

ARTICLE



EFL adolescents' engagement in artificial intelligence and peer interaction

Juhee Lee, *Hongik University*

Abstract

This study investigates how intelligent personal assistants (IPAs, e.g., Google Assistant), AI language learning applications, and peer interactions shape the learning experiences of 201 seventh-grade Korean EFL students. The students participated in a 12-week intervention and were categorized into four groups: Google Assistant, AI app, peer interactions, and textbook-based instruction (control group). Quantitative and qualitative data were collected through questionnaires to examine the cognitive, behavioral, and emotional engagement. The findings highlight the distinct contributions of each intervention to student engagement. The IPAs were praised for creating interactive, humorous, and natural conversations that felt like speaking with real native speakers. This feature significantly enhanced students' emotional engagement compared to the control group. With their structured and gamified learning features, including immediate feedback and repetitive practice, AI apps motivated students but were considered less interactive. Peer interactions provided rich opportunities for authentic communication and socioemotional connections, which AI tools could not replicate. Although the AI tools effectively enhanced student engagement, peer interactions supported developing human connections. These connections manifested through collaborative learning, peer support, and forming a learning community and were instrumental in fostering socioemotional engagement. The findings offer insight for educators aiming to balance technology and interpersonal collaboration in EFL instruction.

Keywords: artificial intelligence; peer interaction; EFL adolescents; engagement

Language(s) Learned in This Study: English

APA Citation: Lee, J. (2025). EFL adolescents' engagement in artificial intelligence and peer interaction. *Language Learning & Technology*, 29(1), 1–21. <https://doi.org/10.64152/10125/73643>

Introduction

The advancement of artificial intelligence (AI) has profoundly transformed second or foreign language (L2) learning, reshaping both pedagogical approaches and learners' experiences (Jeon, 2024; Wu, 2024). Traditionally, human interaction has been considered central to L2 acquisition (Li & Jeong, 2020). However, recent developments in AI-powered tools—such as intelligent personal assistants (IPAs) like Google Assistant and AI-driven language learning applications like Duolingo—offer interactive, flexible, and personalized alternatives to human interlocutors (Fryer et al., 2020; Loewen et al., 2020). These tools differ in functionality: IPAs provide context-aware, interactive conversational support through natural language processing and machine learning (Belda-Medina & Calvo-Ferrer, 2022; Tai, 2024), while AI-based applications offer structured, goal-oriented lessons with adaptive feedback tailored to learner needs (Lord, 2016; Rosell-Aguilar, 2018). IPAs simulate authentic conversations with native speakers, whereas AI applications focus on guided skill practice. Both tools have demonstrated potential in supporting the development of English as a foreign language (EFL) skills, including pronunciation (Han, 2020), speaking (Fathi et al., 2024; Hsu et al., 2023; Kim et al., 2021; Tai, 2024; Yang et al., 2024; Zou et al., 2023), listening (Tai & Chen, 2024), and vocabulary acquisition (Alrajhi, 2024).

Despite the growing integration of AI in language education, concerns remain regarding its potential to foster only superficial engagement, hinder cognitive development, and reduce peer interaction—ultimately weakening meaningful engagement in classroom settings (Dai & Liu, 2024; Zhai et al., 2024). Student engagement—defined as cognitive involvement, active participation, and sustained interest in learning—is a key determinant of positive educational outcomes such as academic persistence, higher achievement, and lower dropout rates (Christenson et al., 2012; Fredricks et al., 2004; Hiver et al., 2024). As AI tools become increasingly common in EFL learning environments, it is crucial to examine how different AI tools promote engagement compared to conventional teaching methods. Accordingly, the present study aims to (1) investigate how different types of AI tools—specifically IPAs and AI-based language learning applications—foster EFL adolescent engagement compared to peer interactions and teacher-led instruction, and (2) identify the factors that facilitate or hinder engagement across these different learning contexts.

Literature Review

Theoretical Background: Student Engagement

Student engagement is widely recognized as a multidimensional construct, commonly comprising three core components: cognitive, behavioral, and emotional engagement (Hiver et al., 2024; Fredricks et al., 2004). While some studies on task-specific engagement in L2 learning have identified additional dimensions—such as agentic and social engagement (Derakhshan & Zare, 2024; Zare & Derakhshan, 2024)—the present study adopts the widely accepted framework of cognitive, behavioral, and emotional engagement as its analytical lens, rather than task-specific engagement.

Cognitive engagement refers to the mental effort and strategic processes students employ to comprehend and master academic content (Fredricks et al., 2004; Henrie et al., 2015). As cognitive processes are not directly observable, researchers have assessed cognitive engagement using indicators such as metacognitive strategies (e.g., planning, monitoring, and evaluating) and learning strategies (e.g., critical thinking, rehearsal, and elaboration), typically measured through self-report questionnaires or observational methods (Fredricks et al., 2004).

Behavioral engagement refers to students' active participation in learning activities, observable through actions such as time on task, attendance, and class participation (Fredricks et al., 2004; Henrie et al., 2015). While cognitive engagement focuses on internal processes, behavioral engagement reflects outward signs of involvement and effort (Appleton et al., 2008).

Emotional engagement captures students' affective responses to learning, including feelings of enjoyment, interest, or frustration (Fredricks et al., 2004; Henrie et al., 2015). These emotional reactions influence motivation and connection to learning activities, playing a critical role in sustaining participation and fostering academic success (Henry & Thorsen, 2020).

To deepen the theoretical understanding of student engagement, Self-Determination Theory (SDT) offers a valuable theoretical framework. SDT posits that students are more likely to engage in learning when their basic psychological needs for competence, autonomy, and relatedness are fulfilled (Ryan & Deci, 2000, 2017). When students feel capable of succeeding (competence), have a sense of control over their learning (autonomy), and feel socially connected (relatedness), they are more likely to demonstrate sustained motivation and engagement. Therefore, it is essential to examine how various instructional modalities—such as AI-based tools or peer interaction—either support or impede the fulfillment of these needs, thereby influencing the depth and quality of student engagement.

Student engagement has attracted growing attention from scholars, policymakers, and educators (Zhou et al., 2021). High levels of engagement are associated with deeper learning, active participation, meaningful interaction, and sustained effort—all of which contribute to improved academic performance, greater persistence, enhanced motivation, and increased enjoyment (Fredricks et al., 2004; Hiver et al.,

2024; Ladd & Dinella, 2009; Wang & Fredricks, 2014). Engagement is also widely recognized as a malleable construct that can be influenced through targeted interventions, generating substantial interest in both educational and social contexts (Appleton et al., 2008). Given its changeable nature, it is essential to investigate whether the use of AI technologies in education enhances or undermines student engagement prior to their widespread adoption in school settings.

The Impact of AI Tools on EFL Student Engagement and Attitudes

Recent research highlights the growing potential of AI technologies to enhance student engagement and attitudes in language learning. AI-based language learning applications are known to offer personalized learning experiences, real-time adaptive feedback, and interactive activities tailored to learners' needs and proficiency levels (Ouyang et al., 2024; Yuan & Liu, 2025). For instance, Ouyang et al. (2024) reported that Chinese EFL university students who used Duolingo for 30 minutes daily over two months exhibited significantly higher cognitive, behavioral, and emotional engagement than a control group without access to the app. Similarly, Yuan and Liu (2025) found increased engagement and enjoyment among students who used Duolingo for 12 weeks compared to those receiving traditional instruction.

In the case of intelligent personal assistants, such as Google Assistant (GA), two studies investigated students' attitudes. Chen et al. (2020) found that Taiwanese EFL learners enjoyed interacting with GA and perceived it as useful and motivating for improving their speaking and listening skills. Likewise, Tai and Chen (2023) observed that two weeks of GA use enhanced eighth-grade students' willingness to communicate and reduced their speaking anxiety. Students especially valued GA's engaging features, such as games and jokes, which contributed to increased motivation, engagement, and confidence.

Other studies have explored the use of AI chatbots. Wang and Xue (2024) examined the impact of AI chatbots like TalkAI and SpeakG, finding significant improvements in cognitive, behavioral, and emotional engagement among Chinese EFL learners who used the tools for four weeks, compared to a control group receiving regular instruction. Similarly, Jeon (2024) found that Korean elementary students appreciated Google Dialogflow chatbots for their immediate feedback and anxiety-reducing features, while Fathi et al. (2024) reported that Iranian EFL adult learners valued the low-pressure speaking practice and personalized feedback provided by AI chatbots. However, both studies also noted that some learners still preferred human interaction—due to challenges in language proficiency (Jeon, 2024) or the lack of culturally nuanced feedback from AI (Fathi et al., 2024).

