UC Irvine
UC Irvine Previously Published Works

Title
Dialogue with a conversational agent promotes children's story comprehension via enhancing engagement

Permalink
https://escholarship.org/uc/item/1rw6t612

Journal
Child Development, 93(2)

ISSN
0009-3920

Authors
Xu, Ying
Aubele, Joseph
Vigil, Valery
et al.

Publication Date
2022-03-01

DOI
10.1111/cdev.13708

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

DOI: 10.1111/cdev.13708

EMPIRICAL ARTICLE

CHILD DEVELOPMENT

Dialogue with a conversational agent promotes children's story comprehension via enhancing engagement

Ying Xu | **Joseph Aubele** | **Valery Vigil** | **Andres S. Bustamante** | **Young-Suk Kim** | **Mark Warschauer**

School of Education, University of California, Irvine, Irvine, California, USA

Correspondence
Ying Xu, School of Education at the University of California, Irvine, Irvine, CA 92697, USA.
Email: ying.xu@uci.edu

Funding information
This research was supported in part by the National Science Foundation under Grant No. 1906321.

Abstract

Dialogic reading, when children are read a storybook and engaged in relevant conversation, is a powerful strategy for fostering language development. With the development of artificial intelligence, conversational agents can engage children in elements of dialogic reading. This study examined whether a conversational agent can improve children's story comprehension and engagement, as compared to an adult reading partner. Using a 2 (dialogic reading or non-dialogic reading) × 2 (agent or human) factorial design, a total of 117 three- to six-year-olds (50% Female, 37% White, 31% Asian, 21% multi-ethnic) were randomly assigned into one of the four conditions. Results revealed that a conversational agent can replicate the benefits of dialogic reading with a human partner by enhancing children's narrative-relevant vocalizations, reducing irrelevant vocalizations, and improving story comprehension.

Preschool years are a critical time for developing language skills that are needed to succeed in school. Storybook reading with adults, typically caregivers or teachers, provides a prime context to bolster children's language development. In line with the Vygotskian principle of scaffolding (Berk & Winsler, 1995), the benefits of storybook reading are amplified by engaging children in contingent, structured interactions that revolve around story narratives and facilitate conversation about content just above the child's current level of understanding. This interactive reading style—termed dialogic reading—includes asking open-ended questions to stimulate children's thinking and providing feedback for child participation (Arnold et al., 1994; Whitehurst, 1992). Dialogic reading interventions with caregivers and teachers have confirmed the value of using dialogue for enhancing children's engagement during reading and supporting children's vocabulary learning, comprehension, and expressive language (for reviews, see Flack et al., 2018; Mol et al., 2008; Noble et al., 2019; Towson et al., 2017). However, the quantity and quality of storybook reading children are exposed to depend on the training opportunities, availability, skills, and inclinations of their caregivers or teachers. Unequal access to high-quality reading experiences is believed to contribute to the language and literacy divide among children in the United States. (Farver et al., 2013; Phillips & Lonigan, 2009).

With the rapid development of artificial intelligence, children are increasingly interacting with non-human intelligent agents through speech, gesture, or writing. Conversational agents that support natural speech interaction may be especially valuable for young children, whose lack of proficiency in literacy or fine motor skills causes them difficulty in navigating digital environments (Lovato & Piper, 2019). Conversational agents

Abbreviations: IRR, interrater-reliability; SEM, structural equation modeling; DR, dialogic reading.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
© 2021 The Authors. *Child Development* published by Wiley Periodicals LLC on behalf of Society for Research in Child Development.

comprehend speech, thus enabling complex dialogue that mimics human-to-human conversation. Familiar examples of speech-based agents include Apple Siri, Google Assistant, and Amazon Alexa. Due to these products' growing prevalence, the developmental consequences of children interacting with speech-based agents have spurred much research interest (e.g., Garg & Sengupta, 2020; Sciuto et al., 2018; Yuan et al., 2019). Some argue that machine-mediated communication afforded by agents is a form of social interactions akin to interpersonal communication, with the agent thus assuming the role of a partner or guide in children's language learning (e.g., Roseberry et al., 2014). However, there is little evidence as to whether and how interacting with conversational agents supports language development.

This experimental study directly examines this issue. We focus on the impact and mechanism of learning and engagement in storybook reading by young children when interacting with a conversational agent compared to a human partner. Evidence that conversational agents can emulate the benefits of an adult co-reader would offer a promising mechanism for supporting children's language development in daily life (see discussion in Sengupta & Garg, 2019). In the following section, we discuss the theoretical perspectives underpinning this study and prior work that led to the formation of our research questions and hypotheses.

The social nature of language development

Sociocultural theory views language development as a mediated process in which children acquire language skills through collaborative dialogue with more knowledgeable members of society in everyday activities (John-Steiner & Mahn, 1996). Through back-and-forth conversation with more knowledgeable language partners who provide scaffolding and facilitate active participation, children internalize knowledge by focusing attention, expressing thoughts, and critically reflecting on the topic being discussed (Golinkoff et al., 2019). Moreover, sociocultural theory emphasizes that the experienced adult should purposely craft a language environment that is developmentally appropriate to the child (Bodrova & Leong, 2005). In other words, the adult should assume the role of a language guide and scaffold children's participation in the conversation. A great deal of research has adopted this perspective and designed socially interactive learning experiences to support children's language development, either in face-to-face settings (e.g., dialogic reading with an adult) or computer-mediated environments (e.g., conversational agents). This relevant literature is reviewed in the following sections.

Dialogic questioning during reading

Whitehurst and colleagues established the interactive "dialogic reading" paradigm that involves adults using elaborative questioning and feedback techniques to encourage children's oral contributions (Arnold & Whitehurst, 1994; Whitehurst, 1992). Specifically, during dialogic reading sessions, the adult uses elaborative "wh-" and open-ended questions, repetition of good responses, and expansion of incomplete responses to model sentence formation. The benefits of dialogic reading are supported by a large volume of correlational, experimental, and intervention research (for reviews, see Flack et al., 2018; Mol et al., 2008; Noble et al., 2019; Towson et al., 2017). The studies reported on in these papers look at a broad range of short-term outcomes pertaining to the specific books being read as well as long-term outcomes including expressive and receptive language skills and reading attitudes. For example, Lever and Sénéchal (2011) found that dialogic reading with parents improved children's story comprehension indicated by children's accuracy and linguistic complexity in oral retelling of the story. In another study, children were asked questions requiring them to label illustrations representing target vocabulary words during storybook reading, and they were able to comprehend and produce more of those words than children who simply listened to the same story without any prompted interactions (Sénéchal et al., 1995).

One important question is how dialogic reading affects children across different subgroups. In theory, children who have lower language proficiency or are younger may benefit from language scaffolding more than other children. And indeed, many studies have reported dialogic reading's positive outcomes for children with lower language proficiency. For example, Hargrave and Sénéchal (2000) found that students with limited vocabulary who were exposed to dialogic reading made greater gains both in content-specific vocabulary they encountered and on a standard expressive vocabulary assessment, compared to students who were not exposed to dialogic reading. However, fewer studies have tested whether children's prior language skills moderate the effects of dialogic reading interventions. When meta-analytic methods have been used to synthesize the results across studies, language skills within typically developing children have played a negligible role in the variability of intervention effects (Flack et al., 2018). However, another meta-analysis suggested that children who were at risk of language impairment (i.e., low family income and parental education) benefited less from dialogic reading (Mol et al., 2008). Notably, most studies on dialogic reading tend to target preschool-aged children. While these individual studies suggest preschoolers benefit from dialogic reading, few studies have explicitly tested whether these benefits vary by age within this group. Findings from several meta-analyses were mixed: while Mol et al. (2008) indicated that the positive effects of storybook reading were smaller for studies involving older preschool children (4–5 years old), Flack et al. (2018)

and Noble et al. (2019) suggested that the effects do not seem to vary by children's age. However, as Flack et al. suggested, meta-analyses may not be sufficient for identifying heterogeneous effects of dialogic reading among children with different language abilities and of different ages. This is particularly true if the reading interventions carried out in the original studies were already specifically tailored to the children's language development or age. Overall, while the extant literature shows that dialogic reading is generally an effective method for promoting literacy and language development, the evidence is inconclusive regarding which subgroups of children may benefit the most (Lever & Sénéchal, 2011; McNeill & Flower, 1999; Zevenbergen, & Whitehurst, 2003).

Theory suggests that the learning benefits of dialogic reading occur, at least partially, as a result of children's enhanced engagement during reading. According to Guthrie and Klauda's (2014) well-cited framework, reading engagement consists of behavioral, emotional, and cognitive dimensions. Behavioral engagement refers to how attentive students are during the reading session and is usually measured by how much children visually attend to the reading materials; emotional engagement refers to students' enthusiasm and feelings about what they are reading, and is usually measure by their emotional expressions; and cognitive engagement refers to a child actively thinking in order to comprehend the story and participate in discussion, and is often measured using vocalization as a proxy for cognitive engagement (Neuman et al., 2019; Troseth et al., 2020; Xu et al., 2020; Zhou & Yadav, 2017). A number of studies have suggested that increased reading engagement resulting from dialogic interaction is associated with enhanced outcomes. For example, Neuman et al. (2019) used eye-tracking to show that dialogic co-viewing, where an adult prompted children's attention using techniques such as repeating words, pointing to objects, or providing brief recaps of certain plot points, enhanced children's visual attention to narrative content and resulted in enhanced word learning. Troseth et al. (2020) found that, when parents were prompted to utilize dialogic questioning strategies, children were more cognitively engaged as they talked more and used more diverse vocabulary. Zhou and Yadav (2017) found that children who were asked questions during story book reading showed higher levels of behavioral engagement (as indicated by remaining seated and looking at the book), cognitive engagement (as indicated by meaningfully responding to adult questions and reading along with the adult), and emotional engagement (as indicated by showing positive facial expression or showing empathy with story characters), and developed a better understanding of the story plot and vocabulary. Taken together, these studies have established theoretical and empirical models to examine engagement and its mediating role during storybook reading.

