Check for updates

Received: 29 November 2024 | Accepted: 17 June 2025

DOI: 10.1111/bjet.13615

ORIGINAL ARTICLE

British Journal of Educational Technology BERA

Parent-led vs. AI-guided dialogic reading: Evidence from a randomized controlled trial in children's e-book context

Feiwen Xiao[1] | **Ellen Wenting Zou**[2] | **Jiaju Lin**[2] | **Zhaohui Li**[3] | **Dandan Yang**[4,5]

[1]Department of Learning and Performance Systems, The Pennsylvania State University, State College, Pennsylvania, USA

[2]Department of Educational Psychology, Counseling, and Special Education, The Pennsylvania State University, State College, Pennsylvania, USA

[3]National AI Institute for Exceptional Education, Institute for Artificial Intelligence and Data Science (IAD), University at Buffalo, Buffalo, New York, USA

[4]School of Education, University of California, Irvine, Irvine, California, USA

[5]Center for Community Schooling, University of California, Los Angeles, Los Angeles, California, USA

Correspondence
Feiwen Xiao, Department of Learning and Performance Systems, The Pennsylvania State University, State College, PA, USA.
Email: ffx5014@psu.edu

Abstract

Large language model (LLM)-based conversational agents (CAs), with their advanced generative capabilities and human-like conversational interfaces, can serve as reading partners for children during dialogic reading and have shown promise in enhancing children's comprehension and conversational skills. However, there is limited research on the efficacy of LLM-based bilingual CAs in children's language acquisition in English as a Foreign Language (EFL) contexts. This randomized controlled trial study investigated the effectiveness of LLM-powered CAs compared with traditional parent–child shared reading in promoting engagement and improving learning outcomes among children with EFL. An interactive e-book featuring a LLM-powered CA was developed to engage children in dialogic reading through questioning and scaffolding. Sixty-seven children, aged 5 to 8, were randomly assigned to either an experimental (AI-led) group or a control (parent-led) group. The study found that children in the experimental group outperformed the control group in reading comprehension, with comparable benefits in vocabulary acquisition and story retelling, both immediately and in delayed tests. In the meantime, this study unpacks children's different engagement patterns when reading with the CA versus reading with their parents. Children reading with the CA demonstrated higher

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.
© 2025 The Author(s). *British Journal of Educational Technology* published by John Wiley & Sons Ltd on behalf of British Educational Research Association.

behavioural engagement and visual attention, while those in the parent-led group showed greater affective engagement and narrative-relevant vocalizations. The findings highlighted insights into the potential of LLM-powered CAs in children's language acquisition and suggested key design implications for developing better CAs for children from multilingual backgrounds.

KEYWORDS
artificial intelligence, children, conversational agent, interactive media, language education, large language model

INTRODUCTION

In today's rapidly globalizing world, the number of English Language Learners (ELLs) is growing at an unprecedented rate. As of 2020, more than 600 million individuals are learning English as a foreign language (EFL). English remains the most sought-after second language, serving as a global lingua franca and providing access to broader communication, educational and professional opportunities.

However, in the EFL context, many children face unique barriers. Since English is rarely used as the primary language for daily communication in their immediate environments or school systems, exposure to the language is often minimal. Therefore, target language exposure at home can significantly and directly impact children's language learning outcomes (Chow et al., 2010). A substantial body of literature has documented that EFL-relevant home literacy practices (HLP), such as exposure to English literacy materials, shared reading and direct instruction in English, play a crucial role in children's foreign language development (Dong & Chow, 2022). One of the most effective literacy activities is story reading, which can enhance children's language proficiency while also motivating them to engage in language learning (Caesar & Nelson, 2014; DesJardin et al., 2017).

Dialogic reading, an evidence-based shared reading practice, typically involves an adult asking children story-related questions, giving them feedback and enriching their verbalizations. It is closely related to various aspects of children's native and foreign language development, including vocabulary learning (Hargrave & Sénéchal, 2000), phonological awareness (Chow et al., 2010) and literacy (Fitton et al., 2018). Despite its benefits, dialogic reading has been especially challenging for families with children learning English as a foreign language due to parents' limited English skills, inadequate conversational skills and low inclination to read in English.

Recently, with the advancement of natural language processing (NLP) and artificial intelligence (AI), conversational agents (CAs) have demonstrated significant potential as reading partners for children. Emergent research has shown that intelligent systems with a conversational interface can provide children with additional language learning opportunities (Xu, 2023). These conversational interfaces were either deployed on voice assistants (e.g., Google Home, Amazon Echo) (Xu & Warschauer, 2019; Xu, Wang, et al., 2021), mobile applications (Lin et al., 2022; Liu et al., 2022; Xu, Branham, et al., 2021; Zhang et al., 2022) and children's television programmes (Xu, Vigil, et al., 2022). However, almost all of them targeted children's native language and literacy. There is a research gap in designing CAs for children's language learning in the EFL context.

14678535, 2025, 5, Downloaded from https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13615, Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

14678535, 2025, 5, Downloaded from https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13615, Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Practitioner notes

What is already known about this topic

- Dialogic reading enhances children's language and literacy development.
- Conversational agents (CAs) have shown potential as conversation partners in children's reading, replicating some benefits of adult-led dialogic reading.
- Children with EFL often face challenges in language acquisition due to limited exposure, and bilingual approaches are critical for bridging gaps in their first and target languages.
- Well-designed e-book applications offer interactive features, including animations, narration and embedded prompts, that support comprehension, vocabulary learning and engagement in young readers.

What this paper adds

- Demonstrates that a bilingual CA embedded in an interactive e-book can enhance children's reading comprehension and vocabulary comparably to parent-led reading in the EFL setting.
- Highlights distinct engagement patterns between children reading with CAs versus with parents in an e-book reading context.
- Provides evidence for the effectiveness of LLM-powered CAs in EFL contexts, addressing a research gap in bilingual conversational agents.

Implications for practice and/or policy

- Bilingual CAs can serve as effective supplementary learning tools for children with EFL, particularly in contexts where parental support is limited.
- The design of educational CAs should incorporate multimodal scaffolding, emotional responsiveness and bilingual features to optimise engagement and learning outcomes.
- Policymakers and educators can leverage AI-based tools to create accessible and scalable language learning solutions for diverse linguistic and cultural settings.

Recent advancements in large language models (LLMs) have opened up a new avenue for designing conversational agents capable of being bilingual conversational partners, an area that is currently underexplored. Besides, many children's language learning experiences have shifted towards interactive media, given the built-in multimedia features' affordances towards language learning. Incorporating CAs with multimodal features—such as text, audio and visuals—can facilitate meaning-making in the target language and help reduce communication breakdowns for English language learners (Jeon et al., 2023; Xiao et al., 2024). Additionally, CA can be particularly beneficial for children's language learning when it aligns with their cultural and linguistic practices (Xu et al., 2023). Therefore, this study proposed the design of an e-book application with an embedded bilingual CA as children's dialogic reading partners for children with EFL. It seeks to explore whether such a design has the potential to provide children with comparable benefits of parent-led dialogic reading. This research aimed to provide in-depth design implications for incorporating CAs

into multimodal learning environments, as well as to offer evidence of the effectiveness of CA-enhanced learning media in children's language development.

THEORETICAL FOUNDATIONS OF INTERACTIVE MULTIMEDIA LEARNING: COGNITIVE, AFFECTIVE AND SOCIAL PERSPECTIVES

With the advancement and widespread use of information technology, multimedia learning environments have transformed how children access information and acquire language. Mayer's Cognitive Theory of Multimedia Learning (CTML, 2002, 2003) posits that learning is most effective when information is presented through both verbal (eg, text and speech) and visual (e.g., images and animations) modalities, as the human mind processes these channels separately but integrates them for deeper comprehension and retention. Optimal learning occurs when multimedia materials minimize cognitive overload, facilitate active cognitive processing (i.e., selection, organization and integration of information) and align with the brain's natural cognitive architecture. However, effective and deep learning in multimedia environments relies not only on the presentation of information but also on active interaction and engagement with the learning content. Active learning involves knowledge application, feedback reception, meaningful dialogue and test of conceptualization (Cairncross & Mannion, 2001; Gilakjani, 2012; Mayer, 1997).

Building upon CTML, Moreno and Mayer's (2007) Cognitive-Affective Theory of Learning with Media (CATLM) highlights the importance of interactivity, multimodal engagement and guided instruction for promoting behavioural engagement and active knowledge construction (Markus, 1990). The theory also underscores that affective factors, such as motivation and engagement, influence cognitive processing and knowledge retention. Consistent with CATLM's assumptions, Moreno and Mayer (2005, 2007) argued that guided learning through a pedagogical agent can enhance students' active cognitive processing and generative processing, as well as sustain cognitive engagement and motivation, which are important mediators of learning (Pintrich, 2003). The agent provides explanatory feedback, which is not just evaluative but includes principle-based explanations—why an answer is right or wrong. Additionally, by guiding learners who might otherwise be overwhelmed, agents prevent affective disengagement that results from frustration or confusion caused by potential cognitive overload. This guided approach aligns with social constructivist perspectives (Vygotsky, 1978), reinforcing that children's language learning is an active and dynamic process, driven by scaffolded interaction and cognitive engagement.

Additionally, the integration of intelligent agents in multimedia learning environments is further supported by Schneider et al.'s (2022) Cognitive-Affective-Social Theory of Learning in Digital Environments (CASTLE), a recent theoretical extension to CATLM that highlights the social dimensions of multimedia learning. It posits that learners perceive interactions with digital agents as social experiences, triggering learners' social schemata in long-term memory that promote para-social, emotional and metacognitive engagement, ultimately enhancing learning and retention. According to social agency theory, learners engage more deeply when digital learning environments incorporate human-like social cues (Mayer et al., 2003), such as voice (Atkinson et al., 2005), personalized language (Rogers et al., 1977) and interactive dialogue (Moreno & Mayer, 2004). These cues elicit social presence and activate social learning mechanisms, enhance motivation and promote deeper engagement (Seitchik et al., 2017). Such cues can promote deeper engagement by eliciting social facilitation, though they may also induce social inhibition if the social presence becomes distracting or anxiety-provoking (Seitchik et al., 2017; Spatola et al., 2019). These theories collectively support the rationale of integrating a conversational agent into children's e-books. Also,

14678535, 2025, 5, Downloaded from https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13615, Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

given that social presence, affective processes and engagement are closely intercorrelated with each other (Beege et al., 2023), understanding how CAs-enhanced learning environments influence these factors is critical to maximizing language learning outcomes.

LITERATURE REVIEW

Dialogic reading and children's language learning

Dialogic reading is an evidence-based practice that promotes children's language development by actively engaging them in guided conversations during storybook reading. It involves structured questioning techniques used by parents or educators to encourage children to engage in storytelling while providing feedback in various formats, such as correction, modelling and elaboration (Whitehurst et al., 1988). Two core principles of dialogic reading are known as 'CROWD' and 'PEER' (Whitehurst & Lonigan, 1998; Whitehurst et al., 1988). The PEER technique encourages children to participate in storytelling by prompting them, evaluating their responses, expanding on their ideas and repeating prompts to reinforce understanding. CROWD questions—comprising completion, recall, open-ended, wh-questions and personal distancing prompts—are designed to elicit language production and deeper engagement.

Extensive research has demonstrated that dialogic reading benefits children's language development, not only by facilitating immediate learning but also by supporting long-term outcomes, such as story comprehension, vocabulary acquisition, and both expressive and receptive language skills (Flack et al., 2018; Mol et al., 2008; Zevenbergen & Whitehurst, 2003). For instance, Lever and Sénéchal (2011) found that dialogic reading with parents improved children's expressive vocabulary and comprehension, as evidenced by greater accuracy and complexity in story retelling. Simsek and Erdogan (2015) reported higher gains in both receptive and productive vocabulary among children who engaged in dialogic reading.

Furthermore, the literature underscores that the dialogic reading enhances learning benefits by increasing children's engagement during the reading process (Whitehurst et al., 1988). Reading engagement is conceptualized as a multidimensional construct that maps behavioural, cognitive, and affective dimensions (Guthrie & Klauda, 2014), and it is believed to be associated with language and literacy learning outcomes in dialogic reading (Pillinger & Vardy, 2022). Behavioural engagement refers to children's attention and persistence in tasks, typically measured by children's visual attention toward reading materials (Fredricks et al., 2016; O'Toole & Kannass, 2018; Reeve et al., 2020). Affective engagement involves emotional responses, such as enthusiasm and frustration, elicited by the task (Reich et al., 2019); cognitive engagement pertains to children's cognitive involvement and is typically indexed by children's on-task utterances (Moody et al., 2010; Munzer et al., 2019; Xu, Aubele, et al., 2022; Zhou & Yadav, 2017). Research indicates that children whose parents employ dialogic strategies are more likely to demonstrate greater cognitive engagement, as evidenced by a more diverse vocabulary (Troseth et al., 2020) and content-related verbalizations (Munzer et al., 2019). Similarly, greater visual attention during dialogic reading corresponds with better word-learning outcomes (Strouse & Ganea, 2017).