Several studies have also examined the combined use of AI applications and generative AI tools. Xu and Li (2024) reported that tools such as Grammarly, Duolingo, and ChatGPT enhanced Chinese EFL university students' cognitive, emotional, social, and behavioral engagement. Notably, the greatest improvements were seen among students with initially low engagement, suggesting AI's potential to re-engage disengaged learners. Similarly, Guo and Wang (2024) studied generative AI tools—including Bodoudou, Wenxin Yiyuan, and ChatGPT—and found significant gains in cognitive, emotional, and social engagement compared to traditional instruction, although behavioral engagement remained unchanged. While some students experienced anxiety or frustration, most expressed positive emotions such as enjoyment, anticipation, and satisfaction, which contributed to increased academic engagement. Additionally, Zhou and Hou (2024), using qualitative interviews, reported that AI tools like ChatGPT and Wenxin Yiyuan enhanced behavioral, cognitive, and emotional engagement among Chinese EFL teachers and tertiary-level students.

Despite these benefits, several studies also highlight the limitations of AI in language learning. Fryer et al. (2017) found that Japanese EFL university students' interest declined over 12 weeks when using a chatbot, while interest increased through sustained human interaction. This suggests that AI-based communication may have only short-term effects compared to the longer-lasting impact of human interaction. Zou et al. (2023) reported that although Chinese EFL students appreciated AI applications for their portability and immediate feedback, they remained skeptical about AI replacing face-to-face instruction due to issues such as poor recognition of nonnative pronunciation. Similarly, Dai and Liu

(2024) warned that although AI promotes engagement, motivation, and self-directed learning, it may also inhibit critical thinking, reduce creativity, and increase learner dependency on technology. These findings underscore the continued importance of human interaction in L2 learning to sustain interest and address AI's current limitations.

Methods

Research Design

While AI tools offer promising solutions to the limited speaking opportunities in EFL contexts—where English use outside the classroom is often minimal—their impact on student engagement compared to traditional methods, such as peer interactions and teacher-led instruction, remains underexplored. In particular, it is important to examine the differential effects of distinct types of AI tools, including IPA and AI-based learning applications, to advance the understanding of AI-supported EFL education in the digital age. Moreover, to provide insight into underrepresented population in existing research, this study focuses on middle school EFL students in South Korea. The following research questions guide this study:

1. How does EFL adolescents' engagement differ across AI tools (IPAs and AI applications), peer interactions, and teacher-led instruction?
2. What factors facilitate or hinder engagement in AI-based classes, peer interactions, and teacher-led instruction?

To answer these questions, a 12-week quasi-experimental study was conducted involving three experimental groups (IPA, AI applications, and peer interactions) and one control group (teacher-led instruction).

Participants and Research Context

This study involved 201 seventh-grade middle school students (102 boys and 99 girls, ages 12–13) from two public schools in South Korea. Initially, 234 students participated, but some were excluded due to incomplete questionnaires or school transfers. All students had been studying English since the third grade of elementary school. As shown in [Table 1](#), six homeroom classes from School A (27–30 students per class) were divided into three experimental groups, taught by English Teacher A. Two classes from School B (25–27), taught by English Teacher B, served as the control group. Before the intervention, students took a district-administered diagnostic listening test, which showed no significant differences among the four groups, $F(3, 190) = 2.11, p > .05$.

Table 1

Participant Information (N = 201)

	IPA (Google Assistant)	AI app (Buttertime)	Peer interactions	Control (Teacher-led)
School A (six classes)	53	51	52	
School B (two classes)				45

Note. IPA: Intelligent Personal Assistant.

A background survey revealed that 96% of students had little to no exposure to English-speaking cultures, and only 2.5% (5 students) had stayed abroad for more than six months. Additionally, about 25% of students reported prior experience with AI-assisted English learning. Both schools were located in the same city, used the same English textbook, and followed the national curriculum, ensuring consistency in instructional practices. The English teachers at both schools were similarly qualified: both were female, in their twenties, held bachelor's degrees in English education from the same university, and began teaching in the same year.

Treatments

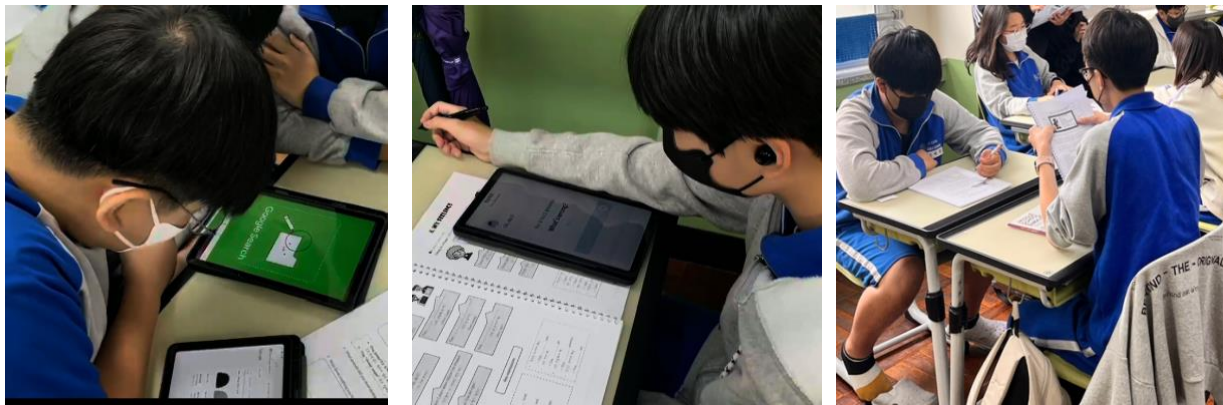
The 201 students participated in three 45-minute English classes each week: two regular lessons and one treatment session. Regular lessons were the same across all groups, focusing on listening, speaking, reading, vocabulary, and grammar using the same English textbook. For the treatment classes, students were assigned to one of four groups: IPA, AI app, peer interactions, or control. Over 12 weeks, three experimental groups used a researcher-designed workbook tailored to their respective instructional methods, with identical class structures (5-minute introduction, 35-minute practice, 5-minute reflection), while the control group (teacher-led instruction) continued with the textbook. Further details on the structure and content of the treatment classes are provided below.

IPA: Google Assistant

The Google Assistant (GA) group participated in weekly speaking activities using GA and a supplementary workbook. Based on Kim et al. (2019), the workbook tasks involved small talk, giving commands, seeking information, and problem-solving. Students completed these activities individually, in pairs, or in groups, using tablets equipped with earphones and microphones (see [Figure 1](#)). Each session followed a consistent structure: the teacher introduced the day's topic (5 minutes), facilitated student interaction with GA (35 minutes), and concluded the class by discussing students' experiences with the tool (5 minutes).

Figure 1

Treatment Snapshots: Google Assistant (left), Butvertime (middle), Peer Interactions (right)