Social learning with artificially intelligent agents

Artificial intelligence has powered agents that allow for communication using natural spoken language. These conversational agents possess different properties, including those with (e.g., robots, avatars) and without embodiment (e.g., phone-based voice assistants, smart speakers; Lee et al., 2006). A number of studies have found that children engage in natural conversation with such agents. For example, through analyzing audio recordings of children talking with the smart speakers deployed in their home, Beneteau et al. (2020) identified three common purposes of children's interaction, namely entertainment, assistance, and information seeking. Interview studies suggest that children attribute human properties to the agents. For example, our earlier study found the majority of preschool-aged children perceived conversational agents to possess cognitive abilities, which they believed enabled the agents to comprehend speech (Xu & Warschauer, 2020a). Together, these studies point to the feasibility of conversational agents as social partners for children (e.g., Roseberry et al., 2014).

Along these lines, studies have specifically explored the use of conversational agents to accompany children during learning processes. For example, Kory and Breazeal (2014) studied how a robot, operated by a human experimenter behind the scenes, could support children's story creation by prompting children to draw attention to the main elements of stories (e.g., what, where, who). This robot taught children story structures and facilitated children's telling of more complex stories. Targeting slightly older children, Michaelis and Mutlu (2018) developed a robot companion to promote elementary school students' interest in reading; the robot was designed to make preprogrammed comments intermittently as children read aloud and to provide non-verbal cues (e.g., eye gaze, semi-randomized idle movements) to demonstrate good listening. This in-home study found the robot motivated children to read and elicited children's social response (i.e., affliction). Our research team also conducted another study to investigate the use of an intelligent media character to engage children in science-related talk during an animated video and found that it helped children learn scientific vocabulary. (Xu & Warschauer, 2020b) These studies demonstrate the role artificial intelligence may play in enriching children's early learning experiences.

Furthermore, several studies suggest that properly designed agents can be equally effective as human language partners. Most of these studies have involved embodied conversational agents, such as robots or on-screen intelligent avatars. For example, Westlund et al. (2017) found that children learned unfamiliar words equally well with a robot or a human interlocutor. Hong et al. (2016) demonstrated that incorporating a robot teaching assistant in a classroom led to similar levels of reading and writing improvement as compared to having a human assistant. To our knowledge, there is

only one study focusing on the comparison between a disembodied conversational agent and a human partner (Aeschlimann et al., 2020). Children collaborated with either a voice assistant (i.e., a smart speaker) or an adult experimenter in a treasure hunt game, which required children to provide necessary information to their respective collaborator. Children supplied more information to the adult experimenter than to the voice assistant. However, this study was carried out in a game-play setting and thus was not able to answer the questions of specific language learning benefits resulting from interaction with a disembodied conversational agent during book reading.

The present study

This study is the first to focus on preschool-aged children's engagement with and learning from a disembodied conversational agent as a dialogic reading partner compared to their engagement and learning from reading with an adult. We explored three research questions:

RQ1: What is the effect of dialogic reading on children's story comprehension? Does the effect of dialogic reading vary by whether children read with a conversational agent versus with a human partner?
RQ2: What is the effect of dialogic reading on children's reading engagement? Does the effect of dialogic reading vary by whether children read with a conversational agent versus with a human partner?
RQ3: Does children's reading engagement serve as a mechanism through which dialogic reading with a conversational agent affects story comprehension?

To answer these questions, we conducted a two-by-two factorial experiment, with the two factors being whether children had dialogic or non-dialogic reading and whether children were partnered with a conversational agent or an adult. One hundred and seventeen children aged 3–6 were randomly assigned into one of the four conditions. Children's story comprehension was measured after reading, and their engagement was analyzed from video recordings of the reading sessions.

For RQ1 and RQ2, we hypothesized that children in the dialogic reading groups would be more engaged in the reading and comprehend the story better than those in the non-dialogic reading groups. This is expected given the advantages documented by dialogic interactions (Hargrave & Sénéchal, 2000; Şimşek & Işıkoğlu Erdoğan, 2015). Regarding the effects of dialogic reading with a human or agent partner, while an in-person partner has long been viewed as more natural than artificially intelligent agents (e.g., Aeschlimann et al., 2020), studies have repeatedly shown that properly designed agents enhance engagement and learning (e.g., Tewari & Canny, 2014). Thus, we expect that having dialogic reading with our agent partner will benefit children's story comprehension and engagement similarly to with a human partner. Specifically, we formed the following hypotheses (H) for our research questions:

For RQ1 focusing on comprehension, we made two hypotheses:

H 1a *Dialogic reading will increase children's story comprehension compared to non-dialogic reading.*
H 1b *The effect of dialogic reading on story comprehension will not vary by reading partner (human vs. agent).*

For RQ2 focusing on engagement, we also made two hypotheses:

H 2a *Dialogic reading will increase children's reading engagement compared to non-dialogic reading.*
H 2b *The effect of dialogic reading on reading engagement will not vary by reading partner (human vs. agent).*

For RQ3, we hypothesized that engagement would be a significant mechanism through which conversational agents enhance learning. Engagement has been posited as a key factor in enhancing reading comprehension, and engaged children are more often motivated to understand the story content with a higher level of cognitive efforts (Guthrie & Klauda, 2014). Specifically:

H 3 *The impact of dialogic reading with a conversational agent on children's story comprehension will be mediated by child engagement during the storybook reading.*

This study focused on preschool-aged children for two reasons. First, children develop their language skills rapidly during the preschool years, and interventions targeting this age group can have long-lasting consequences on children's later reading development and academic achievement (Shanahan et al., 2006). As such, many studies have focused on early interventions that support children's language development, and those involving dialogic reading are proven to be effective. Second, preschool-aged children may particularly benefit from dialogic reading with conversational agents as these children are not yet able to read or write, and providing them with novel voice-based interaction opportunities thus allows them to engage in scaffolded reading tailored to both their educational needs and their available modes of interaction.

METHOD

Participants

One hundred and twenty-two children aged 3–6 years were recruited from five child-care centers serving middle-class communities and participated in the

experiment (data collection: 2/2019–8/2019). To recruit these children, we reached out to the directors of the child-care centers, and, with their approval, set up a recruitment booth at each site during pick-up times to gather parent signatures and answer any questions parents may have had. Parents or guardians also completed a brief survey on demographic characteristics and information related to their child's prior experiences with conversational technologies. Five children who participated in the study were excluded due to data loss resulting from technological problems with the recording device, which resulted in an analytic sample consisting of 117 children (age range = 37–81 months, M = 58.10 months, SD = 9.53 months). Fifty percent of the children were girls and children represented a wide variety of ethnic backgrounds. Almost 80% of these children predominantly spoke English at home. Table 1 presents participants' background information.

Study design

This study used a two (reading partner as conversational agent vs. adult) by two (dialogic reading vs. non-dialogic reading) factorial design, with participants randomly assigned into one of four conditions. Specifically, we utilized a randomized block design, in which participants in each school site were randomly assigned into an experimental condition. The purpose of such a design is to increase the homogeneity of experimental units, thus reducing experimental errors and increasing the power for detecting treatment factor effects. The four conditions were as follows:

- *Agent Dialogic Reading (Agent DR)* where the agent narrated the story to a child and engaged the child in dialogue by asking questions and providing feedback.
- *Agent Non-Dialogic Reading (Agent Non-DR)* where the agent merely narrated the same story to a child but did not ask any questions to engage the child in dialogue.
- *Human Dialogic Reading (Human DR)* where an adult narrated the story to a child and engaged the child in dialogue by asking questions and providing feedback.
- *Human Non-Dialogic Reading (Human Non-DR)* where an adult merely narrated the same story to a child but did not ask any questions to engage the child in dialogue.

In the "Human DR" condition, the human experimenter followed the same dialogue script that was designed for the agent. Adherence to the script ensured that the verbal exposure in the two dialogic reading conditions (Agent DR and Human DR) was comparable, thus increasing the internal validity of the study findings, albeit potentially limiting ecological validity, as we discuss later in this paper.

TABLE 1 Background information by condition

	Full sample	Agent DR	Agent non-DR	Human DR	Human non-DR	ANOVA/χ^2
Age	57.63 (9.53)	59.97 (9.05)	57.59 (10.41)	58.29 (8.82)	53.92 (9.42)	$F(3, 116) = 2.09$, $p = .11$
EOWPVT	69.18 (17.22)	70.58 (17.43)	70.70 (19.71)	66.71 (17.09)	68.77 (14.82)	$F(3, 116) = 0.34$, $p = .80$
Predominant home language						$\chi^2(3) = 1.40$, $p = .70$
English	78.63%	75.76%	85.19%	80.65%	73.08	
Other	21.37%	24.24%	14.81%	19.35%	26.92%	
Female	49.57%	57.58%	48.15%	48.39%	42.31%	$\chi^2(3) = 1.19$, $p = .75$
Race						$\chi^2(18) = 18.28$, $p = .44$
White	36.75%	33.33%	44.44%	32.26%	38.46%	
Asian	30.77%	27.27%	33.33%	41.94%	19.23%	
Hispanic	6.84%	12.12%	3.70%	0.00%	11.54%	
Black	0.85%	0.00%	0.00%	3.23%	0.00%	
Two or more	21.37%	24.24%	11.11%	22.58%	26.92%	
Other	1.71%	0.00%	3.70%	0.00%	3.85%	
Decline	0.85%	3.03%	3.70%	0.00%	00.00%	
Regular conversational agent usage						$\chi^2(3) = 0.54$, $p = .91$
Yes	43.59%	45.45%	40.74%	48.39%	38.46%	
No	55.56%	54.55%	59.26%	51.61%	57.69%	
Decline	0.85%	0.00%	0.00%	0.00%	3.85%	
N	117	33	27	31	26	

Note: Standard deviation in parentheses.
Abbreviations: DR, dialogic reading; EOWPVT, Expressive One Word Picture Vocabulary Test.