Although dialogic reading has been less widely adopted in the EFL learning context, evidence suggests that it holds significant promise. A substantial body of literature (for reviews, see Fitton et al., 2018; Kennedy & McLoughlin, 2023) has confirmed that dialogic reading benefits children learning English as a second language, particularly in enhancing their oral language skills and literacy outcomes. Studies in the EFL context have similarly shown improvements in word reading, phonological awareness and overall language engagement (Chow et al., 2008). However, children with EFL often face greater challenges in vocabulary

14678535, 2025, 5, Downloaded from https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13615, Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

development, which may stem from less frequent or less sophisticated parent–child interactions in English. Parents with EFL may struggle with limited English proficiency, making it difficult to engage their children in rich verbal interactions during reading (Chow et al., 2010). Additionally, work demands and time constraints may further restrict parents' ability to engage in consistent dialogic reading practices with their children.

Conversational agent as children's dialogic reading partner

Conversational agents (CAs) refer to intelligent systems that are capable of conducting dialogue with human beings. Children nowadays interact with conversational agents daily, including voice assistants (e.g., Apple Siri, Amazon Echo and Google Assistant) (Catania et al., 2021), all kinds of learning applications with conversational interfaces (Xu, Branham, et al., 2021) and social robots (Ali Mehenni et al., 2020).

With the rapid advancement of artificial intelligence, CAs have shown great potential to serve as children's conversational partners in dialogic reading. Recent research endeavours have designed CAs as children's dialogic reading partners, yielding promising results. For example, Xu, Vigil, et al. (2022) investigated children's dialogic reading with pre-trained conversational agents embedded in voice assistants (e.g., Google Home Mini). They found that this experience benefited children's reading comprehension as much as reading with an adult. In another study, Xu, Vigil, et al. (2022) found that CAs can replicate the benefits of improving children's reading comprehension with a human by enhancing behavioural engagement, specifically increasing narrative-relevant vocalization and reducing irrelevant talk. Structural equation modelling (SEM) analysis further confirmed that engagement serves as a crucial mechanism through which dialogic reading enhances children's story comprehension, regardless of whether the reading partner is a human or a CA. However, to maintain internal validity, both studies utilized highly scripted child–adult conversations, which may not reflect the natural, spontaneous interactions that typically occur during parent–child reading, thus limiting our understanding of how CAs might perform in real-world reading contexts.

Cheng et al. (2024) designed an embodied CA for reading with children from mainland China, showing improvements in reading comprehension and vocabulary acquisition across different age groups. Similarly, Liu et al. (2022) developed a tablet-based, disembodied CA that engaged children in on-screen ‘book talks’ after English reading sessions, which sustained children's interest in reading over 6 weeks.

Despite these promising findings, most of the current research focuses on monolingual CAs designed for children's first-language development. The potential for CAs to effectively support the language development of children with EFL remains underexplored.

Children's e-book reading and conversational agent

In recent years, children's language learning experience has shifted towards interactive media. E-books have emerged as a popular learning medium among children, offering interactive features, providing multimodal support for sensemaking in another language and engaging children in ways traditional print books cannot. Research has demonstrated the benefits of using well-designed e-book reading applications in enhancing children's phonological awareness (Chera & Wood, 2003), understanding and recalling complex new words (Smeets & Bus, 2015), reading comprehension and retelling (Bus et al., 2019; Takacs et al., 2015; Yang et al., 2022), emergent literacy skills (Korat & Shamir, 2007), and enhance learning engagement (Richter & Courage, 2017; Zhou & Yadav, 2017).

14678535, 2025, 5, Downloaded from https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13615, Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

E-books for children often come equipped with interactive features, such as narration, built-in dictionaries, animation and prompts, which can contribute to an enhanced reading engagement experience and improved learning outcomes (Moody et al., 2010). For example, the audio narration and extratextual animated cues for clickable vocabulary enabled children to make sense of vocabulary (Kucirkova & Littleton, 2016; Smeets & Bus, 2015). Built-in discussion prompts were found to be helpful for children's comprehension, yet children reading independently are more likely to skip prompts, which minimizes their potential benefits (Dore et al., 2018). Furthermore, a recent systematic review paper indicated that e-books with enhancements fostered higher overall level of behavioural, cognitive and affective engagement compared with print format. Content-relevant enhancement, especially when applied to children's comprehension, appears to be beneficial by increasing behavioural engagement, which in turn enhances attention and persistence (Clinton-Lisell et al., 2024).

The advancement of conversational AI has spurred research into incorporating conversational agents (CAs) in children's e-books. For example, Xu et al. (2023) embedded a bilingual CA in an e-book story to support Spanish-English parent–child shared reading, where the CA asked questions directed at both children and parents to enhance joint engagement. Similarly, Zhang et al. (2022) developed an e-book with a CA that posed questions during the story and provided one-turn feedback. In both cases, children and parents reported enhanced comprehension and vocabulary gains. However, children often experienced communication challenges or even breakdowns due to limited language skills, and the rule-based nature of the CA's natural language understanding (Du et al., 2021).

For children learning English as an additional language, the challenge of accented speech can further complicate interactions with the CA. Additionally, in EFL contexts, children's English speaking anxiety may reduce their willingness to communicate and to engage with the CA (Ayedoun et al., 2019). Embedding CAs within a multimodal learning environment may support EFL learners by enriching sensemaking and reducing dependence on language-based communication alone, thereby alleviating children's foreign language learning anxiety (Xiao et al., 2024). Children with lower vocabulary knowledge may also rely more on multimodal cues, which helps them interact more effectively with CAs and enhances learning outcomes (Cheng et al., 2024). Therefore, integrating CAs into interactive e-book reading platforms as reading companions may be a promising approach to enriching the language experience of children with EFL and enhancing their engagement with learning content.

Bilingual CA as conversation partner for children with EFL

For language learning purposes, CAs can be effective when it is designed in line with children's linguistic and cultural practices (Xu et al., 2023). In the EFL context, code switching and translation are common practices adopted by parents during dialogic reading with their children (Yang et al., 2022) and used as a discourse strategy in the classroom for effective language learning (Bhatti et al., 2018; Qian et al., 2009). For EFL learners, bilingual CAs serve as an ideal conversation partner by bridging the gap between their first language and the target language. These systems can provide scaffolding in the learners' first language when necessary, ensuring comprehension and self-efficacy, while progressively encouraging more interaction in the target language (Inbar-Lourie, 2010). Bilingual support can be particularly beneficial for younger learners, whose cognitive and linguistic skills are still developing. Research has shown that such dual-language approaches foster not only linguistic proficiency but also cognitive flexibility and intercultural awareness (García & Wei, 2015).

14678535, 2025, 5, Downloaded from https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13615, Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Traditional rule-based systems required extensive manual programming and lacked flexibility, limiting their ability to respond dynamically to varied learner inputs or support natural dialogue flow (Divekar et al., 2021). With LLM multilingual language understanding and recognition, bilingual conversational agents can now dynamically engage in natural and context-sensitive interactions with young ELLs (Lai et al., 2023). With the recent advancements in AI, LLM-powered CAs are expected to engage in highly flexible and adaptive conversation with children during dialogic reading. The research gap offers a valuable opportunity to explore how CAs can be better designed to enhance language development of children with EFL.

Research questions

Building on theoretical distinctions between parent-led and AI-guided dialogic reading, each approach capitalizes on different strengths. In parent-led reading, adults can adaptively calibrate contingent and emotionally resonant scaffolding tailored to the child's zone of proximal development (Vygotsky, 1978; Whitehurst et al., 1988). Parents' familiarity with the child's interests and developmental level often facilitates sustained engagement (Bus et al., 1995). However, time constraints and variability in language proficiency may reduce the consistency or richness of these interactions (Rodríguez & Tamis-LeMonda, 2011).

By contrast, a conversational agent (CA) embedded within an e-book can offer systematic knowledge support on specific language targets, maintaining high fidelity to dialogic reading principles without the potential drift or distraction that can arise in human-led interactions (Xu, Aubele, et al., 2022). Its 'always-on' availability, limitless patience and novelty appeal may enhance motivation, particularly for digital natives (Kory-Westlund & Breazeal, 2019). However, because the CA lacks the deep socioemotional bonding and shared contextual knowledge characteristic of human interactions, it may foster engagement through different mechanisms.

Taken together, this study seeks to evaluate the effectiveness of an interactive e-book story application with an embedded bilingual conversational agent powered by LLM. By examining both learning outcomes and engagement differences between CA-led and parent-led reading experiences, this study aims to provide empirical insights into the feasibility of AI as a reading partner in EFL contexts, particularly for families where parents may have limited availability and English language skills. Rather than simply testing a theoretical model, this research aims to evaluate the real-world feasibility and effectiveness of bilingual CAs in the EFL learning contexts, providing insights that can inform the design and implementation of AI-driven language learning interventions. The following research questions guide the evaluation of this application:

1. How does reading with a CA-embedded e-book impact children's English vocabulary, comprehension, and retelling? How does it compare to reading the e-book with a parent?
2. How does children's engagement level differ when reading with a CA versus reading with a parent?
3. To what extent does children's engagement correlate with their English vocabulary, comprehension, and retelling?

DESIGN OF STORIO: THE E-BOOK SYSTEM WITH CA

Storio is an e-book enhanced by a bilingual conversational agent (CA) named Mia, designed to support dialogic reading for young learners. The design is built upon extensive research

14678535, 2025, 5, Downloaded from https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13615, Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

into dialogic reading techniques (Zevenbergen & Whitehurst, 2003) and leverages the capabilities of LLMs. The system's refinement went through multiple iterations, including usability study and field testing involving a total of 24 child–parent pairs, as well as a think-aloud study and heuristic evaluation with two experts in learning design and language education. Based on the feedback from these sessions and theories, the design followed key principles aimed at: sustaining children's attention within the story, incorporating multimodal elements (eg, animations, illustrations) to support affective and cognitive engagement, and mirroring dialogic bilingual reading practices through the agent's interactions.

Mia, a friendly yellow cartoon character powered by GPT-4, interacts with children through story-related questions, scaffolding and multimodal cues, such as animations and illustrations to sustain attention and support comprehension. Mia narrated The Story of an Orange Oakleaf developed for another study (Yang et al., 2022). With a total of 451 English words, the story unfolds the tale of an oakleaf butterfly who, initially dismayed by its dry-leaf appearance, soon comes to realize that this very characteristic is the key to rescuing its friends from a sinister black bird. The Automated Readability Index (Smith & Senter, 1967) and Flesch–Kincaid Grade Level (Flesch, 2007) indicated that the story's text readability is 'easy' and it is suitable for children at the lower elementary level (kindergarten to 3rd grade).

At interaction moments, a flashing question mark prompts children to click on Mia, triggering a chat box where they can respond by speaking their preferred language. Guided by the dialogic reading technique (Zevenbergen & Whitehurst, 2003), Mia poses 10 story-related questions of three types: (1) wh-questions (eg, 'Why did Little Oak cry again?'), (2) recall questions to reinforce storyline comprehension and (3) distancing prompts that connect the story to children's personal experiences (eg, 'Have you ever had tears of joy?').