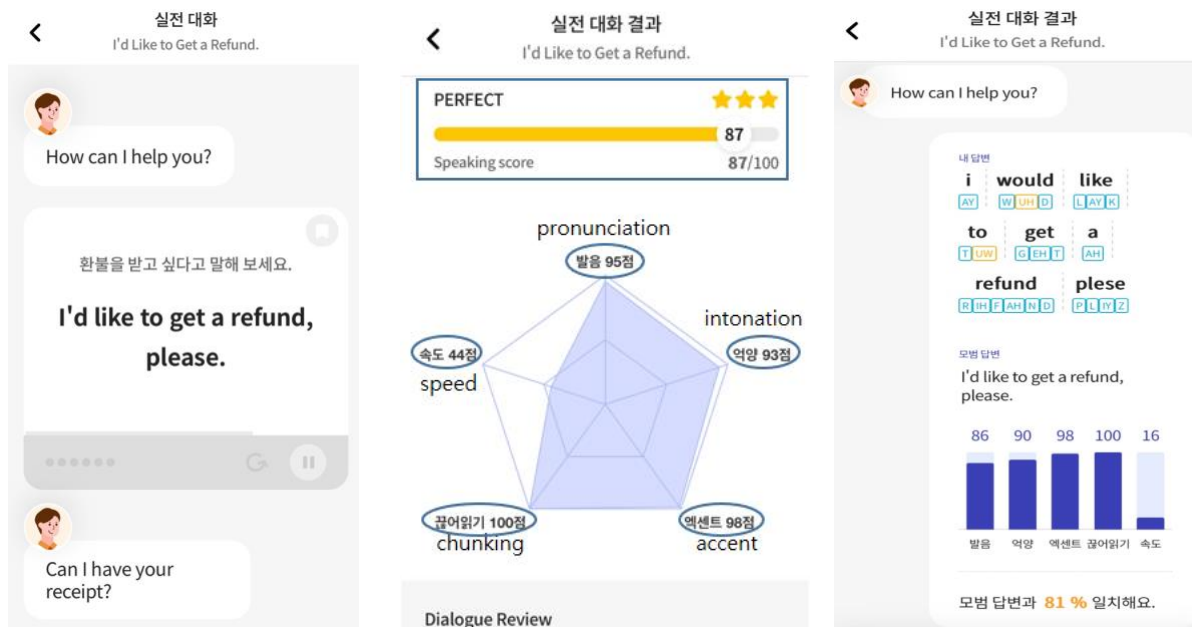
AI App: Butvertime

Butvertime is an AI-powered English tutoring app developed by LG CNS, designed to facilitate conversational practice in everyday situations such as giving directions, expressing emotions, and discussing the weather. Prior to the study, the researcher reviewed over 700 available units from the app and selected 36, organizing them into 12 themed lessons. For example, the “shopping” theme included expressions like “It’s a little expensive,” “It is on sale,” and “I’d like to get a refund.” As shown in [Figure 2](#), Butvertime uses a multistep approach: students listen to native speakers, practice target expressions in varied contexts, engage in role-play activities (evaluating pronunciation, speed, intonation, accent, and

chunking), and receive feedback through a scoring system. The lesson structure mirrored that of the GA group. Each session began with a 5-minute teacher-led introduction to the theme, followed by 35 minutes of individual practice on Butvertime using tablets and a workbook (Figure 1). The session concluded with a 5-minute reflection, during which the teacher collected student feedback on the learning experience.

Figure 2

Snapshots of the Butvertime Application



Peer Interactions

Peer-interaction sessions followed a similar structure to the AI-mediated lessons and also utilized a workbook. Each class began with a brief introduction to the day's theme by the teacher (5 minutes), followed by pair or group speaking activities (35 minutes), and ended with a class discussion on students' experiences (5 minutes). As in the AI sessions, the teacher served as a facilitator, guiding and supporting student interaction rather than directly leading the activities.

Control (Teacher-led Textbook-based Instruction)

The control group received conventional teacher-led instruction, focusing on the listening and speaking sections of the English textbook. Lessons included listening to audio recordings of target expressions, followed by comprehension questions. Students then practiced the target expressions through fixed dialogues with peers. These classes were predominantly teacher-led, with limited opportunities for peer interaction or extended speaking practice.

Instruments

Workbooks

Drawing on Chandra and Hayati's findings (2017) on the effectiveness of guided speaking workbooks in improving EFL adolescents' speaking skills, three types of student workbooks were developed as supplementary materials. These workbooks were designed to support and reinforce conversational English practice with either AI tools or peer interactions. Teacher versions, including suggested answers and instructional guidance, were also offered to teachers to support effective classroom implementation. As shown in Figure 3, each workbook featured 12 themes relevant to students' daily lives, such as making

friends, expressing feelings, sharing experiences, and discussing K-pop stars. Discussion topics were kept consistent across groups whenever possible. Each unit spanned four to five pages and included key vocabulary and model sentences to help students initiate and sustain conversations.

Figure 3

Workbook Snapshots: Google Assistant (left), Butvertime (middle), Peer Interactions (right)

12. My K-pop star

■ (Individual work & whole class) **생각제표:** 여러문은 좋아하는 K-pop 스타가 있나요? 인포지온은 K-pop 스타에 대해 얼마나 알고 있습니까? Google Assistant에게 본인이 좋아하는 K-pop 스타를 معرفی하고, 전체 학생들과 생각을 나누어 봅시다.

■ (Individual work) **서점:** **My K-pop star에 대해 Google Assistant에게 물어보세요.**

Facts about my K-pop star

- Who's in it? (이 그룹의 멤버는 누구인가요?)
- What is the BTS fandom name? (BTS의 팬덤 이름은 무엇인가요?)
- What does BTS stand for? (BTS는 무엇을 의미하나요?)
- When did they debut? (어떤 해에 데뷔했나요?)
- How many albums has BTS sold worldwide? (전 세계적으로 BTS가 얼마나 많은 앨범을 팔았나요?)
- How old are they? Where are they from? (그들은 몇 살이 되었나요? 어디에서 왔나요?)
- What is the most sold album? (가장 많이 팔린 앨범은 무엇인가요?)
- What is the name of the first album? (첫 번째 앨범의 이름은 무엇인가요?)
- How old did BTS become so successful? (BTS가 얼마나 빨리 성공했나요?)
- How many followers do BTS have on Twitter? (BTS는 트위터에 얼마나 많은 팔로워를 가지고 있나요?)
- What is BTS's best-selling song? (BTS의 가장 많이 팔린 노래는 무엇인가요?)
- How did BTS become so successful? (BTS가 어떻게 성공했나요?)

12. EXAM IS OVER

■ **All tutor가 대화해 보세요.**
Talking to Your Mom (부모님 친동생이 20. My exams are over. 섹션)

Mom: Where are you? (이 어디인가요?)

You: (학교가 있어요.) (Hint: work)

Mom: Oh, did you finish your exams? (시험은 다 끝났나요?)

You: (네, 모두 시험이 끝났어요.) (Hint: over)

Mom: Congratulations! Where are you now? (축하해! 지금 어디에 있나요?)

You: (지하철역에 있어요.) (Hint: subway)

Mom: Okay, see you soon. (그래, 곧 보자.)

Key expressions

- 시험이 끝났어요.
- 영어 수업이 끝났어.
- English class is 그 곳은 끝났어.
- The show is 축구 경기는 끝났어.
- The soccer game is
- 지하철역에 있어요.
- I'm the subway station. 그녀는 지하철역에 있어요.
- She is the train station. 그곳은 도서관에 있어요.
- They are the library. 책방에 있어요.
- I'm department store.

12. Our K-pop star

■ (Group activity) **조각의 미션:** 내가 좋아하는 그룹 홍보하기! 각 그룹별(4-5명)로 좋아하는 보이그룹 또는 걸그룹을 정하세요. 포스터 만들고 모든 걸그룹을 하나씩 알아 다음에 대해 알아보고, 아래 빈 칸에 정리해 봅시다.

Facts about our K-pop star

- How many are there in the group? (그룹에 몇 명?)
- What is the name of the group? (그룹 이름?)
- When did they debut? (어떤 해에 데뷔?)
- What are their names? (멤버 이름?)
- How old are they? Where are they from? (몇 살이 되었나요? 어디에서 왔나요?)
- When was/ will be their first concert? And where? (처음 콘서트 언제 어디에서?)
- What is the title of the latest song? (가장 최근 노래 제목?)
- What is the first album's name? (첫 번째 앨범 이름?)

Group Name	Member Names
First Album	Age/Hometown
Latest Song	First Concert
Fandom name	Debut year

Questionnaires

Online questionnaires were designed using Qualtrics. The pre-survey collected the participants' background information, including age, gender, experience in English-speaking countries, and prior exposure to AI technology for language learning.

The post-survey included two sections. The first section assessed student engagement across three dimensions: cognitive, behavioral, and emotional dimensions. It included 18 items, with six items per dimension, rated on a six-point Likert scale (1 = strongly disagree, 6 = strongly agree). A six-point Likert scale was selected to encourage participants to take a definitive stance, minimizing neutral responses, and to enhance discrimination and reliability compared to a five-point scale (Chomeya, 2010). Cognitive engagement measured the psychological and intellectual effort students invested in tasks, focusing on their use of knowledge and strategies (Fredricks et al., 2004, 2005; Henrie et al., 2015; Wu & Huang, 2007). Behavioral engagement captured active participation and observable behaviors, such as time spent on tasks or effort exerted (Fredricks et al., 2004, 2005; Henrie et al., 2015; Wu & Huang, 2007). Emotional engagement explored students' affective responses, including feelings of enjoyment, satisfaction, or boredom (Fredricks et al., 2004, 2005; Henrie et al., 2015; Wu & Huang, 2007). All 18 items were adapted from previously validated instruments and tailored to reflect the specific context of AI-mediated instruction, peer interaction, and teacher-led instruction. To ensure age-appropriateness and contextual relevance, the items were reviewed by the participating teachers in light of the students' age and learning environment. A complete list of the items is presented in [Table 2](#).

