term case studies, small sample sizes, and less reliable indicators of learning gains (Dizon, 2020; Randall, 2019).

Notes on Contributors

Xinghua Wang is a professor in the Normal College at Qingdao University. He received his PhD in Learning Sciences and Technologies from Nanyang Technological University, Singapore. His research is related to technology-enhanced learning, human-computer interaction, and psychometrics.

Hui Pang is an associate professor in the Normal College at Qingdao University. Her research interest covers computer-assisted language learning, psycholinguistics, and teacher education.

Matthew P. Wallace is an assistant professor of applied linguistics in the Department of English, University of Macau, China. His research up to this point has focused on second language listening comprehension, language assessment fairness, and language learner motivation.

Qiyun Wang is a professor in Learning Sciences and Assessment Academic Group. He is also the Programme Leader of the Master of Education (Learning Sciences and Technologies), and the Chair of the Ethnic Clearance Committee of the Academic Group. He specialises in blended synchronous learning (using real-time video conferencing), technology-supported learning environment, and education design-based research.

Wenli Chen is a professor and the head of Learning Sciences and Assessment Academic Group at the National Institute of Education, Nanyang Technological University Singapore. Her research interests include Computer-Supported Collaborative Learning (CSCL), learning analytics, AI for Education, and mobile learning.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Xinghua Wang http://orcid.org/0000-0002-4246-8600
Hui Pang http://orcid.org/0000-0002-6479-6074
Matthew P. Wallace http://orcid.org/0000-0002-3509-2983
Qiyun Wang http://orcid.org/0000-0001-5891-4997
Wenli Chen http://orcid.org/0000-0003-3196-3400

References

Anderson, T., Rourke, L., Garrison, D. R., & Archer, W. (2001). Assessing teaching presence in a computer conferencing context. *Journal of Asynchronous Learning Networks*, *5*(2), 1–17. https://auspace.athabascau.ca/handle/2149/725

Annand, D. (2011). Social presence within the community of inquiry framework. *International Review of Research in Open and Distributed Learning*, *12*(5), 40–56. https://doi.org/10.19173/irrodl.v12i5.924

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

Bibauw, S., François, T., & Desmet, P. (2019). Discussing with a computer to practice a foreign language: Research synthesis and conceptual framework of dialogue-based CALL. *Computer Assisted Language Learning*, *32*(8), 827–877. https://doi.org/10.1080/09588221.2018.1535508

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

Chen, X., Zou, D., Xie, H. R., & Su, F. (2021). Twenty-five years of computer-assisted language learning: A topic modeling analysis. *Language Learning & Technology*, *25*(3), 151–185. http://hdl.handle.net/10125/73454

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

Dewaele, J.-M., Witney, J., Saito, K., & Dewaele, L. (2018). Foreign language enjoyment and anxiety: The effect of teacher and learner variables. *Language Teaching Research*, *22*(6), 676–697. https://doi.org/10.1177/1362168817692161

Divekar*, R. R., Drozdal*, J., Chabot*, S., Zhou, Y., Su, H., Chen, Y., … Braasch, J. (2021). Foreign language acquisition via artificial intelligence and extended reality: design and evaluation. *Computer Assisted Language Learning*, 1–29. https://doi.org/10.1080/09588221.2021.1879162

Dizon, G. (2017). Using intelligent personal assistants for second language learning: A case study of Alexa. *Tesol Journal*, *8*(4), 811–830. https://doi.org/10.1002/tesj.353

Dizon, G. (2020). Evaluating intelligent personal assistants for L2 listening and speaking development. *Language Learning & Technology*, *24*(1), 16–26. https://doi.org/10.10125/44705

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

Garrison, D. R. (2017). *E-learning in the 21st century: A community of inquiry framework for research and practice*. Routledge.

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*(1–2), 5–9. https://doi.org/10.1016/j.iheduc.2009.10.003

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

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

Jiang, Y., & Dewaele, J.-M. (2019). How unique is the foreign language classroom enjoyment and anxiety of Chinese EFL learners? *System*, *82*, 13–25. https://doi.org/10.1016/j.system.2019.02.017

Ke, F. (2010). Examining online teaching, cognitive, and social presence for adult students. *Computers & Education*, *55*(2), 808–820. https://doi.org/10.1016/j.compedu.2010.03.013

Köse, H., Uluer, P., Akalın, N., Yorgancı, R., Özkul, A., & Ince, G. (2015). The effect of embodiment in sign language tutoring with assistive humanoid robots. *International Journal of Social Robotics*, *7*(4), 537–548. https://doi.org/10.1007/s12369-015-0311-1