To provide adaptive scaffolding, Mia tailors its responses based on children's input, powered by few-shot learning (see Figure 1 for details) (Perez et al., 2021; Wei et al., 2021). The 100 scaffolding examples used to guide the LLM's responses were curated by two language education experts, who collected frequent child responses from field testing with 18 children. Child–AI interactions were then analysed and incorporated into the system, refining Mia's ability to deliver personalized and contextually appropriate feedback. Mia is further instructed to deliver concise, 25-word corrective feedback alongside scaffolding to assist the child's response to specific questions. Additionally, the CA was prompted to take on the role of posing questions to guide children's responses and offer constructive feedback

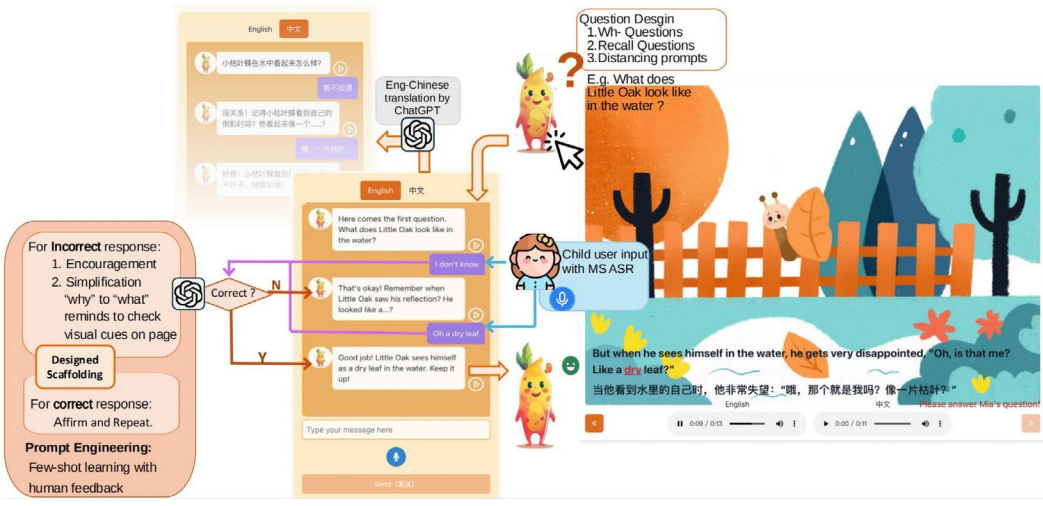


FIGURE 1 Storio system and interface design.

14678535, 2025, 5, Downloaded from https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13615, Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

responsive to children's developmental and cognitive levels. Mia is bilingual in Chinese and English, allowing children to interact in either language or a mix of both. Chinese translations were added beneath each sentence in the e-book to facilitate bilingual comprehension. Additionally, a multimedia dictionary for key vocabulary allows children to access pronunciation and visual aids with a simple click. In consideration of data privacy, all testing and implementation stages were conducted with data stored locally. Prompts are processed via the GPT API to ensure data security and prevent leakage, addressing concerns about potential data sharing with third parties (Wu et al., 2024). Additionally, all student sensitive demographic information (eg, names, contact), is excluded from the prompts to further safeguard privacy.

To further ensure the system's usability and credibility for intervention, the two educators reviewed the final implementation and design. We also conducted a pilot test with six children. In addition to administering knowledge tests, we gathered qualitative feedback from both children and their parents and analysed the chat logs. Feedback indicated that the system was easy to use and that the bilingual conversational agent helped support language learning and comprehension of the story.

METHOD

Experiment design and procedures

This study features a randomized controlled trial with repeated measures (see Figure 2). Each child participated in a single reading session, lasting approximately 20–40 min, depending on the child's pace. Children participants ($N = 67$) were randomly assigned to (1) the experimental group (CA group), where they ($N = 37$) read an e-book story with Mia, the CA, on touchscreen tablets, and (2) the control group ($N = 30$) (Parent group), where they read the same e-book story using touchscreen tablets but with their parents. In the CA group, children completed a tutorial on e-book navigation and practised interacting with Mia for 5 min before reading the e-book. This was to ensure that children mastered the basic skills to interact with CA and the system. During the reading session, Mia narrated the story and posed dialogic questions at key moments. This narration ensured that the content was accessible to all children, regardless of their reading ability. Children responded verbally in English, Chinese, or a mix of both. Mia provided scaffolded feedback tailored to their answers, offering up to three hints (See Figure 1). Mia can communicate in Chinese, English or a mix of both.

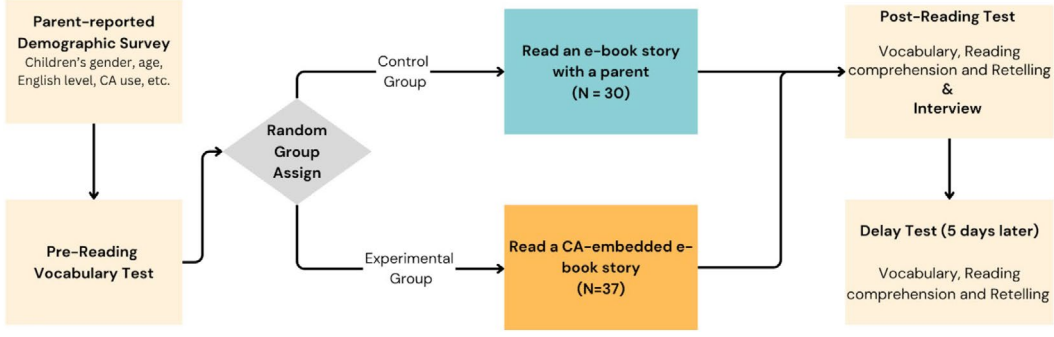


FIGURE 2 RCT workflow.

14678535, 2025, 5, Downloaded from https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13615, Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

In the Parent group, before reading, parents received a tutorial on using embedded prompt cues in the e-book, which reminded them to ask the same questions Mia posed at the same time point with the CA group. At the same point where Mia asked about Little Oak's appearance, parents were prompted: 'Ask your child: 'What does Little Oak look like in the water?'. The initial questions asked by both Mia and the parents were in English. Parents were encouraged to follow up on their child's responses, but their interactions varied. Regarding language use, parents were instructed to interact naturally in their preferred language. In practice, they primarily used a mix of simple English with occasional Chinese to support comprehension.

Before the intervention, parents were asked to complete a demographic survey. Children underwent a vocabulary test before the reading session, immediately after, and 5 days later, along with a Story Comprehension Test and Story Retelling Test. Right after the reading session, children in both groups were interviewed about their reading experience, while parents, if present, were interviewed about their impressions of the design and children's learning experience. Additionally, the CA group was interviewed about their perception of Mia, such as, their conversation experience with Mia and design suggestions for CA. All the reading sessions and assessments took place at children's homes, were administered remotely via Zoom, and were transcribed for further analysis. Researchers took detailed field notes for each reading session, specifically focusing on child–AI and child–parent interactions. The system automatically logged all child–CA interactions, including chat transcripts, response times and clickstream data.

Measures

Story comprehension test

Adapted from (Yang et al., 2022), this assessment consisted of seven question sets to evaluate children's recall and inference of story content. Three sets included a yes/no question followed by an open-ended question (eg, 'Did Little Oak use his camouflage to help his friends? How did he do it?'), while four sets included only an open-ended question. Yes/no questions were scored 1 point for correct answers and 0 for incorrect or no responses. Open-ended questions were scored on a 0–3 scale, with a maximum test score of 24 points. The test was administered immediately after the reading session and again 5 days later, with item order shuffled to reduce bias. Questions were read in both English and Chinese, and children could respond in either language. Two trained coders independently scored children's responses. Double-scoring yielded over 95% interrater reliability, and the test also demonstrated strong internal consistency (Cronbach's $\alpha = 0.91$).

English story vocabulary test

This test was used to assess children's word knowledge of the story-related vocabulary after the intervention. The vocabulary assessment consisted of two parts: **Receptive Vocabulary** and **Expressive Vocabulary**, each containing 25 items, for a total of 50 questions (Author, Year) (Appendix A). The test was administered before the reading session, immediately after, and 5 days later, with a shuffled item order in the delayed test. The target words, which include nouns, verbs and adjectives that could be distinctly represented through images, were drawn directly from the story and generally ranged from one to three syllables. In the Receptive Vocabulary section, children heard an English

14678535, 2025, 5, Downloaded from https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13615, Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

word or phrase from the story and selected the matching image from four options. The experimenter asked the children, 'Please tell me which of the following pictures is [target word]' consistently for all the words, using the child's preferred language as determined through a brief check at the start of the session. In the Expressive Vocabulary section, children viewed images featuring 25 story-related elements, and the experimenter pointed to an item while prompting the child to name it in English. Correct answers scored 1 point, while incorrect or missing responses scored 0. Instructions were provided in Chinese. The test yielded a Cronbach's Alpha of 0.98.

Story retelling test

This assessment aimed to evaluate the children's ability to recall and narrate the story (Yang et al., 2022). It was conducted both as an immediate posttest and as a delayed posttest. Using selected illustrations from the story as visual cues, children were encouraged to retell the entire narrative. The test administrator initiated the task with the prompt: 'Hi, I missed the story. Could you tell me what happened? Feel free to use the illustrations to help you remember'. If the child paused or struggled to continue, the experimenter gently prompted them with, 'And what happened next?'. All retellings were transcribed and analysed based on critical idea units, following the framework of Capotosto and Kim (2016). The scoring process involved double-scoring among three independent raters (Cohen's Kappa = 0.92).

Engagement

Children's engagement was coded from videos of their reading sessions by two researchers. Each video was segmented into 8-s segments and coded based on the framework from general book reading engagement (Guthrie & Klauda, 2014), e-book reading engagement (Reich et al., 2019) and CA-facilitated reading engagement (Xu, Aubele, et al., 2022). As a result, children's engagement was evaluated from a detailed investigation of their behaviours based on an itemized coding system and holistic scoring based on children's global engagement (details see Appendix A). As for the itemized coding system, children's engagement based on their visual attention to the story (Cohen's Kappa = 0.93), vocalization (i.e., narrative-irrelevant, narrative-relevant, device-related) (Cohen's Kappa = 0.95) and affect (i.e., negative, positive) (Cohen's Kappa = 0.91). Children's global engagement was coded based on a 5-point scale (ICC = 0.88) (Appendix B).

Participants

All the child–parent dyads were recruited using snowball sampling within the researchers' social network. Eligible participants were children aged 5 to 8 years whose native language was Chinese, who had prior experience using tablets and who did not have any diagnosed cognitive disabilities. This study aimed at recruiting 50 children based on a power analysis to detect a medium effect size of 0.3 (Cohen, 2013). Initially, 88 families from China ($N = 86$), United States ($N = 1$) and Canada ($N = 1$) signed up for the study.[1] However, the burden of parental involvement, especially for the Parent group, led to attrition, with dropout reasons, including scheduling, the demands of dialogic reading and the coordination of multiple assessments. Consequently, only 67 child–parent dyads were able to participate in all the required tests and sessions. Most of the child participants (19.1%) had previously learned basic English words, and used chatbots at least once (94%). They were randomly assigned

14678535, 2025, 5, Downloaded from https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13615, Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

to the experimental group ($N = 37$) and the control group ($N = 30$) (details see Figure 3). *T*-tests and chi-squared analyses indicated no significant demographic differences between the groups, confirming the comparability of the final samples (Appendix E).

Qualitative data analysis

During each session, researchers took detailed field notes to capture child–CA and child–parent interactions (eg, content of verbal responses and turn-taking). These notes were taken using a semi-structured protocol that focuses on behaviours and language use, such as behavioural cues (eg, gestures, posture and device manipulation), affective expressions (eg, smiles and frustration), interaction strategies (eg, self-repair) and other contextual factors (eg, voice recognition difficulty). These notes were timestamped and synchronized with video recordings (8 s per segment) for cross-validation. The first and second authors reviewed video recordings and systematically annotated instances of each behaviour using a coding scheme. The initial codes were informed by prior literature on child–computer interaction (Xu et al., 2023; Xu, Aubele, et al., 2022) and children–adult dialogic shared reading

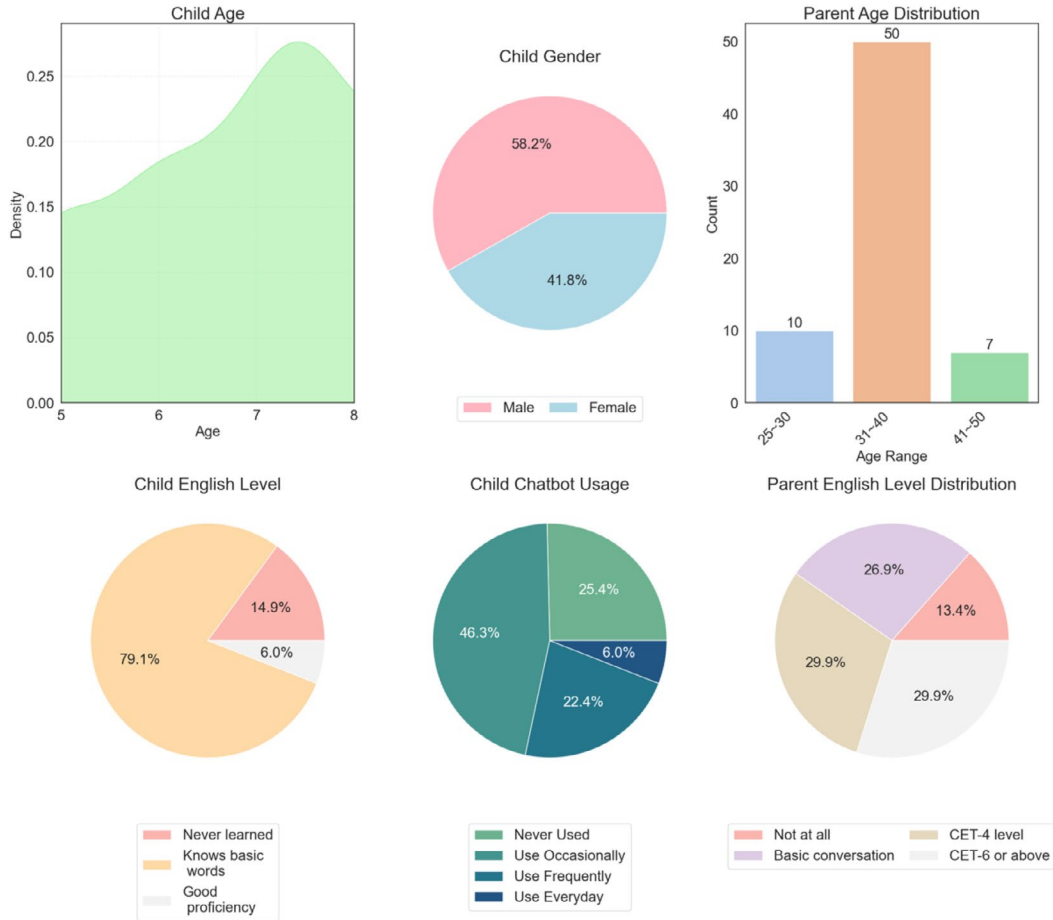


FIGURE 3 Demographic information of child–parent participants. CET-4 and CET-6 stand for 'College English Test Band 4' and 'Band 6'. Groupwise details and other relevant demographic can be seen from Appendix D.