The second section examined student perceptions of the treatment classes through seven open-ended questions, adapted from previous research (Dai & Liu, 2024; Tai & Chen, 2023, 2024). It explored their language learning experiences with AI tools, peers, and the teacher; their opinions on the workbook or textbook; and the language skills they felt were improved. Additionally, it addressed moments when they

experienced a sense of partnership in class, their most valuable experiences, and how their questions were addressed by AI, peers, or the textbook. The section concluded with an invitation to share final thoughts, providing comprehensive insights into their overall learning experience. The questions are provided in [Appendix A](#).

Table 2

Student Engagement Questionnaire Items (N = 187)

Items	Loadings
Cognitive engagement (six items)	
4. The interactions with AI (peers/in textbook-based instruction) helped me think deeply.	.82
6. The interactions with AI (peers/ in textbook-based instruction) motivated me to challenge [myself] intellectually.	.72
9. I used different types of strategies to complete workbook(textbook) tasks.	.81
12. The interactions with AI (peers/ in textbook-based instruction) led me to engage in higher-order thinking processes, such as critical thinking, reasoning, and decision-making.	.75
16. I found the interactions with AI (peers/in textbook-based instruction) experience extremely rewarding.	.75
18. When I communicated with AI (peers/in textbook-based instruction), I asked myself or other friends to ensure I understood what it was about.	.66
Behavioral engagement (six items)	
1. I spent time speaking in English whenever I had a chance.	.71
5. I participated actively in English-speaking activities.	.80
7. I made extra effort to participate in English speaking.	.76
10. I worked very hard on completing speaking activities in a workbook.	.83
11. I continued participating in speaking activities persistently when confronted with difficulties related to unknown words, unfamiliar structures, or understanding sentences.	.78
14. I paid full attention to English-speaking activities.	.64
Emotional engagement (six items)	
2. While speaking with AI (peers/in textbook-based instruction), I felt that the way time passed seemed different from normal.	.71
3. I was highly motivated to speak with AI (peers/in textbook-based instruction).	.79
8. I enjoyed speaking with AI (peers/in textbook-based instruction).	.82
13. The interactions with AI (peers/in textbook-based instruction) were so interesting that they motivated me to engage in the task.	.72
15. I felt I had a great deal of control in the interactions with AI (peers/in textbook-based instruction).	.68
17. I was satisfied with the interactions with AI (peers/in textbook-based instruction).	.79

Note. Factor loading values are from the confirmatory factor analysis.

Procedures

As shown in [Figure 4](#), the study was conducted over a 14-week period, following a four-step data collection procedure:

1. Student recruitment. Students were chosen from two public middle schools. As the interventions were implemented as part of regular classroom instruction, formal recruitment was not required. However, informed consent was obtained from all students to ensure their understanding of the study's purpose and their rights.

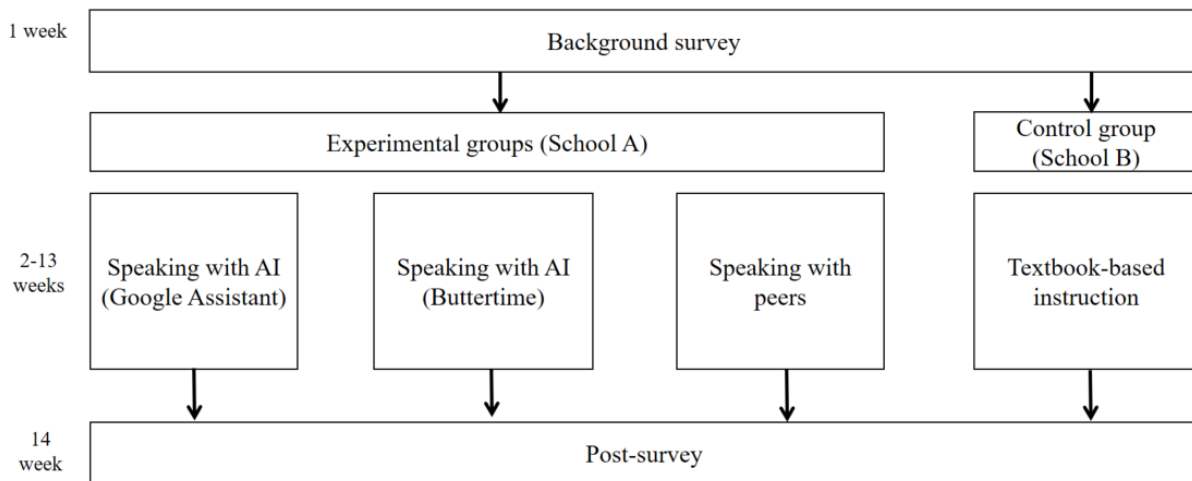
2. Pre-Questionnaire. In the first week, students completed a background survey during regular class time.

3. Interventions. Students were assigned to either experimental or control groups and participated in a 45-minute treatment class each week for 12 weeks. At School A, six classes were randomly assigned to one of three experimental groups—Google Assistant, Butvertime, or peer interactions—and received instruction from Teacher A. The control group received textbook-based English instruction from Teacher B.

4. Post-Questionnaire. In the 14th week, students completed a post-survey measuring engagement and responded to open-ended questions during regular class time.

Figure 4

Experimental Procedure



Data Analyses

This study employed a mixed-methods approach. Quantitative data were analyzed in three stages using SPSS 21.0 and Amos 21.0. First, descriptive statistics were calculated to assess assumptions such as data normality. Second, confirmatory factor analysis was conducted to verify whether the questionnaire items accurately represented the engagement constructs (cognitive, behavioral, and emotional). Model fit was evaluated using several indices: chi-square (χ^2 ; nonsignificant values indicating good fit), Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Goodness-of-Fit Index (GFI; values ≥ 0.90 indicating good fit), and standardized root mean square residual (SRMR; values < 0.08 considered good) (Bentler, 1990; Byrne, 2010; Hu & Bentler, 1999). After confirming the reliability and validity of the measures, a one-way analysis of variance (ANOVA) was conducted to examine the effects of the treatments on student engagement, with post hoc tests using Bonferroni adjustments.

Qualitative data from the open-ended questions were analyzed using NVivo, following the procedures of Miles and Huberman (1984) and Guba and Lincoln (1994). An initial coding scheme was developed to categorize emerging themes, which was refined through iterative analysis in collaboration with a second coder to ensure consistency and depth of interpretation. The finalized codes were then organized into broader dimensions of student engagement.

Ensuring Trustworthiness

To ensure the trustworthiness of the qualitative analysis, several procedures were implemented based on the criteria outlined by Guba and Lincoln (1994). Credibility was established through triangulation, which involved collecting data from both Likert-scale questionnaires and open-ended responses. Additionally, peer debriefing was conducted with a professional outside the research context to validate the researcher's interpretations and analytical decisions. Also, member checking was conducted with the two English teachers to verify the accuracy of the interpretations and ensure that the findings accurately reflected the students' perspectives and classroom realities. Transferability was supported by providing detailed descriptions of the study context, participant demographics, and the nature of the instructional interventions, allowing readers to assess the applicability of the findings to other EFL educational settings. Dependability was ensured through systematic documentation and the development of an audit trail that recorded coding decisions and any modifications made during the analysis process. Confirmability was addressed by maintaining reflective memos throughout the analysis to monitor potential researcher bias. Furthermore, to strengthen analytic rigor, a second coder independently analyzed the student responses. The two coders then compared their interpretations, refined the coding scheme, and resolved discrepancies through iterative discussions. As all differences were reconciled and full consensus was achieved, inter-coder reliability was not calculated.

Findings

Student Engagement

The engagement questionnaire was administered after the intervention, as its items specifically addressed students' experiences with AI, peers, or the textbook, making pre-intervention administration infeasible. Of the 201 students who participated, 187 completed the engagement questionnaire, and all 201 responded to the open-ended questions. To maintain data integrity, all responses were included in the analysis. Descriptive statistics for the engagement variable are presented in [Table 3](#). Skewness and kurtosis values fell within the acceptable range (± 3), confirming normal distribution. Across all four groups, mean scores for cognitive, behavioral, and emotional engagement exceeded the midpoint (3.5) on the six-point Likert scale, indicating generally high levels of student engagement.