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

Lee, J. S., & Lee, K. (2021). The role of informal digital learning of English and L2 motivational self system in foreign language enjoyment. *British Journal of Educational Technology*, *52*(1), 358–373. https://doi.org/10.1111/bjet.12955

Lee, S., Noh, H., Lee, J., Lee, K., Lee, G. G., Sagong, S., & Kim, M. (2011). On the effectiveness of robot-assisted language learning. *ReCALL*, *23*(1), 25–58. https://doi.org/10.1017/S0958344010000273

Leeuwestein, H., Barking, M., Sodacı, H., Oudgenoeg-Paz, O., Verhagen, J., Vogt, P., … de Wit, J. (2020). Teaching Turkish-Dutch kindergartners Dutch vocabulary with a social robot: Does the robot's use of Turkish translations benefit children's Dutch vocabulary learning? *Journal of Computer Assisted Learning*, *37*(3), 603–620. https://doi.org/10.1111/jcal.12510

Lee, C., Yeung, A. S., & Ip, T. (2016). Use of computer technology for English language learning: do learning styles, gender, and age matter? *Computer Assisted Language Learning*, *29*(5), 1035–1051. https://doi.org/10.1080/09588221.2016.1140655

Li, N., & Kirkup, G. (2007). Gender and cultural differences in Internet use: A study of China and the UK. *Computers & Education*, *48*(2), 301–317. https://doi.org/10.1016/j.compedu.2005.01.007

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

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. (2021). 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

Park, S. J., Han, J. H., Kang, B. H., & Shin, K. C. (2011). *Teaching assistant robot, ROBOSEM, in English class and practical issues for its diffusion* [Paper presentation]. Paper Presented at the Advanced Robotics and Its Social Impacts.

Pavelescu, L. M., & Petric, B. (2018). Love and enjoyment in context: Four case studies of adolescent EFL learners. *Studies in Second Language Learning and Teaching*, *8*(1), 73–101. https://doi.org/10.14746/ssllt.2018.8.1.4

Pennington, M. C., & Rogerson-Revell, P. (2019). Using technology for pronunciation teaching, learning, and assessment. In M. C. Pennington & P. Rogerson-Revell (Eds.), *English pronunciation teaching and research* (pp. 235–286). Springer.

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

Shea, P., Hayes, S., Smith, S. U., Vickers, J., Bidjerano, T., Pickett, A., … Jian, S. (2012). Learning presence: Additional research on a new conceptual element within the Community of Inquiry (CoI) framework. *The Internet and Higher Education*, *15*(2), 89–95. https://doi.org/10.1016/j.iheduc.2011.08.002

Smidt, E., Chau, M. H., Rinehimer, E., & Leever, P. (2021). Exploring engagement of users of Global Englishes in a community of inquiry. *System, 98*, 102477. https://doi.org/10.1016/j.system.2021.102477

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

Underwood, J. (2017 *Exploring AI language assistants with primary EFL students* [Paper presentation].Paper Presented at the CALL in a Climate of Change: adapting to Turbulent Global Conditions-Short Papers from EUROCALL,. https://files.eric.ed.gov/fulltext/ED578302.pdf

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

Vrieze, S. I. (2012). Model selection and psychological theory: a discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychological Methods*, *17*(2), 228–243. https://doi.org/10.1037/a0027127

Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology*, *45*(3), 223–235. https://doi.org/10.1080/17439884.2020.1798995

Wu, W.-C., 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. https://www.ds.unipi.gr/et&s/journals/20_2/12.pdf

Wu, W.-C., Wang, R.-J., & Chen, N.-S. (2015). Instructional design using an in-house built teaching assistant robot to enhance elementary school English-as-a-foreign-language learning. *Interactive Learning Environments*, *23*(6), 696–714. https://doi.org/10.1080/10494820.2013.792844

Young, S. S. C., Wang, Y. H., & Jang, J. S. R. (2010). Exploring perceptions of integrating tangible learning companions in learning English conversation. *British Journal of Educational Technology*, *41*(5), E78–E83. https://doi.org/10.1111/j.1467-8535.2009.00989.x

Yung, K. W. H. (2015). Learning English in the shadows: Understanding Chinese learners' experiences of private tutoring. *Tesol Quarterly*, *49*(4), 707–732. https://doi.org/10.1002/tesq.193

Yung, K. W. H., & Chiu, M. M. (2020). Secondary school students' enjoyment of English private tutoring: An L2 motivational self perspective. *Language Teaching Research*, 1–23. https://doi.org/10.1177/1362168820962139

Zhang, R., & Zou, D. (2020). Types, purposes, and effectiveness of state-of-the-art technologies for second and foreign language learning. *Computer Assisted Language Learning*, 1–47. https://doi.org/10.1080/09588221.2020.1744666