Experimental stimuli

The story reading materials were adopted from a commercially available picture book, *Three Bears in a Boat*, authored by David Soman. The story is about three little bears who accidentally break their mother's precious seashell and then embark on an adventure to search for a replacement. The story was chosen based on length, potential story interest, the low likelihood that participants would have read the book previously, and appropriate level of narrative complexity. The print book was 16 pages long, with each page consisting of about 6 sentences (an average of 11 words per sentence) accompanied by illustrations. We analyzed the book's narrative complexity using the rubric developed in Petersen et al. (2008) and determined that the book is appropriate for preschool children because it contains (i) main characters with names, (ii) specific places and times where the story took place, and (iii) a clear story sequence with causes and consequences.

Both human and agent dialogic reading conditions followed the exact same dialogue script (i.e., asking the same questions and providing responsive feedback in the same manner). Nine open-ended questions were asked throughout the storytelling. Based on Blewitt et al.'s (2009) suggestion, we incorporated a combination of 6 low-cognitive-demand and 3 high-cognitive-demand questions. For example, the following is a sentence from the story: "One day, when their mother was out, the three bears did something they really shouldn't have, and with a crash, their mother's beautiful blue seashell lay scattered in pieces across the floor." A low-cognitive-demand question asked, "What did the bears break?" And the answer to that question was "seashell", which was found directly in the text. A high-cognitive-demand question asked children to make an inference based on the given information in the story or to summarize the information (e.g., "How did the bears search for the seashell?"). Both the human and agent provided elaborative feedback to children's responses in a way that acknowledged what the children had said and explained the question to solidify children's understanding or clarify any confusion. For example, after children responded to the question of why the bears stopped at an island, the agent first assessed the children's answer, and then explained the reason the bears stopped there as follows, "The bears stopped at this island because they think they can find a blue seashell here. The old salty bear said the blue seashell is on the island shaped like a lumpy hat."

As shown in Figure 1, children in each of the four conditions looked at a physical copy of the storybook while they were read to. Each page of the storybook contained printed text as well as illustrations relevant to that text. We intended to simulate ordinary joint reading activities, in which a child is typically provided with a book while being read to. Even though the children at this age range were not yet able to read, looking at the illustrations while being read to may have facilitated their comprehension of the narration (Kaefer et al., 2017) and made the reading activity more engaging (Ann Evans & Saint-Aubin, 2005). A Google Home Mini device (pictured in the right panel of Figure 1) was utilized in the two agent conditions. In the dialogic reading condition, the Google Home Mini device narrated the story and conversed with the children, while in the non-dialogic reading condition, the device merely narrated the story without asking questions.

Procedure

Children met individually with a trained adult experimenter in a designated quiet area at their school for two sessions. In the first session, the participants received an Expressive One Word Picture Vocabulary Test as a pretest, which was used as the baseline measure of their expressive vocabulary skills.

In the second session, children engaged in the storybook reading activity and answered post-reading assessment questions. Prior to the reading, children interacted with the experimenter or the conversational agent through a structured dialogue, depending on their assigned condition. The dialogue involved the conversational agent or experimenter asking the children their age, favorite color, and simple animal fact questions and then repeating the children's responses (Agent/human: "What is your favorite color?"; child: "I like red the best."; agent/human: "Great choice! My favorite color is also red."). The purpose of including this pre-reading interaction was to build rapport between the child and the reading partner, as well as to provide children in the Agent DR condition with opportunities to practice conversing with the Google Home device.

During the reading session, children were encouraged to take responsibility for turning pages when the narration of a page was finished. An experimenter was present in the room, but interfered only when/if technical issues interrupted the reading. Any time a child asked a question or initiated conversation, the experimenter simply addressed the question or replied "okay," but avoided elaborating or extending the conversation. The reading sessions lasted approximately 15 min and were videotaped.

Following the reading session, children's story comprehension was assessed using a battery developed by the research team. The experimenter asked questions orally, and children responded orally to the questions or identified images presented on laminated cards. Children's answers were recorded on a paper-based checklist.

Measures

Demographic information

A parent survey was used to collect demographic information including children's date of birth (month

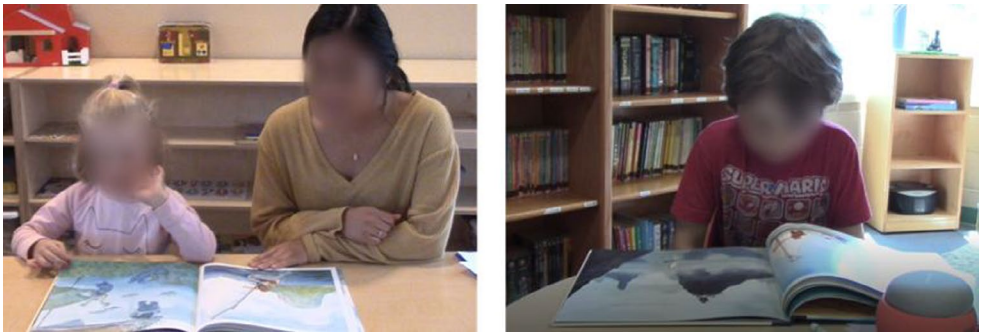

FIGURE 1 Study procedure of Human-dialogic reading (DR) condition (left) and Agent-DR condition (right). *Note*: A child participant in the Human DR condition (left) and another participant in the Agent DR condition (right; note the agent device in the lower right corner)

and year), gender, race/ethnicity, and predominant home language (i.e., English, English as a second language). This survey also asked about children's prior experience with voice technologies because this factor has been found to influence children's interactions with such technologies (Bartneck et al., 2007). If parents indicated that their child used voice technologies at least monthly, the child was categorized as a regular user of voice technologies.

Expressive vocabulary

Children's oral language skills are positively associated with children's comprehension of storybook reading activities (Kendeou et al., 2009). Children's baseline oral language skills were measured by the Expressive One Word Picture Vocabulary Test, fourth edition, which is an experimenter-administered, norm-referenced picture-naming assessment with an internal reliabilty for 3- to 6-year-olds of 0.95 (Martin & Brownell, 2011). Each child was asked to name objects, actions, and concepts that were depicted graphically. The test lasted on average 15–20 min.

Story comprehension

Children's comprehension of the storybook was measured as a proximal learning outcome. A 10-item comprehension measure was developed. These 10 questions, which were different from the ten questions asked during the dialogic reading activity, assessed children's story comprehension in three dimensions: (1) memorization of main story events, (2) inference making skills, and (3) understanding of the narrative sequence. There were five items on story event memorization and three items on inference making. For these eight items, an open-ended question was first asked, then if children could not recall the answer correctly, the research assistant provided three multiple-choice options to choose from. Two points were given to each item that was answered correctly through free recall and one point was given if answered correctly with multiple-choice options. Additionally, there were two items on narrative sequence understanding. The first one was sequence sorting, where children were asked to place four images from the book in the order they occurred in the story. Children earned three points if they correctly placed all four images in order, two points for the correct order of three images, and one point for the correct order of two images. The second prompted children to retell a part of the story. For this item, children earned one point each for mentioning each of the four key elements of a specific portion of the story in order (i.e., the four places the bears went to search to find a new seashell) for a maximum of four.

An overall story comprehension score was calculated by summing the number of points across all the items; this score was used as a dependent variable for the analysis. The range is from 0 to 23 points (16 points maximum for the 8 memorization and inference-making items, 3 points maximum for the single sequence sorting item, and 4 points maximum for the story-retelling item). Cronbach's coefficient alpha was .83 for this overall story comprehension assessment score. Three subscales of story comprehension were also calculated: a story event memorization scale ranged from 0 to 10 points, an inference making scale ranged from 0 to 6 points, and a narrative sequence understanding scale ranged from 0 to 7 points.

Engagement

Children's engagement during story listening was coded from the video-taped reading sessions. Videos were divided into 5-s segments and each segment was coded by trained researchers (Willoughby et al., 2015; Zhou & Yadav, 2017). We determined children's engagement level by analyzing the three engagement dimensions (i.e., behavioral, emotional, and cognitive) from Guthrie and Klauda's (2014) framework. Our analysis looked at these engagement dimensions via both an itemized system providing fine-grained indices of children's specific

behaviors along each separate dimension and a global coding system scoring children's overall engagement holistically. Using these two approaches allowed us to capture a more comprehensive picture of children's engagement while also establishing concurrent validity across our measures. A total of five research assistants were involved in the coding process, and the interrater-reliability (IRR) for all items was above a satisfactory level (see details below).

The itemized coding system

We coded four items on three dimensions of engagement: vocalizations (two items), affective expressions (one item), and visual attention (one item). This coding scheme was adopted from Xu et al.'s (2020) study on young children's reading engagement. For each time segment, we coded whether each item was present (score of 1 if present and 0 if not present). To calculate the proportion of time segments each item was present, we divided the number of time segments an item was present by the total number of time segments in the reading session.

Vocalizations. Children's vocalizations during each 5-s time segment of the reading episode were transcribed and coded as (1) relevant to the story content, which we call narrative-relevant (e.g., "I had lots of beautiful seashells"), and (2) irrelevant to the story content (e.g., "I want to have a snack"). Note that these vocalizations may be spontaneous or prompted by the agent or human experimenter. For each type of comment, segments received a score of 1 if the comment type was present and a score of 0 if it was absent. Every time segment was coded for both types of vocalizations, but the frequency of each type of vocalization in the segment was not coded (e.g., a score of 1 was given for narrative-relevant vocalization whether the child made one narrative-relevant comment during the segment or more than one). The IRR (Cohen's κ) was .89 for narrative-relevant vocalization and .87 for irrelevant vocalization.