The questionnaire items measuring student engagement were validated through confirmatory factor analysis. As shown in [Appendix B](#), most fit indices met acceptable thresholds, confirming the model's adequacy. Although the chi-square statistic did not meet the recommended criteria, this was likely due to the large sample size, which can inflate chi-square values and lead to statistical significance (Byrne, 2010). Alternative fit indices—including CFI, TLI, GFI, and SRMR—demonstrated a reasonable model fit, indicating that the items effectively represented the engagement construct. Reliability scores ranged from .88 to .89, reflecting strong internal consistency. Additionally, all factor loadings exceeded the acceptable threshold of .40 ([Table 2](#)), supporting the validity of the measurement model.

A one-way ANOVA was performed for each engagement factor to investigate the influence of the different treatments on student engagement (see [Table 4](#)). The results indicated that, although the mean scores of the GA group were higher than in the other groups, these differences were not statistically significant for cognitive engagement, $F(3, 183) = 0.94, p = .42$, or behavioral engagement, $F(3, 183) = 1.74, p = .16$. However, a significant difference was observed for emotional engagement, $F(3, 183) = 2.72, p < .05$. A post hoc analysis with Bonferroni adjustments revealed that the GA group reported significantly higher emotional engagement than the control group ($p < .05$). These results suggest that

cognitive and behavioral engagement did not vary significantly among the groups; nevertheless, GA appeared to promote significantly higher emotional engagement than teacher-led instruction. However, no significant differences in emotional engagement were observed between the GA, Butvertime, and peer-interaction groups.

Table 3

Descriptive Statistics of the Engagement Factors (N = 187)

Factors	Items	Min-Max	Mean (SD)	Skewness ^a	Kurtosis ^a
Cognitive	6	1.67–6.00	4.30 (1.05)	-.44 (-.18)	-.29 (.35)
Behavioral	6	2.17–6.00	4.72 (.94)	-.66 (.18)	.03 (.35)
Emotional	6	1.67–6.00	4.37 (1.07)	-.49 (.18)	-.56 (.35)

Note. SD: Standard deviations. ^aNumbers in parentheses are standard errors; Maximum score = 6.

Table 4

Student Engagement Across Groups (N = 187)

Factors	Groups	Mean	SD	F (3, 183)	Sig.	η_p^2
Cognitive	Google Assistant (n = 50)	4.48	1.10	.94	.424	—
	Butvertime (n = 45)	4.13	.95			
	Peer interaction (n = 49)	4.33	.97			
	Control (n = 43)	4.24	1.16			
Behavior	Google Assistant (n = 50)	4.93	.92	1.74	.16	—
	Butvertime (n = 45)	4.71	.76			
	Peer interaction (n = 49)	4.72	1.02			
	Control (n = 43)	4.48	1.01			
Emotional	Google Assistant (n = 50)	4.66	1.00	2.72	.046*	.04
	Butvertime (n = 45)	4.37	1.00			
	Peer interaction (n = 49)	4.38	1.00			
	Control (n = 43)	4.03	1.17			

Note. SD: Standard deviations, * $p < .05$

Factors that Promote or Hinder Student Engagement

To address the second research question, it is essential to identify the factors that promote or hinder student engagement in order to optimize the integration of AI tools and peer interactions in EFL classrooms. A content analysis was conducted using open-ended responses from all 201 students. Through an iterative coding and review process, facilitative or hindering factors were identified and subsequently categorized into the three dimensions of engagement: cognitive, behavioral, and emotional. [Appendix C](#) presents the number and proportion of students who referenced each factor within the different learning contexts, along with representative student responses.

Cognitive Dimension: Mental Effort and Use of Learning Strategies

As shown in [Appendix C](#), students in both the AI and peer interaction groups commonly reported cognitive engagement, including increased mental effort and the use of learning strategies, alongside improvements in speaking, pronunciation, vocabulary, and listening skills. However, the factors contributing to cognitive engagement differed by group, indicating that each intervention promoted cognitive engagement through distinct mechanisms.

Notably, over 70% of students in all three groups reported exerting mental effort and engaging in self-regulated learning to improve their speaking and pronunciation. The perceived sources of improvement differed across contexts: the GA group highlighted the value of realistic conversations and native-like pronunciation; the Butterside group emphasized real-time feedback, scoring, and repeated practice of target expressions; meanwhile, the peer interaction group attributed their progress to mutual feedback and consistent opportunities for speaking practice—factors that differ from the precise modeling and automated feedback offered by AI tools.

Furthermore, over 60% of students across the groups reported using learning strategies to enhance vocabulary and listening skills, though the underlying mechanisms varied. Students in the GA group noted that, while the GA's fast-paced and complex responses often exceeded their proficiency levels, this challenge stimulated curiosity, deepened cognitive engagement, and encouraged self-directed learning. They employed both cognitive and metacognitive strategies, such as monitoring comprehension and seeking external resources. In contrast, students using Butterside credited their gains to the repeated presentation of vocabulary in diverse contexts, which aided retention and application. Repeated exposure to native-speaker pronunciation also familiarized them with natural speech patterns, thereby improving listening comprehension. The peer interaction group pointed to meaningful, ongoing conversations with classmates as the primary source of their vocabulary and listening development, using social strategies to acquire language through authentic communication.

By comparison, few students in the control group mentioned engaging in deeper cognitive processes such as mental effort or strategic learning. A key reason appears to be the limited speaking opportunities in the teacher-led classroom—explicitly noted by 20% of students—which reduced the need for cognitive engagement in developing oral skills.

Behavioral Dimension: Effort and Active Participation

While the factors contributing to cognitive engagement varied across interventions, behavioral engagement was consistently promoted by common elements across all three experimental groups (see [Appendix C](#)). First, more than 75% of students in each group demonstrated effort and persistence when they observed tangible improvements in their English skills. Whether during class activities or real-life situations—such as test-taking or watching English-language movies—students recognized their progress, which motivated them to continue making sustained effort to learn. It is also noteworthy that although 78% of students in the control group (who received textbook-based, teacher-led instruction) reported overall improvement in English proficiency, their responses reflected few signs of sustained effort or persistence in future learning.

Second, exposure to authentic, practical, and engaging content, along with relevant English expressions, encouraged students to actively participate in the interventions, whether interacting with AI tools or peers. All three groups used researcher-designed workbooks featuring 12 themes closely related to students' daily lives, such as making friends, expressing feelings, sharing experiences, and discussing K-pop stars. This engaging content and practical language use contributed to students' active involvement.

Emotional Dimension: Increased Interests, Enjoyment, and Sense of Belonging

Across all three experimental groups, over 70% of students reported increased interest, enjoyment, and confidence as a result of their respective interventions, although the sources of emotional engagement varied by group. Students in the GA group described the interaction as novel and engaging, highlighting

the assistant's ability to understand their input and respond in a human-like, sometimes humorous manner. Given the limited opportunities EFL learners have to communicate with native English speakers, interacting with GA, which was perceived as a native-like speaker, proved emotionally stimulating. In the Butterside group, emotional engagement was primarily driven by the app's immediate feedback through ratings and scores. Positive reinforcement, such as earning high scores, fostered a sense of achievement and sustained motivation. Although some students expressed frustration with challenges in obtaining high scores due to pronunciation errors or voice recognition issues, many found overcoming these difficulties rewarding. In contrast, emotional engagement in the peer interaction group stemmed from collaborative learning experiences. Students highlighted the enjoyment of sharing ideas, offering support, and creating a positive and focused classroom atmosphere with their peers.

Another key factor contributing to emotional engagement across all groups was students' perception of AI or human peers as learning partners. This sense of belonging increased students' willingness to interact, collaborate, and participate actively in speaking tasks. In the GA group, 80% of students reported feeling they had a genuine learning partner, citing the assistant's capacity for two-way conversations as a motivating factor. Similarly, 60% of students in the Butterside group perceived the app as a supportive partner, appreciating its real-time, personalized feedback and its ability to offer praise and constructive corrections. Meanwhile, 65% of students in the peer interaction group viewed their classmates as true learning partners. As shown in [Appendix C](#), peer interactions fostered a deeper sense of partnership built on collective growth, trust, and mutual respect—qualitatively different from AI-based partnerships. While students in the AI groups emphasized responsiveness and two-way interaction, those in the peer group valued mutual assistance, shared responsibility, and emotional support, which AI tools could not fully replicate.