14678535, 2025, 5, Downloaded from https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13615, Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

(Abdelhadi, 2020; Son & Tineo, 2016; Troseth et al., 2020), and were then iteratively refined and expanded based on patterns that emerged from the data during analysis. Discrepancies in the coding process were resolved through iterative discussions until consensus was achieved. To ensure robustness and interpretative validity, field notes were triangulated with chat logs and interviews. For instance, when a child struggled to answer the CA's question verbally but turned to an illustration for clarification, the chat logs revealed repeated attempts and clickstream data showing navigation back to the relevant image, while interview comments illuminated the child's or parent's perspective. Such convergence revealed multimodal repair strategies, integrating visual and verbal cues for sensemaking. These qualitative findings were then compared against quantitative engagement metrics to clarify how specific behaviours (eg, repeated speech attempts) influenced engagement and learning outcomes.

RESULTS

The impact of CA on children's e-book reading

Descriptively, the CA group performed better in comprehension and retelling, with 3.07 points and 2.27 (two idea units) higher, respectively. Both groups gained approximately the same amount of vocabulary after reading, with around five more receptive vocabularies and two more productive vocabularies (Figure 4). To assess between-group differences

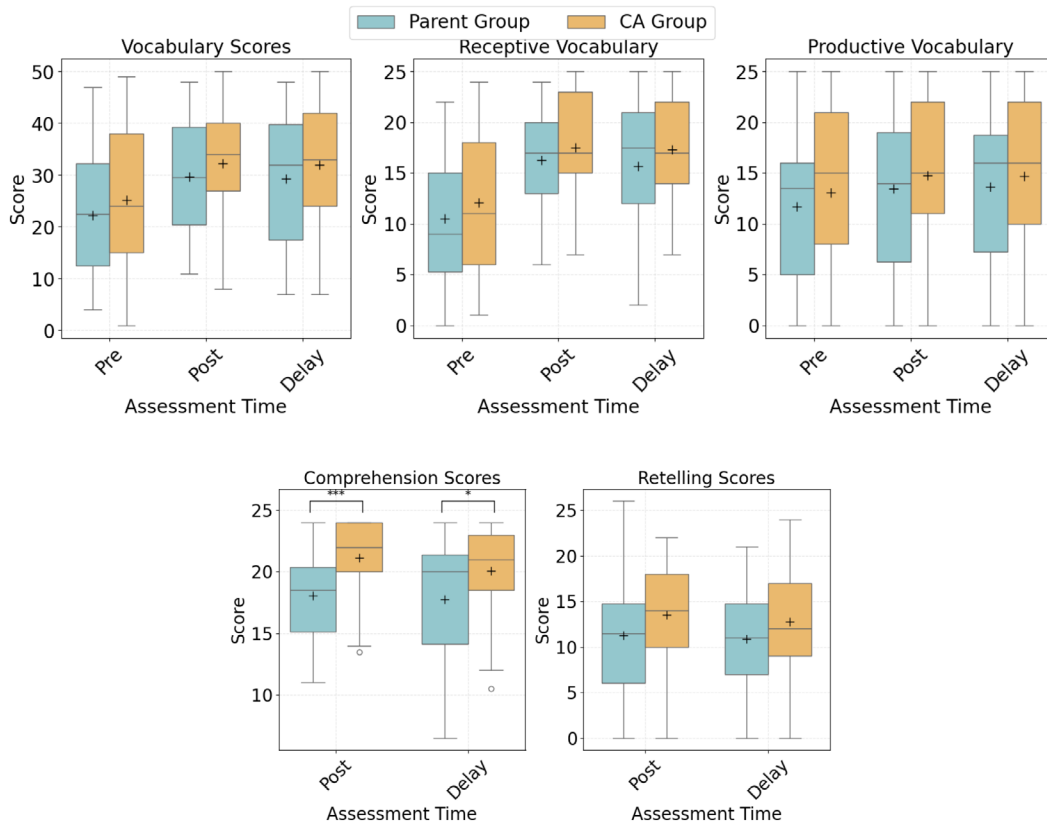


FIGURE 4 Group comparison for vocabulary, comprehension and retelling at different time point. * denotes $p < 0.05$, ** for $p < 0.01$, *** for $p < 0.001$. Numbers in parentheses are standard deviation.

14678535, 2025, 5, Downloaded from https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13615, Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

across measures, multiple Mann–Whitney *U*-tests were conducted (See Appendix C). The results revealed that only comprehension showed a statistically significant difference. At the posttest and delayed test, the CA group demonstrated better performance than the parent group ($U=266$, $p<0.001$). However, the difference between the two groups became less pronounced at the delayed test ($U=401$, $p=0.03$).

To further examine group differences and change over time, three separate linear mixed-effects models were used to analyze the three outcome variables, respectively, after controlling children's demographic variables (Figure 5). The results revealed a main effect of time on vocabulary. A within-group ANOVA captured a significant increase from pretest to posttest ($\beta=7.53$, $p<0.001$) with medium effect size (Cohen's $d=0.55$), and from the pretest to the delay test ($\beta=7.13$, $p<0.001$) with medium effect size (Cohen's $d=0.52$). However, there was no significant difference between the two groups ($\beta=1.0$, $p=0.7$), and the interaction effects for the posttest ($\beta=-0.44$, $p=0.942$) and delayed time point ($\beta=-0.27$, $p=0.942$) were not significant as well ($\beta=-0.13$, $p=0.82$). This indicated that both groups benefited similarly from the intervention regarding vocabulary acquisition.

Besides, there is a significant main effect of Group on comprehension scores, $\beta=2.91$ ($p<0.01$), indicating reading with CA leading to 2.91 *SD* increase in children's comprehension score. Although there is a slight decline in comprehension in the delay test, the difference over time was not statistically significant ($p=0.14$). No significant interaction between Time and Group ($\beta=-0.72$, $p=0.35$), with a marginal time effect ($U=406$, $p=0.06$).

For retelling, there was no significant group effect, with the CA group showing slightly higher retelling scores compared with the Parent Group, with a marginal time effect ($\beta=1.09$,

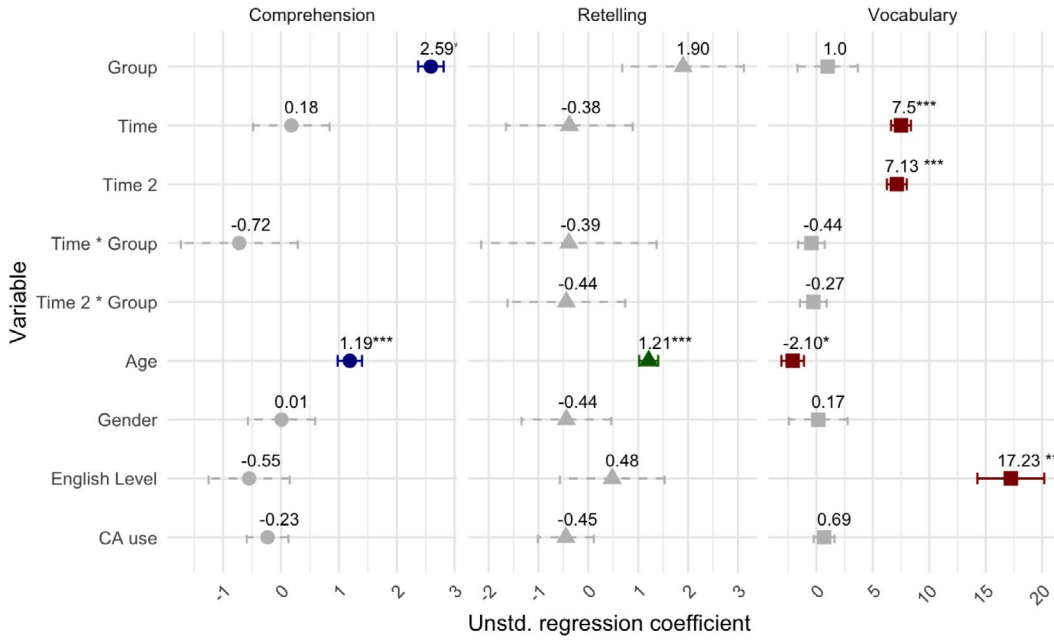


FIGURE 5 Fixed effects model for vocabulary, retelling and comprehension at each time point. * denotes $p<0.05$, ** for $p<0.01$, *** for $p<0.001$. Coefficients of significant variables were colour coded. Time: posttest to pretest; Time 2: delayed test to pretest. For vocabulary, a separate model using the pretest (time 0) as the reference group was used. Time * Group: if significant, these terms indicate that the effect of time on the outcome differs between groups.

14678535, 2025, 5, Downloaded from https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13615, Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

$p = 0.15$). There is no significant change from the posttest to the delay test ($t = -1.40$, $p = 0.15$), and the interaction between time and group was not significant ($\beta = -0.39$, $p = 0.2$).

To ensure the robustness of the results, a Bonferroni correction was applied to adjust for multiple comparisons across the three outcome domains, resulting in an adjusted alpha level of $0.05/3 \approx 0.017$. All primary findings remained consistent under this correction, including the significant vocabulary gains over time and the group effect on comprehension. In summary, reading with the CA significantly enhanced children's story comprehension, with a marginal improvement observed in retelling scores. Both groups showed vocabulary gains following the reading sessions, and all measures demonstrated good retention over time, as demonstrated in the delayed test.

Children's engagement patterns with the CA and parents

To unpack children's engagement patterns with the CA and parents, a series of Mann–Whitney *U*-tests (see Appendix D) were conducted to compare the engagement of the CA and Parent group regarding their vocalization, visual attention, affect and overall engagement. Across all time segments, children's global engagement reading with CA is higher than with parents. However, children reading with CA showed fewer vocalizations than those reading with a parent, in both narrative-relevant ($U = 767.0$, $p = 0.008$) and irrelevant vocalizations ($U = 755.00$, $p = 0.007$) (Figure 6). Furthermore, children reading with CA demonstrated significantly higher visual attention compared to those reading with their parents ($U = 240.0$, $p < 0.001$). As for the display of affect, children demonstrated more positive affect than those reading with their parents ($U = 733.0$, $p = 0.025$); however, this difference did not remain significant after applying a Bonferroni correction for multiple comparisons (adjusted $\alpha = 0.007$), while the other differences remained statistically significant. Descriptive statistics of each variable and Bonferroni-adjusted results can be found in Appendix D.