Affective expressions. Affective expressions were indicated by the presence or absence of children's positive expressions during each 5-s segment. Positive expression was scored (score of 1) if the child showed at least one of the following 16 expressive 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). The IRR (Cohen's κ) was .73 for positive expression.

Visual attention. Attention was coded as children's complete *visual* attention to the book during the 5-s segment. If children maintained orientation to the book during the entire time segment, their visual attention was coded as present (score of 1). If children shifted their orientation away from the book at any point, their visual attention was coded as absent (score of 0). The IRR (Cohen's κ) for this item was .86.

Global scale of child engagement

The global scale was based on coders' broader holistic assessments of each child's engagement. For each time segment, we provided a 5-point rating based on a child's posture, facial expression, eye gaze, distractibility, verbal and nonverbal comments, and responsiveness to the adult or agent's direction (e.g., turning pages; Kaderavek et al., 2014). A score of 5 indicated the highest level of engagement (e.g., showing clear signs of excitement that stems from the reading, making large movements with hands to illustrate a point). A score of 3 indicated a medium level of engagement where a child did the minimum work required to follow protocols (e.g., listening, remaining seated). A score of 1 was the lowest level of engagement where a child was clearly distracted and had little interest in the story. An average global engagement rating was calculated by the mean of the ratings across all time segments in each child's reading session. The IRR (calculated by Intraclass Correlation because this was a numeric variable) was .82 for this global coding.

RESULTS

The results section first presents the descriptive statistics of outcome measures for the full sample (to access data, see Xu & Warschauer, 2021). The data analyses for the three research questions (i.e., effects of dialogic reading with a conversational agent on comprehension and engagement and the mechanisms of story comprehension from dialogic reading with a conversational agent) are then presented sequentially.

Descriptive statistics of outcome measures for full sample

The descriptive statistics for the full sample are presented in Table 2. Children's mean story comprehension score was 10.6, indicating that these children on average correctly answered about half of the post-test questions. In terms of attention, children on average were visually attentive to the print book about 60% of the time. In terms of emotion, children showed obvious positive expression about 11% of the time. In terms of vocalization, children across the four conditions were observed to make narrative-relevant comments in 11.0% of the time segments, while the frequency of irrelevant comments

TABLE 2 Descriptive statistics of outcome measures for full sample

	Comp	Mem	Inf	Seq	Global	RV	IRV	PE	*M*	*SD*	Range
Comprehension	1								10.4	4.39	(0, 20)
Event memorization (Mem)	.83***	1							3.07	2.12	(0, 8)
Inference making (Inf)	.81***	.55***	1						3.20	1.74	(0, 6)
Sequence understanding (Seq)	.82***	.48***	.51***	1					4.16	2.14	(0, 7)
Global engagement	.18†	.07	.20*	.18	1				3.04	0.19	(2.32, 3.71)
Relevant vocalization (RV)	.23*	.18	.22*	.18*	.46***	1			0.11	0.09	(0, 0.34)
Irrelevant vocalization (IRV)	−.21*	−.12	−.24*	−.17*	−.21*	.10	1		0.02	0.04	(0, 0.22)
Positive expression (PE)	−.02	−.07	.03	.00	.53***	.41***	.05	1	0.11	0.17	(0, 1)
Visual attention	.02	.02	.03	−.01	.16	−.31***	−.27**	−.01	0.75	0.16	(0.27, 0.98)

Note: Coefficients are Pearson correlations.
†$p < .1$.
*$p < .05$; **$p < .01$; ***$p < .001$.

(1.9%) was much lower. In terms of the global engagement rating, the average score was 3.0 across the four conditions, which represented a medium level of engagement in our coding system (1–5).

We also looked at the correlations between the outcome variables (Table 2). The overall story comprehension score was strongly correlated with the three subscales, confirming the internal consistency of this assessment. In terms of relations between comprehension and engagement, the overall story comprehension score was positively correlated with the frequency of narrative-relevant comments ($r(115) = .23$, $p < .05$) and was negatively correlated with the frequency of irrelevant comments ($r(115) = -.21$, $p < .05$). In terms of the relations between the global engagement rating and itemized coding (i.e., vocalizations, positive expression, visual attention), global engagement was positively correlated with narrative-relevant vocalization ($r(115) = .44$, $p < .001$), positive expression ($r(115) = .53$, $p < .001$), and visual attention ($r(115) = .36$, $p < .001$), and negatively correlated with irrelevant vocalization ($r(115) = -.20$, $p < .05$).

Effects of reading condition on story comprehension (RQ1)

The observed mean, standard deviation, and range of children's story comprehension scores across the four reading conditions are displayed in Table 3. To assess the effects of reading conditions on story comprehension, we fitted two regression models for each outcome. The first one is the main effect model that included the two experimental factors (i.e., dialogic reading as DR, reading partner as Agent) as the main predictor. The second one is an interaction effect model that included two experiment factors as well as the interaction between these two factors (i.e., dialogic reading × reading partner as DR × Agent). For both models, we controlled for children's baseline language proficiency, age, and whether children were regular users of conversational technologies. As shown in the left panel of Table 4, the dialogic reading factor had a significant effect on the story comprehension, $\beta = 0.51$ ($p < .001$), suggesting that dialogic reading, with a human partner or a conversational agent, led to 0.51 *SD* increase on children's story comprehension score. Reading partner (human vs. agent) did not have a significant effect on story comprehension ($\beta = -0.14$, $p = .25$). The interaction model suggested that dialogic reading with an agent induced a comparable level of positive effect on children's story comprehension as an adult reader ($\beta = 0.22$, $p = .35$). When breaking down to subscales, dialogic reading had a significant main effect on all three subscales, with the effect size being the larger for event memorization ($\beta = 0.53$, $p < .001$) as compared to inference making ($\beta = 0.38$, $p < .05$) and sequence understanding ($\beta = 0.34$, $p < .05$).

Based on the interaction models used above, we calculated the estimated marginal means of each total comprehension score and subscales across the four reading conditions, as shown in the right panel of Table 4. We conducted pairwise comparisons on these estimated marginal mean scores. In particular, the estimated total comprehension score of children in the Agent DR condition was substantially higher than those in the Agent DR condition ($p < .001$) and those in the Human non-DR condition ($p = .06$). The difference in the estimated scores between the Agent DR and Human DR groups was very small ($p = .78$).

Possible interaction effects

We also examined whether the effects of dialogic reading varied based on children's age and language proficiency. We did so by interacting children's expressive vocabulary score and age, respectively, with our two experimental factors (i.e., dialogic reading and reading partner) as well as with the products between them (i.e., dialogic reading × reading partner). Regarding language proficiency, as shown in Table 5 Model 2, the

TABLE 3 Observed outcome measures by condition

	Agent DR	Agent non-DR	Human DR	Human non-DR
Story comprehension				
M (*SD*)	11.6 (4.64)	8.67 (4.90)	11.2 (5.32)	9.88 (4.48)
Median [min, max]	12 [2, 20]	7 [1, 17]	11 [0, 20]	9.50 [2, 19]
Event memorization				
M (*SD*)	3.42 (2.17)	2.33 (1.86)	3.61 (2.22)	2.73 (2.03)
Median [min, max]	3 [0, 7]	2 [0, 7]	3 [0, 8]	3 [0, 7]
Inference making				
M (*SD*)	3.58 (1.56)	2.74 (1.83)	3.42 (1.80)	2.92 (1.74)
Median [min, max]	4 [1, 6]	3 [0, 6]	4 [0, 6]	3 [0, 6]
Sequence understanding				
M (*SD*)	4.61 (1.84)	3.59 (2.37)	4.13 (2.26)	4.23 (2.07)
Median [min, max]	4 [1, 7]	4 [0, 7]	4 [0, 7]	4 [0, 7]
Global engagement				
M (*SD*)	3.09 (0.16)	2.96 (0.17)	3.03 (0.20)	3.05 (0.21)
Median [min, max]	3.06 [2.76, 3.52]	2.98 [2.53, 3.26]	3.07 [2.32, 3.35]	3.04 [2.54, 3.71]
Relevant vocalization				
M (*SD*)	0.13 (0.04)	0.03 (0.06)	0.18 (0.07)	0.09 (0.11)
Median [min, max]	0.12 [0, 0.24]	0 [0, 0.23]	0.16 [0.02, 0.32]	0.04 [0, 0.34]
Irrelevant vocalization				
M (*SD*)	0.01 (0.01)	0.01 (0.02)	0.02 (0.04)	0.04 (0.06)
Median [min, max]	0 [0, 0.06]	0 [0, 0.08]	0.01 [0, 0.15]	0.01 [0, 0.22]
Positive expression				
M (*SD*)	0.14 (0.21)	0.10 (0.16)	0.10 (0.12)	0.09 (0.15)
Median [min, max]	0.05 [0, 1]	0.01 [0, 0.55]	0.04 [0, 0.39]	0.04 [0, 0.69]
Visual attention				
M (*SD*)	0.74 (0.13)	0.79 (0.15)	0.70 (0.17)	0.76 (0.17)
Median [min, max]	0.75 [0.43, 0.96]	0.83 [0.44, 0.98]	0.74 [0.27, 0.93]	0.82 [0.48, 0.98]

Abbreviation: DR, dialogic reading.

coefficients of the three interaction products with expressive vocabulary were not significant, and inclusion of these interaction terms led to a negligible change of variance explained by the model (Change of R^2 = .00, p = .66). Regarding age, as shown in Table 5 Model 3, the interaction product between age and dialogic reading (i.e., DR × Age) was also not significant (DR × Age: β = 0.42, p = .07), though the moderate effect size suggested that the benefits of dialogic reading might be larger for the older children in this sample. However, the inclusion of all three interaction products only increased the R^2 by .01, which was not significant (p = .38). Overall, this set of analyses failed to confirm that the effects of dialogic reading with an agent varied by children's language ability or age.