Despite these positive responses, some students in the Butterside and control groups reported negative emotions, including boredom and frustration. In the Butterside group, 53% of students expressed dissatisfaction with the app's limited and preset responses, which hindered dynamic interaction and reduced the perception of the AI as an autonomous conversational partner. Many felt that Butterside was better suited for beginners and lacked the adaptability and natural tone needed for authentic communication, leading some to prefer human interactions. Likewise, 35% of control group students described textbook-based instruction as unengaging and limited in expression variety, citing a lack of interactive or stimulating activities. These findings underscore the importance of designing speaking tasks that are interactive, contextually relevant, and aligned with learners' interests to enhance emotional engagement.

Discussion

This study investigates how EFL adolescents' engagement differs when using AI tools compared to peer interactions and teacher-led instruction. Questionnaire results showed no significant differences in cognitive or behavioral engagement across the groups. However, students in the GA group reported significantly higher emotional engagement than those in the control group. Additionally, qualitative responses offered deeper insights into students' experiences, complementing the quantitative findings.

IPA: Google Assistant

The GA group demonstrated the highest levels of cognitive ($M = 4.48$), behavioral ($M = 4.93$), and emotional engagement ($M = 4.66$) among the four groups. Emotional engagement is particularly noteworthy, as it was significantly higher in the GA group than in the control group. One key factor contributing to this heightened emotional engagement was GA's ability to simulate realistic, human-like conversations, which students described as similar to speaking with a native English speaker. Given their limited opportunities to interact with foreign speakers, students found GA's two-way, interactive dialogue both novel and engaging. Notably, approximately 80% of students viewed GA as a genuine learning partner, underscoring the extent to which they enjoyed conversing with the tool as if it were a real person.

These findings align with those of Tai and Chen (2023), who reported that most Taiwanese EFL learners experienced excitement and enjoyment when interacting with GA. The present study reinforces their results, highlighting GA's emotionally engaging nature for Korean EFL adolescents. Enhancing emotional engagement is especially critical, as it serves as a foundation for other forms of engagement. Emotional engagement has been identified as a key driver of cognitive engagement (Li & Lerner, 2013), a contributor to motivation and academic achievement (Özhan & Kocadere, 2020), and a predictor of overall student engagement (Henry & Thorsen, 2020). Wang (2022) also found that L2 enjoyment strongly predicts academic engagement among EFL learners. Taken together, the findings from previous research and the current study suggest that integrating GA into EFL classrooms can not only provide authentic speaking opportunities but also foster enjoyable and emotionally engaging learning environments.

Despite its advantages, students reported that GA's fast-paced and complex responses often exceeded their proficiency levels in terms of speed, vocabulary, and sentence complexity. This aligns with findings by Chen et al. (2020), who observed that less proficient EFL learners struggled with GA due to pronunciation issues, while more advanced learners achieved greater mutual comprehensibility. Although many students in the current study found GA's responses challenging, these difficulties paradoxically motivated them to adopt cognitive and metacognitive strategies, such as engaging in additional self-study or verifying their understanding. As Meng et al. (2016) suggest, cognitively demanding tasks can foster intrinsic motivation, encouraging learners to invest greater mental effort. In this regard, GA appears to offer an optimal level of challenge that stimulates students to engage more deeply in the learning process.

Furthermore, promoting behavioral engagement—such as persistence and active participation—requires helping students recognize their own progress and providing authentic, engaging content. In this study, approximately 76% of students reported perceivable improvement in their English skills through interactions with GA, which likely contributed to their sustained effort and motivation to learn. This finding aligns with previous research indicating that when students recognize their progress, they tend to take greater responsibility for their learning and participate more actively (Zimmerman, 2002). In addition, consistent with Tai and Chen's findings (2023) that learners enjoyed GA activities featuring interactive and engaging elements, the topics used in the present study were closely aligned with students' interests and real-life contexts. As a result, 96% of students regarded the content as authentic, practical, and engaging. Taken together, prior research and the current findings underscore the importance of enhancing students' awareness of their learning progress and incorporating meaningful, relevant content to foster behavioral engagement when implementing GA.

AI Applications: Butvertime

Butvertime demonstrated cognitive ($M = 4.13$), behavioral ($M = 4.71$), and emotional ($M = 4.37$) engagement levels comparable to those of GA. This suggests that EFL adolescents were similarly engaged by both types of AI tools—whether simulating free-form conversations like IPAs or offering structured speaking practice through AI-based applications. Indeed, students in the Butvertime group reported perceived improvement in speaking, pronunciation, vocabulary, and listening skills, attributing this to real-time feedback, repeated practice in diverse scenarios, and positive reinforcement—factors that contributed to enhanced cognitive engagement. As students observed tangible progress in their English skills, they expressed increased motivation and persistence to continue learning. Additionally, they valued learning authentic, practical expressions through Butvertime, which were directly applicable to real-life contexts—further promoting sustained participation and behavioral engagement. These results are consistent with previous research suggesting that interactive features, real-time responsiveness, and gamified elements in AI language learning apps foster engagement (Ouyang et al., 2024; Xu & Li, 2024; Yuan & Liu, 2025).

In terms of emotional engagement, students enjoyed receiving immediate feedback and scores, as well as interacting with an AI partner who corrected their errors. However, unlike GA, Butvertime did not lead to significantly higher emotional engagement compared to the control group, which received textbook-

based, teacher-led instruction. This contrasts with prior findings suggesting that AI applications—such as Duolingo, SpeakG, and TALKAI—typically result in greater emotional engagement than traditional textbook-based lessons and lectures (Guo & Wang, 2024; Ouyang et al., 2024; Yuan & Liu, 2025; Wang & Xue, 2024).

One possible explanation for this discrepancy is students' mixed emotional responses toward AI applications. While 83% of students reported increased interest and enjoyment due to Butvertime's immediate feedback and supportive features, 53% pointed out limitations, including restricted responses and less natural communication (see [Appendix C](#)). Some even expressed a preference for human interaction, emphasizing the value of more flexible and human-like conversations. These findings suggest that, to sustain emotional engagement in EFL classrooms, AI applications like Butvertime must further enhance their conversational naturalness and flexibility, distinguishing themselves more clearly from conventional textbook-based instruction.

Peer Interactions

An important question in existing research on AI technologies is whether AI tools can promote levels of engagement comparable to peer interactions. The current study found that, while AI tools offer distinct advantages, their overall effectiveness in fostering engagement is similar to that of peer-based learning. This finding aligns with previous studies highlighting AI's positive impact on EFL students' engagement, motivation, confidence, and reduced anxiety (Chen et al., 2020; Tai & Chen, 2023; Ebadi & Amini, 2024). However, it contrasts with research showing divergent student preferences—some students favor AI for its flexibility and low-pressure environment (Fathi et al., 2024; Kim, 2017), while others prefer peer interaction for its socioemotional benefits and opportunities for collaborative learning (Fryer et al., 2017; Jeon, 2024).

Although AI tools provide personalized instruction and can effectively engage even less motivated learners (Guo & Wang, 2024; Xu & Li, 2024), the benefits of peer interaction remain significant. In this study, students in the peer group emphasized the value of collaborative learning and a supportive community, which fostered not only perceived English improvement but also interpersonal skills such as mutual respect, encouragement, and shared progress. According to Self-Determination Theory (Ryan & Deci, 2000, 2017), student engagement is enhanced when learners experience competence, autonomy, and relatedness. Notably, the sense of social connection and collaboration fostered through peer interaction is a critical element of engagement that current AI tools may struggle to replicate fully.

Overall, the findings suggest that while AI tools can enhance student engagement, peer interactions play a unique role in creating a supportive and socially enriching learning environment. Classrooms that cultivate a caring, supportive atmosphere have been shown to promote greater engagement (Fredricks et al., 2004). For adolescents in particular, a sense of belonging and peer support are crucial, influencing motivation, persistence, and attitudes toward learning (Appleton et al., 2008). Higher levels of relatedness also enhance behavioral and emotional engagement (Appleton et al., 2008; Ryan & Deci, 2000, 2017). Therefore, to optimize the learning experience in AI-mediated classrooms, it is essential to incorporate opportunities for human connection and social interaction.

Conclusion

This study contributes to the growing body of research on technology-enhanced language learning by demonstrating that IPAs, AI language learning applications, and peer interactions significantly enhance student engagement, especially emotional dimension, among EFL adolescents, compared to traditional teacher-led, textbook-based instruction. The findings suggest that AI tools can be effectively integrated into public EFL education to promote student engagement in speaking at levels comparable to those achieved through human interaction.