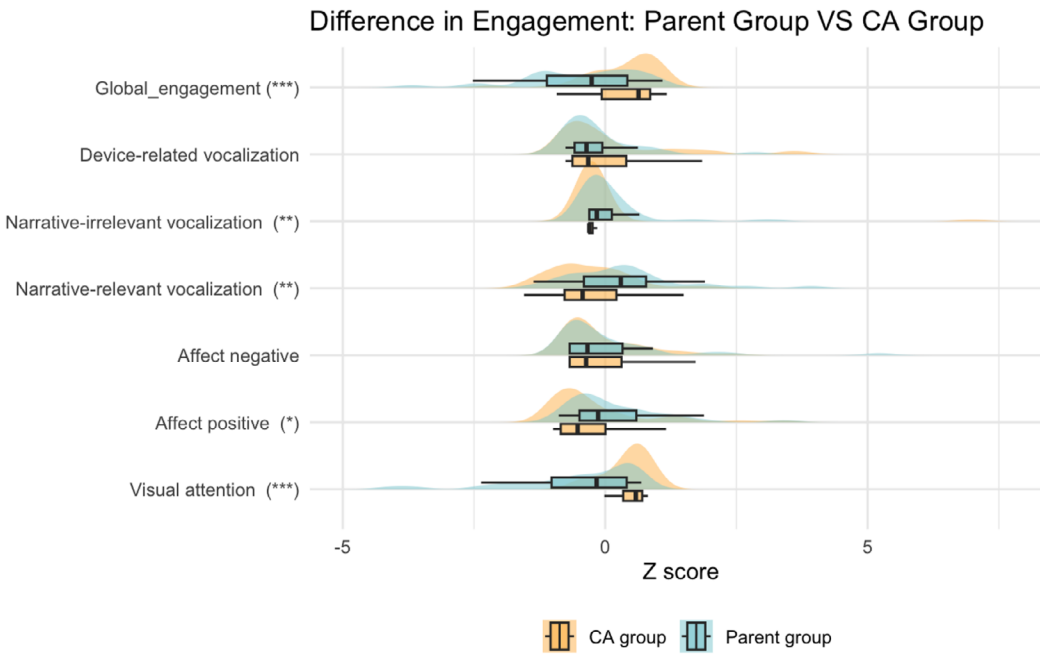


FIGURE 6 Group comparison regarding engagement. * denotes $p < 0.05$, ** for $p < 0.01$, *** for $p < 0.001$.

14678535, 2025, 5, Downloaded from https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13615, Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Qualitative observations during the reading sessions revealed distinct engagement patterns between the two groups. Children interacting with the CA showed higher global engagement, likely due to their curiosity about the new design and agency over the device. They maintained sustained visual attention on the e-book, often fixating on the story, the CA and interactive elements. They were generally quieter, with vocalizations primarily in response to the CA's prompts and occasionally commenting on the story and CA features. Their affective expressions included smiles and enthusiastic gestures when the CA acknowledged their input or provided positive animated feedback, as well as expressions of empathy for the story characters.

In contrast, children reading with their parents displayed more vocalizations, frequently commenting on the story and relating it to their own relevant experiences. Their visual attention shifted between the e-book, their parents and occasionally elsewhere. These children exhibited a broader range of emotions, including laughter, expressions of impatience for their parents to take over the device, surprise and affectionate gestures like hugging or leaning towards their parents. They sometimes laughed when looking at their parents. These behaviours likely reflect the emotional bond inherent in the parent–child relationship.

Association between children's engagement and learning outcomes

Pearson's correlation was used to investigate the association between children's engagement and learning outcomes in both groups and the CA group independently (Figure 7). For the full sample (comprising both groups), overall engagement showed a moderate to strong correlation with comprehension ($r = 0.49$, $p < 0.001$) and retelling ($r = 0.30$, $p < 0.05$). Similarly, visual attention was positively associated with comprehension ($r = 0.42$, $p < 0.001$) and retelling ($r = 0.25$, $p < 0.05$), highlighting the importance of maintaining focus on the e-book visuals in supporting comprehension and retelling of the story. Negative affect demonstrated a moderate negative relationship with retelling ($r = -0.27$, $p < 0.01$), indicating that feelings, such as frustration or disengagement can impair the ability of retelling the story. This pattern is also evident in the CA group, where children's negative affect exhibited a moderate negative relationship with retelling ($r = -0.38$, $p < 0.01$). In addition, the CA group exhibited unique patterns, particularly in device-related vocalizations, which showed a strong negative correlation with reading comprehension ($r = -0.46$, $p < 0.01$). A closer examination

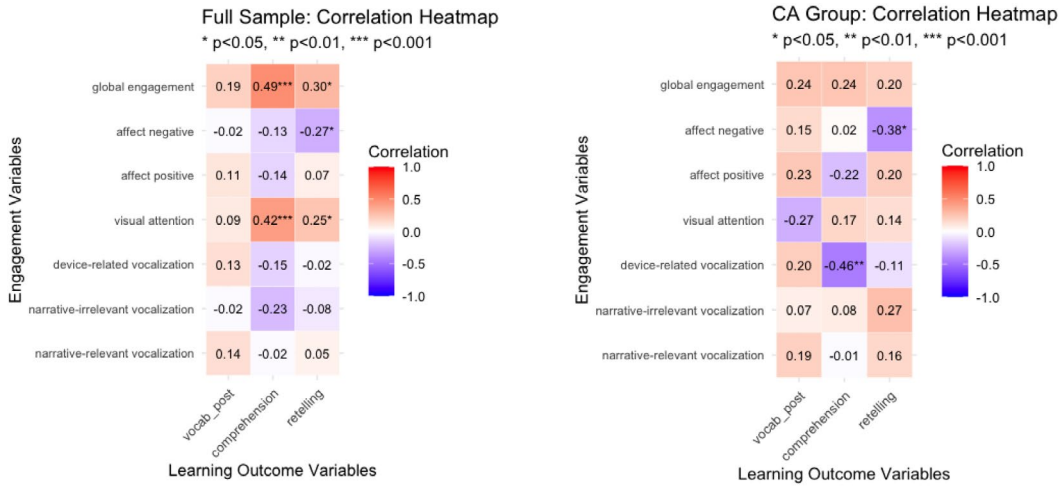


FIGURE 7 Correlation matrix. * denotes $p < 0.05$, ** for $p < 0.01$, *** for $p < 0.001$.

14678535, 2025, 5, Downloaded from https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13615, Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

of instances where children were highly visually engaged, made frequent device-related vocalizations, but showed low comprehension revealed that these children struggled with usability of the e-book, such as failure in voice recognition and slow content loading.

DISCUSSION AND CONCLUSION

This study presents a novel generative AI-powered CA integrated into children's e-book reading experience and evaluates its effectiveness in children's language development compared with parent-guided reading in the EFL context. This study is among the first to investigate how generative AI can be seamlessly integrated into children's daily media use and tailored to enrich their language learning experiences. Additionally, it is the first to examine the impact of bilingual CAs on children's bilingual literacy development in a multilingual context.

The learning benefits of bilingual CA in children's e-book reading

The study demonstrated that dialogic reading with CA or with parents demonstrated comparable benefits for children's vocabulary, comprehension, and retelling in post- and delayed tests in e-book reading. This exploratory RCT offers additional evidence that contributes to prior research. Specifically, it contrasts earlier studies suggesting that dialogic reading with an experimenter using the same CA script yielded similar outcomes in reading comprehension (Cheng et al., 2024; Xu, Wang, et al., 2021).

Reading with CA, children received consistent attention-regulating cues from the CA's scaffolding, functioning similarly to parental involvement. These cues helped regulate distractions commonly associated with interactive media (Dore et al., 2018; Yang et al., 2022), triggering greater visual attention and proactive engagement in the multimodal sensemaking process (Shepard-Carey, 2020; Takacs & Bus, 2018). Additionally, these cues potentially reduced competition for attention between parents and the e-book, which could otherwise hinder vocabulary learning (O'Toole & Kannass, 2018). Moreover, the CA addressed findings from other studies that reading e-books with parents led to better comprehension than when children read alone, as parents provided more social interaction (Dore et al., 2018). The social presence of CA embedded in the e-book facilitated reciprocal social interaction, triggering active cognitive processing and deliberate efforts, which distinguished it from traditional e-books and promoting in-depth engagement with the story (Hirsh-Pasek et al., 2015; Mayer, 2005; Schneider et al., 2022). This suggests that the CA can replicate some of the social and interactive benefits typically associated with parent–child reading sessions. Notably, part of the CA's advantage may also stem from the novelty of its design, which appears to enhance behavioural engagement, thus encouraging more active and effortful cognitive processing of digital content.

Another potential explanation for the greater comprehension in the CA group is that children appear to engage in a higher level of negotiation of meaning, an indication of high cognitive engagement in language learning tasks (Nakamura et al., 2021). Because current LLM-powered CAs may struggle with children's non-native accents, some children needed to repeat or rephrase their inputs, a process that may have reinforced recall of story content and vocabulary (Jeon, 2024). In line with multimedia theories, CA promoted self-explanation, active recall and meaning negotiation, which in turn supported deeper comprehension (Carver & Scheier, 2012; Mayer & Moreno, 2003).

In contrast, regarding the parent group, we speculate the progression of dialogue was primarily driven by parental mediation strategies (e.g., conversation turns, language use),

14678535, 2025, 5, Downloaded from https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13615, Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

which were rooted in their inherent beliefs about learning (Korat et al., 2013; Kucirkova & Flewitt, 2022). Parents with EFL often prioritize English vocabulary acquisition during storytelling (Chow et al., 2022; Yang et al., 2022) and may inadvertently overlook children's focus on comprehending the storyline. Last but not least, children reading with parents may have relied more passively on adult cues, resulting in fewer opportunities for self-initiated engagement and less accountability in maintaining interaction with the device itself. Future research could systematically compare parental scaffolding versus CA scaffolding to clarify these different learning pathways. Additionally, longitudinal studies would explore whether sustained CA interactions continue to facilitate engagement and meaningful language development over extended periods.

Engagement patterns in different reading conditions

This study also highlights the nuances of children's engagement patterns in different reading conditions. Overall, children demonstrated higher overall and more behavioural engagement in the CA group. Consistent with previous findings, children with EFL who read with the CA exhibited more narrative-relevant and fewer irrelevant vocalizations, the latter being potential indicators of distraction (Reich et al., 2019; Xu, Aubele, et al., 2022). However, reading with parents triggered more emotional engagement. Reading with parents fosters both physiological arousal and positive emotional experiences in children, which motivate sustained reading practices (Avelar et al., 2022).

When parents read to their children on e-books, parents appear more attuned to their children's learning styles, better able to address complex queries and adept at managing frustrations, advantages that are difficult to replicate through machines (Dore et al., 2018; Strouse et al., 2023). This could be due, at least in part, to the emotional bond and shared knowledge parents bring to the reading experience. Such highly contingent social interaction may help sustain children's engagement and facilitate long-term learning, which seems especially relevant for EFL learners experiencing speaking anxiety. In contrast, when CAs provided feedback on children's responses, the scaffolding, though designed with affective language, remained primarily cognitive in nature, lacking the emotional resonance and flexibility that characterize parental interactions and parent–child joint media engagement (Ewin et al., 2021). These observations address a gap in the current literature examining child–CA shared reading, where children read with an experimenter, which cannot replicate these special interaction dynamics.

Even if children showed less attention to e-books when reading with their parents, this may not necessarily undermine their overall learning outcomes, particularly in vocabulary acquisition. In line with the literature, e-books with thoughtful design have been found to yield higher levels of parental mediation, particularly in word explanation (Korat & Segal-Drori, 2016; Troseth et al., 2020). Therefore, children might redirect their attention to their parents, who played an active role in scaffolding the reading experience and helping them make sense of the story by listening. This shift in attention does not necessarily indicate disengagement but rather reflects a dynamic interaction where auditory, visual and social cues intersect to support children's comprehension (Neuman et al., 2020). Notably, we found a negative association between children's negative affect and retelling. Children experiencing negative emotions during reading may focus more on their emotional reactions rather than the sequence of events needed for retelling, thus resulting in worse retelling scores. Also, a negative correlation exists between vocalization with comprehension in the CA group. This may be partly due to the generally high comprehension scores within this group, where a few outliers with low comprehension scores may have a disproportionate effect on the correlation coefficient.

14678535, 2025, 5, Downloaded from https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13615, Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Designing better LLM-powered CA and learning media for children

This study sheds light on the implications of designing and implementing LLM-powered CA learning companions for children, particularly from EFL backgrounds.

Firstly, it is important to design context-sensitive CAs that are capable of incorporating multimodal cues from the learning environment to support children's sensemaking. Future conversational agents powered by LLMs could integrate context-sensitive prompt engineering to interact with the learning context dynamically. Such integration could effectively regulate children's attention and empower them to actively make sense of the digital learning environments, reducing their language learning anxiety and leading to better language learning outcomes.