Effects of reading condition on engagement (RQ2)

The descriptive statistics of children's reading engagement (i.e., observed means, *SD*, and range) are displayed in Table 3. Similar to our analysis on story comprehension, we fitted one main effect and one interaction effect regression model for each of the engagement outcomes. The results are displayed in the left panel of Table 6. In terms of the global engagement rating, the dialogic reading led to a significant main effect (β = 0.41, p < .05), but reading partner did not (β = 0.00, p = .99). However, the effect of agent partner on engagement appeared to be dependent upon whether children were engaged in dialogic reading (DR × Agent: β = 0.64, p = .08), meaning that having the dialogic interaction component enhanced the engagement level of those reading with an agent. In terms of vocalization, dialogic reading led to a significantly higher level of narrative-related vocalization (β = 1.11, p < .001), and reading with an agent was associated with a decreased level of narrative-relevant vocalization (β = −0.54, p < .001). Reading with an agent resulted in less irrelevant vocalization (β = −0.63, p < .001). The interaction model also suggested that when reading with a human partner, dialogic reading appeared to help reduce the instances of irrelevant vocalization (β = −0.50, p = .06). The two experimental

TABLE 4 Regression results and estimated marginal means of story comprehension scales by condition

	Regression β (*SE*)			Marginal means by condition *M* (*SE*)			
	DR	**Agent**	**DR × Agent**	**Agent DR**	**Agent non-DR**	**Human DR**	**Human non-DR**
Story comprehension							
Main	**0.51 (.13)*****	−0.14 (.12)					
Int.	**0.38 (.18)***	−0.27 (.18)	0.22 (.24)	0.22 (.14)$_{a}$	−0.39 (.14)$_{b}$	0.26 (.14)$_{a}$	−0.12 (.15)$_{ab}$
Event memorization							
Main	**0.53 (.14)*****	**−0.23 (.14)**†					
Int.	**0.55 (.20)****	−0.21 (.21)	−0.03 (.27)	0.07 (.16)$_{ac}$	−0.45 (.16)$_{b}$	0.31 (.15)$_{ac}$	−0.24 (.17)$_{ab}$
Inference making							
Main	**0.38 (.16)***	−0.10 (.15)					
Int.	0.26 (.23)	−0.23 (.23)	0.22(.30)	0.15 (.18)$_{a}$	−0.33 (.18)$_{b}$	0.16 (.17)$_{a}$	−0.10 (.19)$_{a}$
Sequence understanding							
Main	**0.34 (.16)***	−0.02 (.15)					
Int.	0.14 (.23)	−0.23 (.24)	0.37 (.31)	0.32 (.19)$_{a}$	−0.19 (.21)$_{b}$	0.17 (.21)$_{a}$	0.04 (.24)$_{a}$

Note: All coefficients and estimated marginal means are standardized. "Main" refers to the regression model that includes two experimental factors as predictors (dialogic reading as DR and reading partner as Agent). "Int." refers to the interaction model that includes the two experimental factors as well as the interaction term between them (DR × Agent). For all regression models, covariates included age, expressive vocabulary, and prior usage of agents. Experimenter fixed effects included to adjust any potential confounding introduced by the experimenters. Regression coefficients with *p* values less than .1 are in bold. Pairwise comparisons with Tukey adjustments were conducted to examine the significant differences between the estimated marginal means of each two conditions. Means in the same row that do not share subscripts differ at $p < .05$.

†$p < .1$.

*$p < .05$; **$p < .01$; ***$p < .001$.

TABLE 5 Regression analysis of the condition effects and interaction effects on story comprehension

	Model 1	Model 2	Model 3
DR	**.38 (.18)***	**.40 (.19)***	**.40 (.20)***
Agent	−.24 (.19)	−.26 (.18)	−.23 (.20)
DR × Agent	.22 (.24)	.20 (.24)	.19 (.25)
DR × Expressive Vocab		.23 (.21)	—
Agent × Expressive Vocab		.10 (.19)	—
DR × Agent × Expressive Vocab		−.14 (.26)	—
DR × Age		—	**.41 (.22)**†
Agent × Age		—	.00 (.18)
DR × Agent × Age		—	−.17 (.27)
R^2	.63	.63	.64

Note: Standardized coefficients reported. Standard errors in parentheses. Regression coefficients with *p* values less than .1 are in in bold. For all regression models, covariates included age, expressive vocabulary, and prior usage of agents. Experimenter fixed effects included to adjust any potential confounding introduced by the experimenters.

†$p < .1$.

*$p < .05$.

factors, nor their interactions, had significant effects on positive affect and visual attention.

The marginal means of engagement measures across the four experimental conditions are shown in the right panel of Table 6. A series of pairwise comparisons were conducted to examine the significance of differences between each pair of conditions. In particular, the Agent DR condition led to a significantly higher rating than the Agent non-DR condition ($p < .05$) but did not significantly differ from the other two conditions involving human partners.

Mediating effects of engagement on story comprehension (RQ3)

Finally, we conducted structural equation modeling (SEM) to formally test whether engagement explains the effect of reading condition on story comprehension. Given that the narrative-relevant and irrelevant vocalizations are the only two coded variables significantly correlated with comprehension (see Table 2), we specifically focused on these two variables in the SEM analysis. This choice was also supported by the rationale of the purpose of dialogic reading, which is to increase the amount of vocalization (Hargrave & Sénéchal, 2000). We used the experimental condition as a four-level categorical predictor. Specifically, we converted this condition variable into three dummy variables (i.e., Agent non-DR, Human DR, Human non-DR) with Agent DR condition as the omitted reference group. Agent DR was chosen as the reference group because it was the central group of interest in the study and understanding the possible differences in the mechanisms through which Agent DR sessions are effective was also a study goal.

With this rationale in mind, we fitted a model with narrative-relevant and irrelevant vocalizations as

TABLE 6 Regression results and estimated marginal means of engagement scales by condition

	Regression β (*SE*)			Marginal means by condition *M* (*SE*)			
	DR	**Agent**	**DR × Agent**	**Agent DR**	**Agent non-DR**	**Human DR**	**Human non-DR**
Global engagement							
Main	**0.41 (.20)***	0.00 (.19)					
Int.	0.06 (.28)	−0.36 (.28)	**0.64 (.36)†**	0.25 (.22)$_a$	−0.45 (.22)$_b$	−0.03 (.21)$_a$	−0.09 (.23)$_a$
Relevant vocalization							
Main	**1.11 (.15)*** **	**−0.54 (.15)*** **					
Int.	**1.07 (.22)*** **	**−0.59 (.22)** **	0.08 (.29)	0.53 (.18)$_{ac}$	−0.62 (.20)$_b$	1.04 (.20)$_c$	−0.03 (.22)$_a$
Irrelevant vocalization							
Main	−0.25 (.19)	**−0.63 (.18)*** **					
Int.	**−0.50 (.27)†**	**−0.89 (.27)** **	0.45 (.36)	−0.02 (.22)$_a$	0.03 (.25)$_a$	0.42 (.24)$_{ab}$	0.92 (.27)$_b$
Positive expression							
Main	.23 (.20)	0.22 (.19)					
Int.	.18 (.29)	0.17 (.29)	0.09 (.38)	0.35 (.23)$_a$	0.08 (.26)$_a$	0.09 (.26)$_a$	−0.09 (.29)$_a$
Visual attention							
Main	**−0.37 (.19)†**	0.20 (.19)					
Int.	−0.40 (.28)	0.17 (.28)	0.06 (.37)	−0.36 (.23)$_a$	−0.02 (.26)$_a$	−0.59 (.25)$_a$	−0.19 (.28)$_a$

Note: All coefficients and estimated marginal means are standardized. "Main" refers to the regression model that includes two experimental factors as predictors (dialogic reading as DR and reading partner as Agent). "Int." refers to the interaction model that includes the two experimental factors as well as the interaction term between them (DR × Agent). For all regression models, Covariates including age, expressive vocabulary, and prior usage of agents. Experimenter fixed effects included to adjust any potential confounding introduced by the experimenters. Regression coefficients with *p* values less than .1 are in bold. Pairwise comparisons with Tukey adjustments were conducted to examine the significant differences between the estimated marginal means of each two conditions. Means in the same row that do not share subscripts differ at $p < .05$.

†$p < .1$

*$p < .05$; **$p < .01$; ***$p < .001$.

mediators between the different group assignments and our outcome, comprehension. Our model specification included all three groups having direct paths to the outcome, as well as indirect paths through vocalizations to the outcome (see Figure 2). Three covariates in the regression analysis above—participant age, expressive vocabulary score, and prior experience with conversational technologies—were also included. This model has a great fit ($\chi^2(1) = .74$, $p = .39$, comparative fit index = 1.00, Tucker–Lewis index = 1.03, root mean square error of approximation = .00 [.00, .23], standardized root mean square residual = .01), according to Keith (2014). The path coefficients are displayed in Table 7.

Children's group assignment had differing relationships to each of the mediators in comparison to the Agent DR reference group, consistent with our analysis on engagement. Participants in the Human DR group had, on average, higher rates of narrative-relevant vocalizations ($\beta = 0.55$, $p < .01$) than the Agent DR group, while children in the Agent non-DR group had, on average, substantially lower rates of narrative-relevant comments ($\beta = -1.12$, $p < .001$). Children in the Human non-DR group had, on average, higher rates of irrelevant comments ($\beta = 0.96$, $p < .001$). As for the relations between mediators and story comprehension, the narrative-relevant vocalizations mediator was positively associated with the outcome ($\beta = 0.18$, $p < .05$), and the irrelevant vocalizations mediator was negatively associated with the outcome ($\beta = -0.16$, $p < .05$). In terms of the direct paths from reading conditions to the outcome, there was a marginally significant direct path from Agent non-DR to the story comprehension score ($\beta = -0.34$, $p = .06$), suggesting that children in the Agent non-DR condition may have had lower learning performance compared to those in the Agent DR condition while controlling for irrelevant and relevant vocalization mediators and other covariates.