From a theoretical perspective, the study highlights the central role of emotional engagement—particularly interest, enjoyment, and a sense of belonging—in sustaining and facilitating other dimensions of engagement, such as behavioral and cognitive involvement. This reinforces the view that fostering positive emotional responses to learning is critical in maintaining students’ overall engagement in EFL contexts.

Practically, the findings offer several implications for EFL educators seeking to integrate AI technologies effectively. First, it is essential to design workbooks and tasks that reflect students’ real-life experiences and provide authentic, meaningful opportunities for language use. These materials can help bridge the gap between digital interaction and real-world application. Second, teachers should ensure that students can track their own progress through tools such as self-reflection activities. This allows learners to recognize their improvement and maintain engagement when using AI tools. Third, educators should clearly position AI tools as supportive partners rather than replacements for teacher-led instruction. When framed appropriately, AI tools can offer low-stakes environments that encourage risk-taking, build learner confidence through repetition, and deliver immediate feedback. Finally, while AI tools offer unique advantages, they should complement—not replace—human interaction. While AI offers valuable support, educators should not overlook the benefits of peer collaboration. Incorporating peer interaction—especially after AI-mediated tasks—can foster collaborative learning and contribute to a more supportive classroom environment.

Despite these contributions, the study has several limitations. First, student engagement was assessed primarily through self-report questionnaires, which may be subject to bias. Second, engagement was measured only at the post-intervention stage, without baseline data to account for pre-existing differences in engagement levels. Third, the study focused on specific types of AI tools—namely IPAs and AI learning applications—which may limit the generalizability of the findings to other AI technologies or learning contexts.

Future research should address these limitations by incorporating mixed-method approaches, including interviews and classroom observations, to gain deeper insights into how students engage with AI and peer-based activities. In addition, studies should adopt pre- and post-intervention measures to more accurately assess changes in engagement over time. Expanding the scope of AI tools and exploring their long-term impact on multiple dimensions of engagement would also offer valuable directions for further inquiry.

Acknowledgements

This work was supported by 2025 Hongik University Research Fund.

References

- Alrajhi, A. S. (2024). Intelligent personal assistants in self-access L2 vocabulary learning. *Education and Information Technologies*, 29(1), 1–31. <https://doi.org/10.1007/s10639-024-12967-7>
- Appleton, J. J., Christenson, S. L., & Furlong, M. J. (2008). Student engagement with school: Critical conceptual and methodological issues of the construct. *Psychology in the Schools*, 45(5), 369–386. <https://doi.org/10.1002/pits.20303>
- Belda-Medina, J., & Calvo-Ferrer, J. R. (2022). Using chatbots as AI conversational partners in language learning. *Applied Sciences*, 12(17), 8427. <https://doi.org/10.3390/app12178427>
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107(2), 238–246. <https://doi.org/10.1037/0033-2909.107.2.238>
- Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (2nd ed.). Taylor & Francis Group.

- Chandra, J., & Hayati, A. (2018). Using speaking worksheets in enhancing accuracy in EFL adults' spoken English. In *International Conference on ELT and CALL* (Vol. 2017, p. 113).
- 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*, 28(2), 1–16. <https://doi.org/10.1080/10494820.2020.1833043>
- Chomeya, R. (2010). Quality of psychology test between Likert scale 5 and 6 points. *Journal of Social Sciences*, 6(3), 399–403.
- Christenson, S. L., Reschly, A. L., & Wylie, C. (Eds.). (2012). *Handbook of research on student engagement*. Springer. <https://doi.org/10.1007/978-1-4614-2018-7>
- Dai, K., & Liu, Q. (2024). Leveraging artificial intelligence (AI) in English as a foreign language (EFL) classes: Challenges and opportunities in the spotlight. *Computers in Human Behavior*, 139, 108354. <http://dx.doi.org/10.1016/j.chb.2024.108354>
- Derakhshan, A., & Zare, J. (2024). Self-regulated learning and task engagement: a SEM analysis. *International Review of Applied Linguistics in Language Teaching*. <https://doi.org/10.1515/iral-2024-0009>
- Ebadi, S., & Amini, A. (2024). Examining the roles of social presence and human-likeness on Iranian EFL learners' motivation using artificial intelligence technology: A case of CSIEC chatbot. *Interactive Learning Environments*, 32(2), 655–673. <http://dx.doi.org/10.1080/10494820.2022.2096638>
- Fathi, J., Rahimi, M., & Derakhshan, A. (2024). Improving EFL learners' speaking skills and willingness to communicate via artificial intelligence-mediated interactions. *System*, 121, 103254. <http://dx.doi.org/10.1016/j.system.2024.103254>
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74(1), 59–109. <https://doi.org/10.3102/00346543074001059>
- Fredricks, J. A., Blumenfeld, P. C., Friedel, J., & Paris, A. (2005). School engagement. In K. A. Moore & L. H. Lippman (Eds.), *What do children need to flourish? Conceptualizing and measuring indicators of positive development* (pp. 305–321). Springer. https://doi.org/10.1007/0-387-23823-9_19
- Fryer, L. K., Ainley, M., Thompson, A., Gibson, A., & Sherlock, Z. (2017). Stimulating and sustaining interest in a language course: An experimental comparison of Chatbot and Human task partners. *Computers in Human Behavior*, 75, 461–468. <https://doi.org/10.1016/j.chb.2017.05.045>
- Fryer, L., Coniam, D., Carpenter, R., & Lăpușneanu, D. (2020). Bots for language learning now: Current and future directions. *Language Learning & Technology*, 24(2), 8–22. <http://hdl.handle.net/10125/44719>
- Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105–117). Sage Publications.
- Guo, Y., & Wang, Y. (2024). Exploring the effects of artificial intelligence application on EFL students' academic engagement and emotional experiences: A mixed-methods study. *European Journal of Education*. <https://doi.org/10.1111/ejed.12812>
- Jeon, J. (2024). Exploring AI chatbot affordances in the EFL classroom: Young learners' experiences and perspectives. *Computer Assisted Language Learning*, 37(1–2), 1–26. <http://dx.doi.org/10.1080/09588221.2021.2021241>

- Han, D. E. (2020). The effects of voice-based AI chatbots on Korean EFL middle school students' speaking competence and affective domains. *Asia-pacific Journal of Convergent Research Interchange*, 6(7), 71–80. <http://dx.doi.org/10.47116/apjcri.2020.07.07>
- Henrie, C. R., Halverson, L. R., & Graham, C. R. (2015). Measuring student engagement in technology mediated learning: A review. *Computers & Education*, 90, 36–53. <http://dx.doi.org/10.1016/j.compedu.2015.09.005>
- Henry, A., & Thorsen, C. (2020). Disaffection and agentic engagement: Redesigning activities to enable authentic self-expression. *Language Teaching Research*, 24(4), 456–475. <http://dx.doi.org/10.1177/1362168818795976>
- Hiver, P., Al-Hoorie, A. H., Vitta, J. P., & Wu, J. (2024). Engagement in language learning: A systematic review of 20 years of research methods and definitions. *Language Teaching Research*, 28(1), 201–230. <https://doi.org/10.1177/13621688211001289>
- Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Hsu, H. L., Chen, H. H. J., & Todd, A. G. (2023). Investigating the impact of the Amazon Alexa on the development of L2 listening and speaking skills. *Interactive Learning Environments*, 31(9), 5732–5745. <https://doi.org/10.1080/10494820.2021.2016864>
- Kim, N. Y. (2017). Effects of types of voice-based chat on EFL students' negotiation of meaning according to proficiency levels. *English Teaching*, 72(1), 159–181. <https://doi.org/10.15858/engtea.72.1.201703.159>
- Kim, H., Shin, D., Yang, H., & Lee, J. (2019). A study of AI chatbot as an assistant tool for school English curriculum. *Journal of Learner-Centered Curriculum and Instruction*, 19(1), 89–110. <https://doi.org/10.22251/jlcci.2019.19.1.89>
- Kim, H. S., Kim, N. Y., & Cha, Y. (2021). Is it beneficial to use AI chatbots to improve learners' speaking performance? *Journal of Asia TEFL*, 18(1), 161–178. <http://dx.doi.org/10.18823/asiatefl.2021.18.1.10.161>
- Ladd, G. W., & Dinella, L. M. (2009). Continuity and change in early school engagement: Predictive of children's achievement trajectories from first to eighth grade? *Journal of Educational Psychology*, 101(1), 190–206. <https://doi.org/10.1037/a0013153>
- Li, P., & Jeong, H. (2020). The social brain of language: Grounding second language learning in social interaction. *Science of Learning*, 5(1), 1–9. <http://dx.doi.org/10.1038/s41539-020-0068-7>
- Li, Y., & Lerner, R. M. (2013). Interrelations of behavioral, emotional, and cognitive school engagement in high school students. *Journal of Youth and Adolescence*, 42(1), 20–32. <https://doi.org/10.1007/s10964-012-9857-5>
- Loewen, S., Isbell, D. R., & Sporn, Z. (2020). The effectiveness of app-based language instruction for developing receptive linguistic knowledge and oral communicative ability. *Foreign Language Annals*, 53(2), 209–233. <https://doi.org/10.1111/flan.12454>
- Lord, G. (2016). Rosetta Stone for language learning: An exploratory study. *IALLT Journal of Language Learning Technologies*, 46(1), 1–35. <http://dx.doi.org/10.17161/iallt.v46i1.8552>
- Meng, L., Pei, G., Zheng, J., & Ma, Q. (2016). Close games versus blowouts: Optimal challenge reinforces one's intrinsic motivation to win. *International Journal of Psychophysiology*, 110, 102–108. <http://dx.doi.org/10.1016/j.ijpsycho.2016.11.001>