Secondly, emotional-expressive and emotional-responsive CAs are desired for a better learning experience. Children require a higher level of affective support during technology-mediated learning (Neuman et al., 2020), especially when speaking a foreign language, which can increase cognitive and emotional demands. While the CA in our study provided encouragement through scaffolding, its tone was often perceived as too adult or robotic, diminishing its emotional resonance. Incorporating child-friendly voice modulation, empathetic phrasing, personalized praise, humour and context-aware responses can make interactions more engaging and effective. Voice customization, such as replicating a parent's voice, could further enhance emotional connection. Future designs might also establish a CA-parent collaboration system, enabling parents to provide personalized information to optimize the CA's interactions (Zhang et al., 2022).

Thirdly, designing CAs that incorporate diverse types of prompts that reflect the cultural and linguistic practices of children. In EFL context, repetitive practices are critical for effective language learning (Ismail, 2017). Children learning a foreign language benefit significantly from repetition in the target language, which helps reinforce learning and comprehension (Jeon et al., 2023). While LLM-powered CAs have shown promise in generating cognitively beneficial scaffolding pieces, such as code switching, these features alone are not enough. To better support children with EFL, CAs must incorporate prompts emphasizing repetition in vocabulary to create a more effective learning experience. The limited productive vocabulary outcomes in our study, along with interviews with parents, also highlighted this aspect.

Implications for educators and practitioners

In line with previous studies, these findings underscore the importance of using meaningful prompts to engage children effectively. Research indicates that children achieve lower comprehension outcomes when reading with a conversational agent lacking dialogic interaction (Cheng et al., 2024; Xu, Vigil, et al., 2022). Therefore, it is essential to integrate dialogic prompts and foster conversation-rich reading experiences, particularly in technology-enhanced environments. When guiding children in e-book reading, educators and parents should encourage children's active participation with the device rather than opting for a passive approach. For example, parents can pose thoughtful, open-ended questions and provide scaffolding that helps children leverage multimodal cues (eg, illustrations, animations or audio features) to support their sensemaking and maintain deeper engagement. Importantly, it should be noted that this study does not advocate replacing parents with AI. Instead, it explores a potential supplementary resource for situations where parental involvement in language and literacy is limited or constrained. In such contexts, a well-designed CA could offer an additional avenue for dialogic reading, helping children maintain or enhance their engagement and language practice. Finally, given the growing accessibility of GPT-based tools, we encourage researchers, educators and developers to collaborate on co-designing

14678535, 2025, 5, Downloaded from https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13615, Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

dialogic reading environments that are not only grounded in empirical evidence but also responsive to the real needs and experiences of children, families and practitioners.

Limitations and other future directions

Last but not least, we acknowledge the limitations of our study and propose directions for future research. First, the sample size may limit the generalisability of our findings. Additionally, the children were recruited through snowball sampling, which may have resulted in a sample skewed towards families with higher socioeconomic status (SES) and strong English reading skills. Future studies should aim to replicate this research with a larger and more diverse sample, encompassing children from various linguistic, cultural and SES backgrounds.

Second, although the bilingual CA system outperformed the control condition in some aspects, part of its advantage may stem from a novelty effect. Most children reported having only occasional or no prior experience with similar technology, so their initial excitement could have inflated engagement and short-term gains. Future research should investigate performance across multiple sessions and replicate the study with a longer intervention to determine whether the benefits endure once the novelty has worn off.

Third, although parents in the control group were trained to ask the same questions as those in the CA group, individual differences in parents' language proficiency, beliefs and interaction style may still have influenced children's learning experiences (Kucirkova & Flewitt, 2022). While this training helped minimise variability, it is possible that some parents were more effective in engaging their children in dialogic reading than others. Future research should account for these variations when comparing CA-guided reading and parent-led reading, while ensuring sufficient ecological validity, for instance, by offering more systematic parent training on dialogic reading, comparing subgroups of parents (eg, low vs. high proficiency or statistically controlling for interaction style). Such measures can help clarify whether and how variability in the parent-led group shapes children's learning outcomes.

Additionally, future studies might broaden both the engagement and linguistic measures used to understand children's experiences in CA-enhanced reading. In line with most emergent research in a similar context, this study focuses on vocabulary and comprehension as primary learning outcome variables. However, additional linguistic measures, such as morphological awareness, reading fluency or listening comprehension, could offer a richer perspective on how children with varying proficiency levels engage in interactions. Furthermore, although we relied on behavioural coding of engagement due to participants' young age, future research could also incorporate children's self-reported engagement or other indicators to supplement these objective measures. Integrating advanced tools, such as eye-tracking systems, would likewise yield more nuanced insights into how children focus their attention during CA-enhanced reading, thereby enriching our understanding of the complex interplay between language proficiency, engagement and learning outcomes.

Finally, the CA in our study primarily functioned as a parent-like companion in the storytelling process. Future research could explore diversifying the CA's roles or providing children with the ability to personalize their CA, as suggested by participants in our interviews. Lastly, we acknowledge the challenges current LLM-based CAs face in engaging young children and non-English speakers. This limitation allowed children to practice more on form, yet trading off their learning experience and attention on the meaning (Talamas et al., 1999). Future studies could build on this by incorporating advanced voice modulation techniques to create a more child-friendly and inclusive experience. While our findings highlight the

14678535, 2025, 5, Downloaded from https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13615, Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

potential of CAs for supporting children's language learning in the EFL setting, the evidence remains inconclusive. Nonetheless, we see an opportunity to transform children's language learning paradigms in the era of AI, opening pathways for innovative and adaptive educational media.

FUNDING INFORMATION
This research project did not receive any funding.

CONFLICT OF INTEREST STATEMENT
The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT
The data used in this study will not be shared to public.

ETHICS APPROVAL STATEMENT
This study was approved by the Institutional Review Board (IRB) of Penn State University. All procedures performed in this study involving human participants were conducted in accordance with the ethical standards of the institution and the 1964 Helsinki declaration and its later amendments.

ORCID
Feiwen Xiao https://orcid.org/0000-0002-7284-9527
Jiaju Lin https://orcid.org/0000-0001-6988-7568

ENDNOTE
[1] Despite the geographic locations, all the children come from EFL backgrounds. The family from the United States are visiting scholars who have just arrived in the United States for a week. The participants from Canada dropped out of our study and were not included in the final count.

REFERENCES
Abdelhadi, S. M. (2020). Responses to hotspots during parent–child shared reading of eBook. *International Journal of Child-Computer Interaction*, *23*, 100164.

Ali Mehenni, H., Kobylyanskaya, S., Vasilescu, I., & Devillers, L. (2020). Nudges with conversational agents and social robots: A first experiment with children at a primary school. In *Conversational dialogue systems for the next decade* (pp. 257–270). Springer Singapore.

Atkinson, R. K., Mayer, R. E., & Merrill, M. M. (2005). Fostering social agency in multimedia learning: Examining the impact of an animated agent's voice. *Contemporary Educational Psychology*, *30*(1), 117–139.

Avelar, D., Dore, R. A., Schwichtenberg, A. J., Roben, C. K., Hirsh-Pasek, K., & Golinkoff, R. M. (2022). Children and parents' physiological arousal and emotions during shared and independent e-book reading: A preliminary study. *International Journal of Child-Computer Interaction*, *33*, 100507.

Ayedoun, E., Hayashi, Y., & Seta, K. (2019). Adding communicative and affective strategies to an embodied conversational agent to enhance second language learners' willingness to communicate. *International Journal of Artificial Intelligence in Education*, *29*, 29–57.

Bai, S., Repetti, R., & Sperling, J. (2016). Children's expressions of positive emotion are sustained by smiling, touching, and playing with parents and siblings: A naturalistic observational study of family life. *Developmental Psychology*, *52*, 88–101. https://doi.org/10.1037/a0039854

Beege, M., Schroeder, N. L., Heidig, S., Rey, G. D., & Schneider, S. (2023). The instructor presence effect and its moderators in instructional video: A series of meta-analyses. *Educational Research Review*, *41*, 100564.

Bhatti, A., Shamsudin, S., & Said, S. B. M. (2018). Code-switching: A useful foreign language teaching tool in EFL classrooms. *English Language Teaching*, *11*(6), 93–101.

Bus, A. G., Sarı, B., & Takacs, Z. K. (2019). The promise of multimedia enhancement in children's digital storybooks. In J. E. Kim & B. Hassinger-Das (Eds.), *Reading in the digital age: Young children's experiences with e-books* (pp. 45–57). Springer. https://doi.org/10.1007/978-3-030-20077-0_4

14678535, 2025, 5, Downloaded from https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13615, Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Bus, A. G., Van IJzendoorn, M. H., & Pellegrini, A. D. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. *Review of Educational Research*, *65*(1), 1–21.

Caesar, L. G., & Nelson, N. W. (2014). Parental involvement in language and literacy acquisition: A bilingual journaling approach. *Child Language Teaching and Therapy*, *30*(3), 317–336.

Cairncross, S., & Mannion, M. (2001). Interactive multimedia and learning: Realizing the benefits. *Innovations in Education and Teaching International*, *38*(2), 156–164.

Capotosto, L., & Kim, J. S. (2016). Literacy discussions in low-income families: The effect of parent questions on fourth graders' retellings. *First Language*, *36*(1), 50–70. https://doi.org/10.1177/0142723715626071

Carver, C. S., & Scheier, M. F. (2012). *Attention and self-regulation: A control-theory approach to human behavior*. Springer Science & Business Media.

Catania, F., Spitale, M., Cosentino, G., & Garzotto, F. (2020). Conversational agents to promote children's verbal communication skills. In *International Workshop on Chatbot Research and Design* (pp. 158–172). Springer International Publishing.

Cheng, X., Yin, L., Lin, C., Shi, Z., Zheng, H., Zhu, L., Liu, X., Chen, K., & Dong, R. (2024). Chatbot dialogic reading boosts comprehension for Chinese kindergarteners with higher language skills. *Journal of Experimental Child Psychology*, *240*, 105842.

Chera, P., & Wood, C. (2003). Animated multimedia 'talking books' can promote phonological awareness in children beginning to read. *Learning and Instruction*, *13*(1), 33–52.

Chow, B. W. Y., McBride-Chang, C., & Cheung, H. (2010). Parent–child reading in English as a second language: Effects on language and literacy development of Chinese kindergarteners. *Journal of Research in Reading*, *33*(3), 284–301.

Chow, B. W. Y., McBride-Chang, C., Cheung, H., & Chow, C. S. L. (2008). Dialogic reading and morphology training in Chinese children: Effects on language and literacy. *Developmental Psychology*, *44*(1), 233–244.

Clinton-Lisell, V., Strouse, G., & Langowski, A. M. (2024). Children's engagement during shared reading of ebooks and paper books: A systematic review. *International Journal of Child-Computer Interaction*, *39*, 100632.

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Academic Press.

DesJardin, J. L., Stika, C. J., Eisenberg, L. S., Johnson, K. C., Ganguly, D. M. H., Henning, S. C., & Colson, B. G. (2017). A longitudinal investigation of the home literacy environment and shared book reading in young children with hearing loss. *Ear and Hearing*, *38*(4), 441–454.

Divekar, R. R., Lepp, H., Chopade, P., Albin, A., Brenner, D., & Ramanarayanan, V. (2021). Conversational agents in language education: Where they fit and their research challenges. In *International conference on human-computer interaction* (pp. 272–279). Springer International Publishing.

Dong, Y., & Chow, B. W. Y. (2022). Home literacy environment and English as a second language acquisition: A meta-analysis. *Language Learning and Development*, *18*(4), 485–499.

Dore, R. A., Hassinger-Das, B., Brezack, N., Valladares, T. L., Paller, A., Vu, L., Golinkoff, R. M., & Hirsh-Pasek, K. (2018). The parent advantage in fostering children's e-book comprehension. *Early Childhood Research Quarterly*, *44*, 24–33.

Du, Y., Zhang, K., Ramabadran, S., & Liu, Y. (2021). "Alexa, what is that sound?" a video analysis of child-agent communication from two Amazon Alexa games. In *Proceedings of the 20th annual ACM interaction design and children conference* (pp. 513–520). ACM Interaction Design and Children conference.

Ewin, C. A., Reupert, A. E., McLean, L. A., & Ewin, C. J. (2021). The impact of joint media engagement on parent–child interactions: A systematic review. *Human Behavior and Emerging Technologies*, *3*(2), 230–254.

Fitton, L., McIlraith, A. L., & Wood, C. L. (2018). Shared book reading interventions with English learners: A meta-analysis. *Review of Educational Research*, *88*(5), 712–751.

Flack, Z. M., Field, A. P., & Horst, J. S. (2018). The effects of shared storybook reading on word learning: A meta-analysis. *Developmental Psychology*, *54*(7), 1334–1346.

Flesch, R. (2007). Flesch-Kincaid readability test, *26*(3), 2007.