We also calculated the indirect effects from the condition assignments to story comprehension through mediators using R's Lavaan package (Gana & Broc, 2019). We focused on the two non-dialogic conditions (i.e., Agent non-DR and Human non-DR) given that these two conditions were found to have lower comprehension scores than the reference group, Agent-DR condition. The analysis of indirect effects could point to a mechanism by which the significant differences in story comprehension can be explained. In terms of the Agent non-DR condition, this group's lower comprehension score compared to the Agent DR group could be partially explained by children's lower level of narrative-relevant vocalizations, as there was a significant, indirect path from Agent non-DR group assignment through narrative-relevant vocalization ($\beta = -0.20$, $p < .05$). In terms of the Human

non-DR condition, this group's lower comprehension score compared to the Agent DR condition could be explained by the Human non-DR group's higher rates of irrelevant vocalizations compared to the Agent DR condition, as the indirect effect from Human non-DR through irrelevant vocalization is negative and significant ($\beta = -0.16$, $p < .05$).

DISCUSSION

The purpose of this study was to examine the effects of dialogic reading with a disembodied conversational agent versus an adult on children's reading engagement and story comprehension. Dialogic reading, during which children are read a storybook and engaged in relevant conversation, has long been viewed as an ideal context to foster children's early language and literacy development. Our study demonstrated that a properly designed conversational agent can assume the role of a dialogue partner during children's storybook reading with benefits comparable to that of an adult dialogue partner. Given that smart speakers are affordable and already owned by many families, these findings are promising for the deployment of this technology in supporting children's language development, especially for children from families who may have limited time, language skills, or resources to themselves engage in dialogic reading.

Our first research question examined the effects of dialogic reading and conversational social agents on children's story comprehension. Consistent with prior research (Flack et al., 2018; Mol et al., 2008; Noble et al., 2019; Towson et al., 2017), we found that children who listened to a story together with dialogue outperformed those who listened to the story without dialogue. This validated the design of dialogic strategies (i.e., the questions and feedback) used in our study. Furthermore, our results suggest that the conversational agent replicated the benefits of dialogue with an adult partner, given that the effects of dialogic reading did not vary by dialogue with an adult or the agent. This is in line with the emerging body of research demonstrating the potential benefits of artificially intelligent learning companions. However, in contrast to prior research on these benefits that typically involved robots (e.g., Breazeal et al., 2016; Westlund et al., 2017), the conversational agent used in our study was disembodied and thus not capable of utilizing non-verbal expressions to facilitate the dialogue. That this agent, with only a voice interface, can benefit children's story comprehension as much as face-to-face human partners reinforces the importance of verbal dialogue in promoting children's language skills laid out in Vygotsky's (2012) theory. This might be especially true in the context of dialogic reading (cf. Lever & Sénéchal, 2011).

Our findings also illustrate the relative ineffectiveness of simply listening to a story read by a digital agent *without* any interaction, as children in the Agent non-DR group had the lowest comprehension and engagement scores among the four experimental conditions. This highlights the importance of future research and development to focus on the provision of high-quality interactivity in digital reading.

We did not detect a significant interaction between the children's baseline language proficiency and the effects of dialogic reading with an agent on their story comprehension. While the non-significant interaction effect may suggest the possible robustness of our results across subgroups with varying language proficiency, it may also result from the lower proficiency children in our sample being within the norm for their chronological age. Specifically, the median age-adjusted expressive vocabulary score was 115, which is equivalent to an 83rd percentile rank among the national, normative population and the score of the first quartile was 103, which is still above the 50th percentile rank. As such, the homogeneous high language proficiency of this sample may

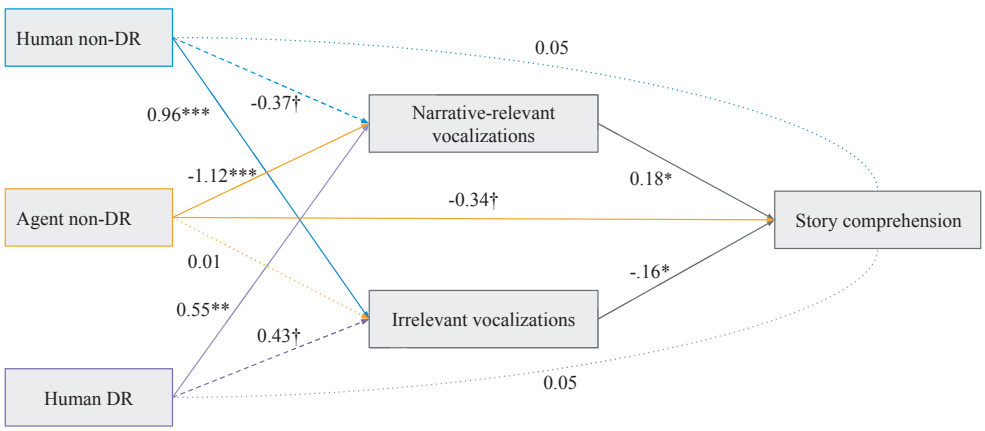

FIGURE 2 Structural equation modeling analysis of reading condition, vocalizations, and story comprehension. *Note*: Solid lines are statistically significant paths, dashed lines are marginally significant paths, and dotted lines are non-significant paths. DR, dialogic reading. †$p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$

TABLE 7 Results of structural equation modeling

Dependent variable		Independent variable	Coefficient	*SE*	*p* Value
Direct paths					
Comprehension	←	Agent non-DR	**−.34**†	(.18)	.06
Comprehension	←	Human non-DR	.05	(.19)	.83
Comprehension	←	Human DR	.05	(.17)	.73
Comprehension	←	Relevant Voc.	**.18***	(.07)	<.05
Comprehension	←	Irrelevant Voc.	**−.16***	(.07)	<.05
Comprehension	←	Expressive Vocab	**.55*****	(.07)	<.001
Comprehension	←	Age	**.26*****	(.07)	<.001
Comprehension	←	Prior CA use	−.19	(.13)	.13
Relevant Voc.	←	Agent non-DR	**−1.12*****	(.21)	<.001
Relevant Voc.	←	Human non-DR	**−.37**†	(.23)	.10
Relevant Voc.	←	Human DR	**.55****	(.20)	<.01
Relevant Voc.	←	Expressive Vocab	.13	(.09)	.13
Relevant Voc.	←	Age	−.13	(.09)	.16
Relevant Voc.	←	Prior CA use	.16	(.15)	.31
Irrelevant Voc.	←	Agent non-DR	.02	(.24)	.95
Irrelevant Voc.	←	Human non-DR	**.96*****	(.26)	<.001
Irrelevant Voc.	←	Human DR	**.43**†	(.23)	.06
Irrelevant Voc.	←	Expressive Vocab	−.09	(.10)	.34
Irrelevant Voc.	←	Age	.05	(.11)	.66
Irrelevant Voc.	←	Prior CA use	**−.34**†	(.18)	.05
Indirect paths					
Comprehension	← Relevant Voc.	← Agent non-DR	**−.20***	(.09)	<.05
Comprehension	← Irrelevant Voc.	← Agent non-DR	−.00	(.04)	.95
Comprehension	← Relevant Voc.	← Human non-DR	−.07	(.05)	.17
Comprehension	← Irrelevant Voc.	← Human non-DR	**−.16***	(.08)	<.05
Comprehension	← Relevant Voc.	← Human DR	**.10**†	(.05)	.07
Comprehension	← Irrelevant Voc.	← Human DR	−.07	(.05)	.14

Note: Standardized coefficient presented. Coefficients with *p* values less than .1 are in bold. Standard errors in parentheses.
Abbreviations: CA, conversational agent; DR, dialogic reading.
†$p < .1$.
*$p < .05$; **$p < .01$; ***$p < .001$.

have obscured our ability to uncover the heterogenous effects of dialogic reading with conversational agents. In addition, though age did not moderate comprehension, the effect size of the interaction between age and dialogic reading was in the moderate range (0.42). This points to the possibility that older children in our sample benefited from dialogic reading to a greater extent than younger children and we were underpowered to detect that effect. It is conceivable that older children are more adept at participating in dialogic interaction, particularly with digital agents (Xu & Warschauer, 2020c), and thus gain more from the interaction.

We also uncovered the effects of dialogic reading and conversational agents on children's engagement. An interesting pattern emerged in terms of global engagement. Non-dialogic reading with an agent is detrimental to children's overall engagement. However, dialogue with an agent increases children's engagement to the levels found when children read with a human. This finding provides empirical support for the notion that opportunities for contingent dialogue with agents may simulate the social presence of a human partner and bring about similar benefits for engaged learning (cf. Brunick et al., 2016).