- Miles, M. B., & Huberman, A. M. (1984). *Qualitative data analysis: A sourcebook of new methods*. SAGE.
- Ouyang, Z., Jiang, Y., & Liu, H. (2024). The effects of Duolingo, an AI-integrated technology, on EFL learners' willingness to communicate and engagement in online classes. *International Review of Research in Open and Distributed Learning*, 25(3), 97–115. <http://dx.doi.org/10.19173/irrodl.v25i3.7677>
- Özhan, Ş. Ç., & Kocadere, S. A. (2020). The effects of flow, emotional engagement, and motivation on success in a gamified online learning environment. *Journal of Educational Computing Research*, 57(8), 2006–2031. <https://doi.org/10.1177/0735633118823159>
- Rosell-Aguilar, F. (2018). Autonomous language learning through a mobile application: A user evaluation of the *busuu* app. *Computer Assisted Language Learning*, 31(8), 854–881. <https://doi.org/10.1080/09588221.2018.1456465>
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55, 68–78. <https://doi.org/10.1037//0003-066x.55.1.68>
- Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. The Guilford Press.
- Tai, T. Y. (2024). Effects of intelligent personal assistants on EFL learners' oral proficiency outside the classroom. *Computer Assisted Language Learning*, 37(5–6), 1281–1310. <https://doi.org/10.1080/09588221.2022.2075013>
- Tai, T. Y., & Chen, H. H. J. (2023). The impact of Google Assistant on adolescent EFL learners' willingness to communicate. *Interactive Learning Environments*, 31(3), 1485–1502. <https://doi.org/10.1080/10494820.2020.1841801>
- Tai, T. Y., & Chen, H. H. J. (2024). The impact of intelligent personal assistants on adolescent EFL learners' listening comprehension. *Computer Assisted Language Learning*, 37(3), 433–460. <https://doi.org/10.1080/09588221.2022.2040536>
- Wang, X. (2022). Enhancing Chinese EFL students' academic engagement: The impact of L2 enjoyment and academic motivation. *Frontiers in Psychology*, 13, 914682. <https://doi.org/10.3389/fpsyg.2022.914682>
- Wang, M. T., & Fredricks, J. A. (2014). The reciprocal links between school engagement, youth problem behaviors, and school dropout during adolescence. *Child Development*, 85(2), 722–737. <https://doi.org/10.1111/cdev.12138>
- Wang, Y., & Xue, L. (2024). Using AI-driven chatbots to foster Chinese EFL students' academic engagement: An intervention study. *Computers in Human Behavior*, 140, 108353. <http://dx.doi.org/10.1016/j.chb.2024.108353>
- Wu, X.-Y. (2024). AI in L2 learning: A meta-analysis of contextual, instructional, and social-emotional moderators. *System*, 126, 103498. <https://doi.org/10.1016/j.system.2024.103498>
- Wu, H. K., & Huang, Y. L. (2007). Ninth-grade student engagement in teacher-centered and student-centered technology-enhanced learning environments. *Science Education*, 91(5), 727–749. <https://doi.org/10.1002/sce.20216>
- Xu, J., & Li, J. (2024). Effects of AI affordances on student engagement in EFL classrooms: A structural equation modelling and latent profile analysis. *European Journal of Education*, 59(4), e12808. <https://doi.org/10.1111/ejed.12808>

- Yang, C. T. Y., Lai, S. L., & Chen, H. H. J. (2024). The impact of intelligent personal assistants on learners' autonomous learning of second language listening and speaking. *Interactive Learning Environments*, 32(5), 2175–2195. <https://doi.org/10.1080/10494820.2022.2141266>
- Yuan, L., & Liu, X. (2025). The effect of artificial intelligence tools on EFL learners' engagement, enjoyment, and motivation. *Computers in Human Behavior*, 162, 108474. <http://dx.doi.org/10.1016/j.chb.2024.108474>
- Zare, J., & Derakhshan, A. (2024). Task engagement in second language acquisition: A questionnaire development and validation study. *Journal of Multilingual and Multicultural Development*, 1–17. <https://doi.org/10.1080/01434632.2024.2306166>
- Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review. *Smart Learning Environments*, 11(28), 1–37. <https://doi.org/10.1186/s40561-024-00316-7>
- Zhou, S., Hiver, P., & Al-Hoorie, A. H. (2021). Measuring L2 engagement: A review of issues and applications. In Hiver, P., Al-Hoorie, A.H., & Mercer, S. (Eds.), *Student engagement in the language classroom* (pp. 75–98). Multilingual Matters.
- Zhou, C., & Hou, F. (2024). Can AI empower L2 education? Exploring its influence on the behavioural, cognitive, and emotional engagement of EFL teachers and language learners. *European Journal of Education*, 59(4), e12750. <https://doi.org/10.1111/ejed.12750>
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, 41(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2
- Zou, B., Du, Y., Wang, Z., Chen, J., & Zhang, W. (2023). An investigation into artificial intelligence speech evaluation programs with automatic feedback for developing EFL learners' speaking skills. *SAGE Open*, 13(3), 1–9. <https://doi.org/10.1177/21582440231193818>

Appendix A. Open-Ended Questions

1. When did you feel that your English speaking skills improved through practice with AI (or a peer or textbook)? Please describe the experience in detail and explain why you felt that way.
2. Which area—speaking, listening, vocabulary, or pronunciation—do you feel improved the most through English speaking practice with AI (or a peer or textbook)? Please explain your choice in detail.
3. At what moment did you feel that AI (or a peer or textbook) was a learning partner while practicing English speaking? Please describe your experience in detail.
4. What was the most enjoyable experience you had while practicing English speaking with AI (or a peer or textbook)? Please describe the activity and explain why it was enjoyable.
5. Did you ever feel that AI (or a peer or textbook) responded appropriately to your questions or answers? If so, please describe the experience in detail.
6. Can you describe a specific moment when AI (or a peer or textbook) helped you concentrate better on English speaking or workbook tasks? Please explain how it supported your learning.
7. Finally, please share any thoughts or feedback you have about participating in the AI (or peer or textbook) English speaking class this semester. Feel free to express your honest opinions.

Appendix B. Fit Indices for Confirmatory Factor Analysis Models

Models	Cronbach's alpha	X ²	df	p	CFI	TLI	GFI	SRMR
Acceptable fit	> .70			> .05	≥ .90	≥ .90	≥ .90	≥ .08
Cognitive	.88	25.53	9	< .05	.97	.95	.96	.04
Behavioral	.88	30.84	9	< .05	.96	.94	.95	.04
Emotional	.89	26.64	9	< .05	.97	.95	.96	.03

Note. CFI: Comparative Fit Index, TLI: Tucker–Lewis Index, GFI: Goodness-of-Fit Index, SRMR: standardized root mean square residual

Appendix C. Factors Contributing to Student Engagement Across Groups: Google Assistant ($n = 54$), Butvertime ($n = 53$), Peer interactions ($n = 54$), and Control ($n = 46$)

Appendix C is available through IRIS: <https://doi.org/10.48316/YPKFJ-F5FBH>

About the Author

Juhee Lee is an Associate Professor in the Department of English Education at Hongik University in South Korea. Her research interests include technology-mediated language learning, reading and writing, learner motivation and attitudes, and teacher education.

E-mail: juheeleecarpediem@gmail.com

ORCID: <http://orcid.org/0000-0002-3109-5140>