Fredricks, J. A., Wang, M.-T., Schall Linn, J., Hofkens, T. L., Sung, H., Parr, A., & Allerton, J. (2016). Using qualitative methods to develop a survey measure of math and science engagement. *Learning and Instruction*, *43*, 5–15. https://doi.org/10.1016/j.learninstruc.2016.01.009

García, O., & Wei, L. (2015). Translanguaging, bilingualism, and bilingual education. In *The handbook of bilingual and multilingual education* (pp. 223–240). Wiley.

Gilakjani, A. P. (2012). The significant role of multimedia in motivating EFL learners' interest in English language learning. *International Journal of Modern Education and Computer Science*, *4*(4), 57–66.

Guthrie, J. T., & Klauda, S. L. (2014). Effects of classroom practices on reading comprehension, engagement, and motivations for adolescents. *Reading Research Quarterly*, *49*(4), 387–416. https://doi.org/10.1002/rrq.81

Hargrave, A. C., & Sénéchal, M. (2000). A book reading intervention with preschool children who have limited vocabularies: The benefits of regular reading and dialogic reading. *Early Childhood Research Quarterly*, *15*(1), 75–90.

14678535, 2025, 5, Downloaded from https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13615, Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Hirsh-Pasek, K., Zosh, J. M., Golinkoff, R. M., Gray, J. H., Robb, M. B., & Kaufman, J. (2015). Putting education in "educational" apps: Lessons from the science of learning. *Psychological Science in the Public Interest*, *16*(1), 3–34.

Inbar-Lourie, O. (2010). English only? The linguistic choices of teachers of young EFL learners. *International Journal of Bilingualism*, *14*(3), 351–367.

Ismail, J. (2017). Language exposure and second language learning. *The English Teacher*, *20*(1), P20–P31.

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.

Jeon, J., Lee, S., & Choe, H. (2023). Beyond ChatGPT: A conceptual framework and systematic review of speech-recognition chatbots for language learning. *Computers & Education*, *206*, 104898.

Kennedy, C., & McLoughlin, A. (2023). Developing the emergent literacy skills of English language learners through dialogic reading: A systematic review. *Early Childhood Education Journal*, *51*(2), 317–332.

Korat, O., & Segal-Drori, O. (2016). E-book and printed book reading in different contexts as emergent literacy facilitator. *Early Education and Development*, *27*(4), 532–550.

Korat, O., & Shamir, A. (2007). Electronic books versus adult readers: Effects on children's emergent literacy as a function of social class. *Journal of Computer Assisted Learning*, *23*(3), 248–259. https://doi.org/10.1111/j.1365-2729.2006.00213.x

Korat, O., Shamir, A., & Heibal, S. (2013). Expanding the boundaries of shared book reading: E-books and printed books in parent-child reading as support for children's language. *First Language*, *33*(5), 504–523.

Kory-Westlund, J. M., & Breazeal, C. (2019). Assessing children's perceptions and acceptance of a social robot. In *Proceedings of the 18th ACM international conference on interaction design and children* (pp. 38–50). ACM Interaction Design and Children conference.

Kucirkova, N., & Flewitt, R. (2022). Understanding parents' conflicting beliefs about children's digital book reading. *Journal of Early Childhood Literacy*, *22*(2), 157–181.

Kucirkova, N., & Littleton, K. (2016). The digital reading habits of children. In *A National Survey of parents' perceptions of and practices in relation to children's reading for pleasure with print and digital books* (pp. 380–387).

Lai, V. D., Ngo, N. T., Veyseh, A. P. B., Man, H., Dernoncourt, F., Bui, T., & Nguyen, T. H. (2023). ChatGPT beyond English: Towards a comprehensive evaluation of large language models in multilingual learning. *arXiv preprint arXiv:2304.05613*.

Lever, R., & Sénéchal, M. (2011). Discussing stories: On how a dialogic reading intervention improves kindergartners' oral narrative construction. *Journal of Experimental Child Psychology*, *108*(1), 1–24. https://doi.org/10.1016/j.jecp.2010.07.002

Lin, G. C., Schoenfeld, I., Thompson, M., Xia, Y., Uz-Bilgin, C., & Leech, K. (2022). "What color are the fish's scales?" Exploring parents' and children's natural interactions with a child-friendly virtual agent during storybook reading. In *Proceedings of the 21st annual ACM interaction design and children conference* (pp. 185–195). ACM Interaction Design and Children conference.

Liu, C. C., Liao, M. G., Chang, C. H., & Lin, H. M. (2022). An analysis of children's interaction with an AI chatbot and its impact on their interest in reading. *Computers & Education*, *189*, 104576.

Markus, M. L. (1990). Toward a "critical mass" theory of interactive media. In J. Fulk & C. Steinfeld (Eds.), *Organization and communication technology* (pp. 194–218). Sage.

Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist*, *32*(1), 1–19.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, *38*(1), 43–52. https://doi.org/10.1207/s15326985ep3801_6

Mayer, R. E. (2002). Multimedia learning. In *Psychology of learning and motivation* (Vol. *41*, pp. 85–139). Academic Press.

Mayer, R. E. (2003). The promise of multimedia learning: Using the same instructional design methods across different media. *Learning and Instruction*, *13*(2), 125–139. https://doi.org/10.1016/S0959-4752(02)00016-6

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In *The Cambridge handbook of multimedia learning* (pp. 31–48). Cambridge University Press.

Mayer, R. E., Sobko, K., & Mautone, P. D. (2003). Social cues in multimedia learning: Role of speaker's voice. *Journal of Educational Psychology*, *95*(2), 419–425.

Mol, S. E., Bus, A. G., De Jong, M. T., & Smeets, D. J. (2008). Added value of dialogic parent–child book readings: A meta-analysis. *Early Education and Development*, *19*(1), 7–26.

Moody, A. K., Justice, L. M., & Cabell, S. Q. (2010). Electronic versus traditional storybooks: Relative influence on preschool children's engagement and communication. *Journal of Early Childhood Literacy*, *10*(3), 294–313. https://doi.org/10.1177/1468798410372162

Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments: Special issue on interactive learning environments: Contemporary issues and trends. *Educational Psychology Review*, *19*, 309–326.

Moreno, R., & Mayer, R. E. (2004). Personalized messages that promote science learning in virtual environments. *Journal of Educational Psychology*, *96*(1), 165–173. https://doi.org/10.1037/0022-0663.96.1.165

14678535, 2025, 5, Downloaded from https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13615, Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Moreno, R., & Mayer, R. E. (2005). Role of guidance, reflection, and interactivity in an agent-based multimedia game. *Journal of Educational Psychology*, *97*(1), 117–128.
Munzer, T. G., Miller, A. L., Weeks, H. M., Kaciroti, N., & Radesky, J. (2019). Differences in parent-toddler interactions with electronic versus print books. *Pediatrics*, *143*(4), e20182012. https://doi.org/10.1542/peds.2018-2012
Nakamura, S., Phung, L., & Reinders, H. (2021). The effect of learner choice on L2 task engagement. *Studies in Second Language Acquisition*, *43*(2), 428–441.
Neuman, S. B., Samudra, P., Wong, K. M., & Kaefer, T. (2020). Scaffolding attention and partial word learning through interactive coviewing of educational media: An eye-tracking study with low-income preschoolers. *Journal of Educational Psychology*, *112*(6), 1100.
O'Toole, K. J., & Kannass, K. N. (2018). Emergent literacy in print and electronic contexts: The influence of book type, narration source, and attention. *Journal of Experimental Child Psychology*, *173*, 100–115.
Perez, E., Kiela, D., & Cho, K. (2021). True few-shot learning with language models. *Advances in Neural Information Processing Systems*, *34*, 11054–11070.
Pillinger, C., & Vardy, E. J. (2022). The story so far: A systematic review of the dialogic reading literature. *Journal of Research in Reading*, *45*(4), 533–548. https://doi.org/10.1111/1467-9817.12407
Pintrich, P. R. (2003). Motivation and classroom learning. In W. M. Reynolds & G. E. Miller (Eds.), *Handbook of psychology: Educational psychology* (pp. 103–122). Wiley.
Qian, X., Tian, G., & Wang, Q. (2009). Codeswitching in the primary EFL classroom in China–two case studies. *System*, *37*(4), 719–730.
Reeve, J., Cheon, S. H., & Yu, T. H. (2020). An autonomy-supportive intervention to develop students' resilience by boosting agentic engagement. *International Journal of Behavioral Development*, *44*(4), 325–338. https://doi.org/10.1177/0165025420911103
Reich, S. M., Yau, J. C., Xu, Y., Muskat, T., Uvalle, J., & Cannata, D. (2019). Digital or print? A comparison of preschoolers' comprehension, vocabulary, and engagement from a print book and an e-book. *AERA Open*, *5*(3), 2332858419878389.
Richter, A., & Courage, M. L. (2017). Comparing electronic and paper storybooks for preschoolers: Attention, engagement, and recall. *Journal of Applied Developmental Psychology*, *48*, 92–102.
Rodríguez, E. T., & Tamis-LeMonda, C. S. (2011). Trajectories of the home learning environment across the first five years: Associations with children's vocabulary and literacy skills at prekindergarten. *Child Development*, *82*(4), 1058–1075.
Rogers, T. B., Kuiper, N. A., & Kirker, W. S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology*, *35*(9), 677–688. https://doi.org/10.1037/0022-3514.35.9.677
Schneider, S., Beege, M., Nebel, S., Schnaubert, L., & Rey, G. D. (2022). The cognitive-affective-social theory of learning in digital environments (CASTLE). *Educational Psychology Review*, *34*(1), 1–38.
Seitchik, A. E., Brown, A. J., & Harkins, S. G. (2017). Social facilitation: Using the molecular to inform the molar. In S. G. Harkins, K. D. Williams, & J. Burger (Eds.), *The Oxford handbook of social influence* (pp. 183–220). Oxford University Press.
Shepard-Carey, L. (2020). Making sense of comprehension practices and pedagogies in multimodal ways: A second-grade emergent bilingual's sensemaking during small-group reading. *Linguistics and Education*, *55*, 100777.
Simsek, Z. C., & Erdogan, N. I. (2015). Effects of the dialogic and traditional reading techniques on children's language development. *Procedia-Social and Behavioral Sciences*, *197*, 754–758.
Smeets, D. J., & Bus, A. G. (2015). The interactive animated e-book as a word learning device for kindergartners. *Applied PsychoLinguistics*, *36*(4), 899–920.
Smith, E. A., & Senter, R. J. (1967). *Automated readability index* (Vol. 66, No. 220). Aerospace Medical Research Laboratories, Aerospace Medical Division, Air Force Systems Command.
Son, S. H. C., & Tineo, M. F. (2016). Mothers' attention-getting utterances during shared book reading: Links to low-income preschoolers' verbal engagement, visual attention, and early literacy. *Infant and Child Development*, *25*(4), 259–282.
Spatola, N., Belletier, C., Chausse, P., Augustinova, M., Normand, A., Barra, V., Ferrand, L., & Huguet, P. (2019). Improved cognitive control in presence of anthropomorphized robots. *International Journal of Social Robotics*, *11*, 463–476.
Strouse, G. A., & Ganea, P. A. (2017). Toddlers' word learning and transfer from electronic and print books. *Journal of Experimental Child Psychology*, *156*, 129–142.
Strouse, G. A., Troseth, G. L., & Stuckelman, Z. D. (2023). Page and screen: Storybook features that promote parent-child talk during shared reading. *Journal of Applied Developmental Psychology*, *86*, 101522.
Takacs, Z. K., & Bus, A. G. (2018). How pictures in picture storybooks support young children's story comprehension: An eye-tracking experiment. *Journal of Experimental Child Psychology*, *174*, 1–12.
Takacs, Z. K., Swart, E. K., & Bus, A. G. (2015). Benefits and pitfalls of multimedia and interactive features in technology-enhanced storybooks: A meta-analysis. *Review of Educational Research*, *85*(4), 698–739.