When examining vocalizations, as expected, dialogic reading resulted in significantly higher levels of narrative-relevant vocalization. This suggests that children were receptive to dialogic reading, as demonstrated repeatedly from studies in both face-to-face settings and computer-based environments (e.g., Calvert et al., 2019; Peebles et al., 2018). Interestingly, it appeared that dialogic reading also reduced the instances of irrelevant vocalizations that may be an indicator of distraction (Reich et al., 2019). This may be because dialogic reading "directs" children's vocalizations along the narrative, thus helping children focus on

the reading. Overall, children that were partnered with a conversational agent did not generate vocalizations, either narrative-relevant or irrelevant, as frequently as those reading with an adult. This finding of fewer child vocalizations with an agent was consistent with Aeschlimann et al. (2020), who found that preschool-aged children were less likely to provide vocal information to a smart speaker than to an adult researcher. There are two possible explanations for this: either children are less knowledgeable about how to talk to non-human agents (Beneteau et al., 2019; Cheng et al., 2018) or they are less interested in doing so (Cameron et al., 2015). Our findings suggest that the social presence of a human partner may encourage children to provide on-topic responses but may also invite children to voluntarily extend the conversation beyond the reading context. Though bringing up irrelevant comments may be developmentally appropriate for young children (Godwin et al., 2016) and generate excitement (Xu & Warschauer, 2020c), our study suggests that doing so may shift children's attention away from the story and dampen learning.

In our analysis, dialogue did not enhance emotional engagement, as there was no significant difference by condition in the frequency of children's positive expressions. This could be due to the nature of the dialogue included in this particular reading activity. The agent's questions primarily asked about specific content in the story, and the agent used matter-of-fact word choices to respond to children (i.e., the agent told children if their answer was correct or not and then gave an explanation). This direct approach may not be typical of how adults interact with children and could have limited the agent's ability to elicit positive emotional responses. Our analysis also suggested that dialogue did not lead to a significantly higher level of visual attention during reading, while other studies suggested that children more frequently fixated on the educational content displayed on the screen when an adult co-viewer commented on the content (Neuman et al., 2019). However, in Neuman et al., the comments were not designed to elicit children's verbal responses, but rather to label and explain the vocabulary. As such, we speculated that the dialogue moments in our study, which elicited verbal responses, may have triggered children to look at their reading partner (either the smart speaker or adult) as they replied to the questions and listened to feedback, thus deviating the children's visual attention from the book. To test this speculation, we recalculated children's visual attention by including their time spent looking at their respective conversational partner. However, the new visual attention variable remained consistent with the original one: Children in the dialogic conditions still had a relatively lower level of visual attention than non-dialogic conditions. Specifically, Agent DR condition had 0.76 of the time looking at the book or the agent ($SD = 0.13$), Agent non-DR condition were attentive for 0.80 of the time ($SD = 0.15$), Human DR condition for 0.70 ($SD = 0.18$), and Human non-DR for 0.76 ($SD = 0.17$). As such, it was not evident that the reduced visual attention to the book in the dialogic conditions was attributed to children looking at their conversational partner. Nevertheless, we did notice from our video recordings that children shifted their eyes away from the book and looked straight ahead when they were thinking hard to formulate their responses. Supporting this observation, Table 2 shows that instances of visual attention were negatively correlated with both narrative-relevant vocalization ($r = -.31$, $p < .001$) and irrelevant vocalization ($r = -.27$, $p < .01$). While many studies have shown a significant positive correlation between children's fixation on the book and their learning (Justice et al., 2008), our findings suggest the importance of a holistic view in understanding children's visual attention and engagement during conversation-rich reading activities, as children may gaze away from the book or interlocutor when formulating a response.

Our mediation analysis corroborated the effects of condition on children's vocalization and points to interesting mechanisms through which dialogic reading with conversational agents may support language development. The advantage of dialogic reading with conversational agents is explained through a two-pronged mechanism: increased narrative-relevant vocalizations (compared to "non-DR" groups) and decreased irrelevant vocalizations (compared to the "Human" groups). The first part of this finding replicates Calvert et al. (2019), which indicated that asking children questions during televised stories promoted learning because of children's increased relevant talk. The second part of the finding suggests that agents can enhance learning through limiting off-task behaviors. However, the covariates we selected for our SEM models were based on significant correlations and model fit. This practice capitalizes on chance fluctuations of our data and may limit the generalizability of the results.

Taken together, our study provides evidence that disembodied conversational agents can effectively engage children in dialogic reading activities. At an applied level, these findings suggest we may take advantage of the prevalence of smart speakers in children's homes and integrate these devices as part of children's informal learning experiences. While we do not recommend that artificial intelligence replaces children's story time with their parents or teachers, properly designed agents may sometimes play the role of an engaging dialogic partner for children when adults are unavailable. Despite these possibilities, more studies are needed to better understand how conversational agents can support high-quality interaction and lead to more fruitful learning (see our discussion below).

Limitations and future directions

There are several directions for future empirical studies to build on the current findings. First, the current study

was carried out in a controlled manner where children had dialogic reading with scripted questions and feedback. Though this design increased the internal validity of the study by holding the conversation consistent between human and agent, it limited the ecological validity of the findings. Future studies could be carried out in a more naturalistic setting, in which a familiar adult reads with the child as they normally would. We would expect variation in how much and how well dialogic questions and feedback were utilized by the adult. As such, we may compare virtual agents against skilled and unskilled human partners who are not constrained to a script.

Second, as discussed above, our sample was limited in both size and scope. The children who participated in our study are from higher socioeconomic backgrounds. Future research should investigate whether dialogic reading with an agent can help at risk children from lower socioeconomic backgrounds. These at-risk children may lag in language and literacy development, potentially making dialogic scaffolding particularly valuable for them. More broadly, our limited sample size and associated lack of power may have manifested itself in the marginally significant coefficients both in regression and SEM analyses, respectively. This may have also limited our ability to detect heterogeneous effects among children sub-groups. Replicating this study with a larger and more diverse sample will allow researchers to investigate whether dialogic reading with agents may have different impacts among children with different language and socioeconomic backgrounds.

Third, our study focused on immediate outcomes after a one-time, short intervention, while future research may want to implement the agent dialogic reading partner for a longer period of time at schools, public libraries, or homes. Given that other longer term dialogic reading interventions (typically lasting 4–8 weeks) have proven successful in promoting children's general language ability, such as receptive and expressive language, vocabulary, and narrative skills, it is plausible that agent-based interventions could also bring similar benefits to children.

Fourth, additional studies should focus on the design of conversational agents. For example, it is important to explore whether the persona of the agent has an impact on learning and engagement. In this study, the conversational agent assumes the role of a knowledgeable adult, yet other studies suggest that children may feel more encouraged to express their ideas when talking to a peer (Zaga et al., 2015). Future studies could compare the impact of different persona designs on children's engagement and learning. Moreover, the dialogic reading literature suggests making story narratives more relatable to children by asking them questions that connect to their personal experiences (Flynn, 2011), for example, "The bears in the story went on a boat ride. Have you ever been on a boat ride before? Was that fun?". Future studies should examine how children respond to personal questions asked by conversational agents and how these questions impact engagement and comprehension. These studies could provide practical design suggestions for the future development of conversational agents.

Finally, we have thus far only considered the role of a conversational agent as a replacement for, rather than as a complement of, a human partner. Study designs that compare non-dialogic and dialogic agents with and without parents present can help shed light on whether CAs supplant adult–child interaction or potentially model and stimulate enhanced adult–child interaction—and what kind of CA designs lead to the latter outcome rather than the former. Given how valuable adult-child interaction is in helping children learn to read, this will be an important area of future research, and one we intend to undertake.

CONCLUSION

This study examined whether and how a smart speaker-based conversational agent can facilitate language development by engaging children in dialogic reading. Our findings suggest that dialogic reading with a disembodied conversational agent can replicate the benefits of an adult partner in facilitating story comprehension. Furthermore, we found that the benefits of dialogic reading with an agent arose from children's increased narrative-relevant vocalizations and decreased irrelevant vocalizations. Given that disembodied conversational agents are already affordable and prevalent, such agents represent a potentially scalable, cost-effective tool for enriching preschool-aged children's early literacy development.

Nevertheless, building conversational agents for young children is a complex endeavor. To maximize the benefits of conversational agents, it is vital to consider children's developing cognitive abilities and specific communication needs. By drawing on well-established research in child development and supplementing it with the growing new genre of research on child–agent interaction, researchers, developers, and educators will be in position to take a proactive, evidence-driven approach to the development and evaluation of conversational agents as children's social learning partners.

CONFLICT OF INTEREST

We have not known conflict of interest to disclose.

ORCID

Ying Xu https://orcid.org/0000-0003-2079-2394
Andres S. Bustamante https://orcid.org/0000-0002-7678-816X
Young-Suk Kim https://orcid.org/0000-0002-4328-3843

REFERENCES

Aeschlimann, S., Bleiker, M., Wechner, M., & Gampe, A. (2020). Communicative and social consequences of interactions with voice assistants. *Computers in Human Behavior*, *112*, 106466. https://doi.org/10.1016/j.chb.2020.106466

Ann Evans, M., & Saint-Aubin, J. (2005). What children are looking at during shared storybook reading: Evidence from eye movement monitoring. *Psychological Science*, *16*, 913–920. https://doi.org/10.1111/j.1467-9280.2005.01636.x

Arnold, D. H., Lonigan, C. J., Whitehurst, G. J., & Epstein, J. N. (1994). Accelerating language development through picture book reading: Replication and extension to a videotape training format. *Journal of Educational Psychology*, *86*, 235. https://doi.org/10.1037/0022-0663.86.2.235

Arnold, D. S., & Whitehurst, G. J. (1994). Accelerating language development through picture book reading: A summary of dialogic reading and its effect. In D. K. Dickinson (Ed.), *Bridges to literacy: Children, families, and schools* (pp. 103–128). Blackwell Publishing.