14678535, 2025, 5, Downloaded from https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13615, Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

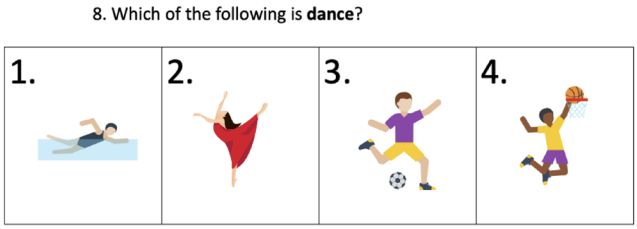
Talamas, A., Kroll, J. F., & Dufour, R. (1999). From form to meaning: Stages in the acquisition of second-language vocabulary. *Bilingualism: Language and Cognition*, *2*(1), 45–58.
Troseth, G. L., Strouse, G. A., Flores, I., Stuckelman, Z. D., & Johnson, C. R. (2020). An enhanced eBook facilitates parent–child talk during shared reading by families of low socioeconomic status. *Early Childhood Research Quarterly*, *50*, 45–58.
Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
Wei, J., Bosma, M., Zhao, V. Y., Guu, K., Yu, A. W., Lester, B., Du, N., Dai, A. M., & Le, Q. V. (2021). Finetuned language models are zero-shot learners. *arXiv preprint arXiv:2109.01652*.
Whitehurst, G. J., Arnold, D. S., Epstein, J. N., Angell, A. L., Smith, M., & Fischel, J. E. (1988). A picture book reading intervention in day care and home for children from low-income families. *Developmental Psychology*, *24*(4), 552–559.
Whitehurst, G. J., & Lonigan, C. J. (1998). Child development and emergent literacy. *Child Development*, *69*(3), 848–872. https://doi.org/10.1111/j.1467-8624.1998.tb06247.x
Wu, X., Duan, R., & Ni, J. (2024). Unveiling security, privacy, and ethical concerns of ChatGPT. *Journal of Information and Intelligence*, *2*(2), 102–115.
Xiao, F., Zhao, P., Sha, H., Yang, D., & Warschauer, M. (2024). Conversational agents in language learning. *Journal of China Computer-Assisted Language Learning*, *4*(2), 300–325.
Xu, Y. (2023). Talking with machines: Can conversational technologies serve as children's social partners? *Child Development Perspectives*, *17*(1), 53–58.
Xu, Y., Aubele, J., Vigil, V., Bustamante, A. S., Kim, Y. S., & Warschauer, M. (2022). Dialogue with a conversational agent promotes children's story comprehension via enhancing engagement. *Child Development*, *93*(2), e149–e167.
Xu, Y., Branham, S., Deng, X., Collins, P., & Warschauer, M. (2021). Are current voice interfaces designed to support children's language development? In *Proceedings of the 2021 CHI conference on human factors in computing systems* (pp. 1–12). ACM Conference on Human Factors in Computing Systems.
Xu, Y., He, K., Vigil, V., Ojeda-Ramirez, S., Liu, X., Levine, J., Cervera, K., & Warschauer, M. (2023). "Rosita reads with my family": Developing a bilingual conversational agent to support parent-child shared Reading. In *Proceedings of the 22nd annual ACM interaction design and children conference* (pp. 160–172). ACM Interaction, Design and Children Conference.
Xu, Y., Vigil, V., Bustamante, A. S., & Warschauer, M. (2022). Contingent interaction with a television character promotes children's science learning and engagement. *Journal of Applied Developmental Psychology*, *81*, 101439.
Xu, Y., Wang, D., Collins, P., Lee, H., & Warschauer, M. (2021). Same benefits, different communication patterns: Comparing Children's reading with a conversational agent vs. a human partner. *Computers & Education*, *161*, 104059.
Xu, Y., & Warschauer, M. (2019). Young children's reading and learning with conversational agents. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–8. https://doi.org/10.1145/3290607.3299035
Yang, D., Xia, C., Collins, P., & Warschauer, M. (2022). The role of bilingual discussion prompts in shared E-book reading. *Computers & Education*, *190*, 104622. https://doi.org/10.1016/j.compedu.2022.104622
Zevenbergen, A. A., & Whitehurst, G. J. (2003). Dialogic reading: A shared picture book reading intervention for preschoolers. *On Reading Books to Children: Parents and Teachers*, *177*, 200.
Zhang, Z., Xu, Y., Wang, Y., Yao, B., Ritchie, D., Wu, T., & Li, T. J. J. (2022). Storybuddy: A human-AI collaborative chatbot for parent-child interactive storytelling with flexible parental involvement. In *Proceedings of the 2022 CHI conference on human factors in computing systems* (pp. 1–21). ACM Conference on Human Factors in Computing Systems.
Zhou, N., & Yadav, A. (2017). Effects of multimedia story reading and questioning on preschoolers' vocabulary learning, story comprehension and reading engagement. *Educational Technology Research and Development*, *65*(6), 1523–1545. https://doi.org/10.1007/s11423-017-9533-2

How to cite this article: Xiao, F., Zou, E. W., Lin, J., Li, Z., & Yang, D. (2025). Parent-led vs. AI-guided dialogic reading: Evidence from a randomized controlled trial in children's e-book context. *British Journal of Educational Technology*, *56*, 1784–1813. https://doi.org/10.1111/bjet.13615

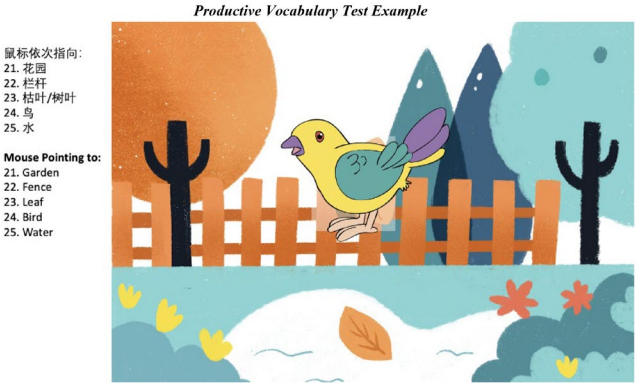
14678535, 2025, 5, Downloaded from https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13615, Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

APPENDIX A

Receptive vocabulary test example



Productive vocabulary test example



APPENDIX B

Engagement coding rubrics

Unit of coding: 8-s as a time segment.

Itemized coding	
Item	**Description**
Vocalizations	
Narrative-relevant vocalization	Children made vocalizations relevant to story/narrative (eg, 'It looks like a dry leaf')
Narrative-irrelevant vocalization	Children made vocalizations irrelevant to story/narrative (eg, 'I want to play with my toy'.)
Device-relevant vocalization	Children made vocalizations relevant to app features, devices, etc (eg, 'It's running too slow.', 'Mia is asking questions', 'I want to listen to Chinese')
Visual attention	
Attention	The child's attention to the story is coded when they focus on the e-book, tablet or main eye contact with the parent. If the child turns away to somewhere else at any time, their visual attention is considered absent and is scored as 0

14678535, 2025, 5, Downloaded from https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13615, Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Itemized coding	
Item	**Description**
Affective response	
Positive affect	Positive affect includes the following 16 expressive emotion displays during the segment: smiling, cheering, clapping, dancing, jumping in excitement, laughing audibly, singing, showing eagerness, giggling, raising cheeks, pulling up lip corners, crinkling eyes, showing affection, smirking, speaking in a warm emotional tone or using terms of endearment (Bai et al., 2016)
Negative affect	Negative affect includes children's explicit display of negative emotions during the segment, such as frowning, crying, showing frustration and confusion, sighing audibly, expressing annoyance, displaying anger, whining, stamping feet, sulking, using a harsh tone of voice or verbalizing negative sentiments
Holistic engagement (1–5 scale)	

- Score of 5: Reflects the highest level of engagement. Indicators include clear signs of excitement stemming from the reading material, expressive hand movements to illustrate points, repeating utterances to ensure understanding and maintain the conversation, leaning towards the screen and sitting attentively.
- Score of 3: Represents a moderate level of engagement where the child fulfils the basic requirements of the protocol. This includes behaviours like listening and remaining seated but not actively interacting or clicking proactively.
- Score of 1: Indicates the lowest level of engagement. The child appears clearly distracted and shows minimal interest in the story, exemplified by random postures instead of sitting attentively.

APPENDIX C

Group comparison for vocabulary, comprehension and retelling at different time point

Measure	Time	Parent group	CA group	*U*-test
Vocabulary	Pretest	22.20 (12.96)	25.16 (14.72)	0.34
	Posttest	29.70 (11.34)	32.22 (12.84)	0.27
	Delay test	29.33 (13.37)	32.03 (12.77)	0.38
Receptive vocabulary	Pretest	10.53 (6.30)	12.08 (6.91)	479.0
	Posttest	16.23 (5.14)	17.49 (5.52)	467.5
	Delay test	15.67 (6.78)	17.32 (5.60)	485.5
Productive vocabulary	Pretest	11.67 (7.12)	13.08 (8.15)	492.0
	Posttest	13.47 (7.31)	14.73 (8.09)	501.5
	Delay test	13.67 (7.44)	14.70 (8.20)	507.0
Comprehension	Posttest	18.08 (3.47)	21.15 (2.98)	266.0***
	Delay test	17.75 (4.84)	20.09 (3.57)	406.0*
Retelling	Posttest	11.23 (5.91)	13.50 (5.26)	416.0 (.)
	Delay test	10.85 (5.38)	12.73 (5.51)	448.5

Note: Numbers in parentheses are standard deviation. * denotes $p < 0.05$, ** for $p < 0.01$, *** for $p < 0.001$, (.) for marginal effect.

14678535, 2025, 5, Downloaded from https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13615, Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

APPENDIX D

Group comparison for engagement variables and *U*-test results

Variable	Group	*n*	Mean	*SD*	Median	Trimmed	*U*	*p*-value	Adjusted *p*-value
Narrative-relevant vocalization	CA	37	0.24	0.14	0.21	0.23	767	0.01*	0.004
	Parent	30	0.36	0.18	0.35	0.34			
Narrative-irrelevant vocalization	CA	37	0.01	0.06	0	0	755	0.01*	0.000
	Parent	30	0.02	0.03	0.01	0.01			
Device-relevant vocalization	CA	37	0.08	0.22	0.02	0.03	520	0.66	0.026
	Parent	30	0.02	0.03	0.01	0.02			
Visual attention	CA	37	0.89	0.16	0.95	0.92	240	0.00*	0.003
	Parent	30	0.76	0.23	0.82	0.79			
Affect positive	CA	37	0.09	0.1	0.05	0.07	733	0.03	0.245
	Parent	30	0.12	0.09	0.09	0.1			
Affect negative	CA	37	0.03	0.03	0.01	0.02	467	0.27	1.000
	Parent	30	0.03	0.05	0.01	0.02			
Global engagement	CA	37	4.5	0.43	4.66	4.55	282	0.00*	0.004
	Parent	30	3.97	0.71	4.1	4.04			

Note: * denotes the group difference remained statistically significant after Bonferroni correction.

APPENDIX E

Demographic characteristic per group

Characteristic	Control group (*N* = 30)	Experimental group (*N* = 37)
Child age (years) (*M*/*SD*)	7.2 (2.1)	7.05 (1.98)
Child gender		
Male (*n*/%)	16 (40%)	14 (34%)
Female (*n*/%)	24 (60%)	27 (66%)
Child English level		
Beginner (no prior learning)	6 (20%)	4 (11%)
Know basic word	23 (73%)	30 (78%)
Learned for a while, good proficiency	1 (7%)	3 (11%)
Child–CA use		
Never	6 (20%)	11 (29.7%)
Rarely	16 (53.3%)	15 (40.5%)
Sometimes	6 (20%)	9 (24.3%)
Frequently	2 (6.67%)	2 (41%)
Parent age (*M*/*SD*)	34.85 (5.62)	33.2 (4.78)
Parent education		
Middle school	0 (0%)	3 (8.11%)
High school	4 (13.3%)	4 (10.8%)

14678535, 2025, 5, Downloaded from https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13615, Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Characteristic	Control group ($N = 30$)	Experimental group ($N = 37$)
Associate degree	2 (6.7%)	5 (13.5%)
Bachelor's degree	13 (43.4%)	10 (27%)
Master's degree or above	11 (36.7%)	15 (40.5%)
Parent English level		
Not at all	3 (10%)	6 (16.3%)
Basic conversation	8 (26.7%)	10 (27%)
CET-4	11 (36.6%)	9 (24.3%)
CET-6 or above	8 (26.7%)	12 (32.4%)
Chinese shared reading fluency		
Rarely	11 (36.7%)	12 (32.4%)
Sometimes	10 (33.3%)	15 (40.5%)
Frequently	9 (30%)	10 (27%)
English shared reading fluency		
Never	4 (13.3%)	5 (13.5%)
Rarely	20 (66.7%)	14 (37.8%)
Sometimes	3 (10%)	11 (29.7%)
Frequently	3 (10%)	7 (18.9%)
Shared English/bilingual E-book reading		
Never	14 (46.7%)	10 (27%)
Rarely	10 (33.3%)	16 (43.2%)
Sometimes	5 (33.3%)	6 (16.2%)
Frequently	1 (3.33%)	5 (13.5%)

14678535, 2025, 5, Downloaded from https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13615, Wiley Online Library on [24/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License