Bai, S., Repetti, R. L., & Sperling, J. B. (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. https://doi.org/10.1037/a0039854

Bartneck, C., Suzuki, T., Kanda, T., & Nomura, T. (2007). The influence of people's culture and prior experiences with Aibo on their attitude towards robots. *Ai & Society*, *21*, 217–230. https://doi.org/10.1007/s00146-006-0052-7

Beneteau, E., Guan, Y., Richards, O. K., Zhang, M. R., Kientz, J. A., Yip, J., & Hiniker, A. (2020). Assumptions checked: How families learn about and use the echo dot. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *4*, 1–23. https://doi.org/10.1145/3380993

Beneteau, E., Richards, O. K., Zhang, M., Kientz, J. A., Yip, J., & Hiniker, A. (2019, May). Communication breakdowns between families and Alexa. In *Proceedings of the 2019 CHI conference on human factors in computing systems* (pp. 1–13). https://doi.org/10.1145/3290605.3300473

Berk, L. E., & Winsler, A. (1995). *Scaffolding children's learning: Vygotsky and early childhood education. NAEYC research into practice series* (Vol. *7*). National Association for the Education of Young Children.

Blewitt, P., Rump, K. M., Shealy, S. E., & Cook, S. A. (2009). Shared book reading: When and how questions affect young children's word learning. *Journal of Educational Psychology*, *101*, 294–304. https://doi.org/10.1037/a0013844

Bodrova, E., & Leong, D. J. (2005). High quality preschool programs: What would Vygotsky say? *Early Education and Development*, *16*, 435–444. https://doi.org/10.1207/s15566935eed1604_4

Breazeal, C., Harris, P. L., DeSteno, D., Kory Westlund, J. M., Dickens, L., & Jeong, S. (2016). Young children treat robots as informants. *Topics in Cognitive Science*, *8*, 481–491. https://doi.org/10.1111/tops.12192

Brunick, K. L., Putnam, M. M., McGarry, L. E., Richards, M. N., & Calvert, S. L. (2016). Children's future parasocial relationships with media characters: The age of intelligent characters. *Journal of Children and Media*, *10*, 181–190. https://doi.org/10.1080/17482798.2015.1127839

Calvert, S. L., Putnam, M. M., Aguiar, N. R., Ryan, R. M., Wright, C. A., Liu, Y. H. A., & Barba, E. (2019). Young children's mathematical learning from intelligent characters. *Child Development*. https://doi.org/10.1111/cdev.13341

Cameron, D., Fernando, S., Collins, E., Millings, A., Moore, R., Sharkey, A., Evers, V., & Prescott, T. (2015, January). Presence of life-like robot expressions influences children's enjoyment of human-robot interactions in the field. In *Proceedings of the AISB convention 2015*, The Society for the Study of Artificial Intelligence and Simulation of Behaviour.

Cheng, Y., Yen, K., Chen, Y., Chen, S., & Hiniker, A. (2018, June). Why doesn't it work? Voice-driven interfaces and young children's communication repair strategies. In *Proceedings of the 17th ACM conference on interaction design and children* (pp. 337–348). https://doi.org/10.1145/3202185.3202749

Farver, J. A. M., Xu, Y., Lonigan, C. J., & Eppe, S. (2013). The home literacy environment and Latino head start children's emergent literacy skills. *Developmental Psychology*, *49*, 775. https://doi.org/10.1037/a0028766

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*, 1334. https://doi.org/10.1037/dev0000512

Flynn, K. S. (2011). Developing children's oral language skills through dialogic reading: Guidelines for implementation. *Teaching Exceptional Children*, *44*, 8–16. https://doi.org/10.1177/005405991104400201

Gana, K., & Broc, G. (2019). *Structural equation modeling with lavaan*. John Wiley & Sons.

Garg, R., & Sengupta, S. (2020, April). Conversational technologies for in-home learning: Using co-design to understand children's and parents' perspectives. In *Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1–13). https://doi.org/10.1145/3313831.3376631

Godwin, K. E., Almeda, M. V., Seltman, H., Kai, S., Skerbetz, M. D., Baker, R. S., & Fisher, A. V. (2016). Off-task behavior in elementary school children. *Learning and Instruction*, *44*, 128–143. https://doi.org/10.1016/j.learninstruc.2016.04.003

Golinkoff, R. M., Hoff, E., Rowe, M. L., Tamis-LeMonda, C. S., & Hirsh-Pasek, K. (2019). Language matters: Denying the existence of the 30-million-word gap has serious consequences. *Child Development*, *90*, 985–992. https://doi.org/10.1111/cdev.13128

Guthrie, J. T., & Klauda, S. L. (2014). Effects of classroom practices on reading comprehension, engagement, and motivations for adolescents. *Reading Research Qsuarterly*, *49*, 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*, 75–90. https://doi.org/10.1016/S0885-2006(99)00038-1

Hong, Z. W., Huang, Y. M., Hsu, M., & Shen, W. W. (2016). Authoring robot-assisted instructional materials for improving learning performance and motivation in EFL classrooms. *Journal of Educational Technology & Society*, *19*, 337–349. https://www.jstor.org/stable/jeductechsoci.19.1.337

John-Steiner, V., & Mahn, H. (1996). Sociocultural approaches to learning and development: A Vygotskian framework. *Educational Psychologist*, *31*, 191–206. https://doi.org/10.1080/00461520.1996.9653266

Justice, L. M., Pullen, P. C., & Pence, K. (2008). Influence of verbal and nonverbal references to print on preschoolers' visual attention to print during storybook reading. *Developmental Psychology*, *44*, 855. https://doi.org/10.1037/0012-1649.44.3.855

Kaderavek, J. N., Guo, Y., & Justice, L. M. (2014). Validity of the children's orientation to book reading rating scale. *Journal of Research in Reading*, *37*(2), 159–178.

Kaefer, T., Pinkham, A. M., & Neuman, S. B. (2017). Seeing and knowing: Attention to illustrations during storybook reading and narrative comprehension in 2-year olds. *Infant and Child Development*, *26*, e2018. https://doi.org/10.1002/icd.2018

Keith, T. Z. (2014). *Multiple regression and beyond: An introduction to multiple regression and structural equation modeling*. Routledge.

Kendeou, P., Van den Broek, P., White, M. J., & Lynch, J. S. (2009). Predicting reading comprehension in early elementary school: The independent contributions of oral language and decoding skills. *Journal of Educational Psychology*, *101*, 765. https://doi.org/10.1037/a0015956

Kory, J., & Breazeal, C. (2014, August). *Storytelling with robots: Learning companions for preschool children's language development*. In The 23rd IEEE international symposium on robot and human interactive communication (pp. 643–648). IEEE. https://doi.org/10.1109/ROMAN.2014.6926325

Lee, K. M., Jung, Y., Kim, J., & Kim, S. R. (2006). Are physically embodied social agents better than disembodied social agents?: The effects of physical embodiment, tactile interaction, and people's loneliness in human–robot interaction. *International Journal of Human- Computer Studies*, *64*, 962–973. https://doi.org/10.1016/j.ijhcs.2006.05.002

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

Lovato, S. B., & Piper, A. M. (2019). Young children and voice search: What we know from human-computer interaction research. *Frontiers in Psychology*, *10*, 8. https://doi.org/10.3389/fpsyg.2019.00008

Martin, N. A., & Brownell, R. (2011). *Expressive One-Word Picture Vocabulary Test-4 (EOWPVT-4)*. Academic Therapy Publications.

McNeill, J., & Flower, S. (1999). Let's talk: Encouraging mother-child conversations during story reading. *Journal of Early Intervention*, *22*, 51–69. https://doi.org/10.1177/105381519902200106

Michaelis, J. E., & Mutlu, B. (2018). Reading socially: Transforming the in-home reading experience with a learning-companion robot. *Science Robotics*, *3*, 1–11. https://doi.org/10.1126/scirobotics.aat5999

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*, 7–26. https://doi.org/10.1080/10409280701838603

Neuman, S. B., Samudra, P., Wong, K. M., & Kaefer, T. (2019). 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*, 1100–1110. https://doi.org/10.1037/edu0000428

Noble, C., Sala, G., Peter, M., Lingwood, J., Rowland, C., Gobet, F., & Pine, J. (2019). The impact of shared book reading on children's language skills: A meta-analysis. *Educational Research Review*, *28*, 100290. https://doi.org/10.1016/j.edurev.2019.100290

Peebles, A., Bonus, J. A., & Mares, M. L. (2018). Questions+ answers+ agency: Interactive touchscreens and Children's learning from a socio-emotional TV story. *Computers in Human Behavior*, *85*, 339–348. https://doi.org/10.1016/j.chb.2018.03.039

Petersen, D. B., Gillam, S. L., & Gillam, R. B. (2008). Emerging procedures in narrative assessment: The index of narrative complexity. *Topics in Language Disorders*, *28*, 115–130. https://doi.org/10.1097/01.TLD.0000318933.46925.86

Phillips, B. M., & Lonigan, C. J. (2009). Variations in the home literacy environment of preschool children: A cluster analytic approach. *Scientific Studies of Reading*, *13*, 146–174. https://doi.org/10.1080/10888430902769533

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*, 2332858419878389. https://doi.org/10.1177/2332858419878389

Roseberry, S., Hirsh-Pasek, K., & Golinkoff, R. M. (2014). Skype me! Socially contingent interactions help toddlers learn language. *Child Development*, *85*, 956–970. https://doi.org/10.1111/cdev.12166

Sciuto, A., Saini, A., Forlizzi, J., & Hong, J. I. (2018, June). "Hey Alexa, what's up?" A mixed-methods studies of in-home conversational agent usage. In *Proceedings of the 2018 designing interactive systems conference* (pp. 857–868). https://doi.org/10.1145/3196709.3196772

Sénéchal, M., Thomas, E., & Monker, J. (1995). Individual differences in 4-year-old children's acquisition of vocabulary during storybook reading. *Journal of Educational Psychology*, *87*, 218–229. https://doi.org/10.1037/0022-0663.87.2.218

Sengupta, S., & Garg, R. (2019, March). *Impact of voice-based interaction on learning practices and behavior of children*. In IUI Workshops.

Shanahan, T., MacArthur, C. A., Graham, S., & Fitzgerald, J. (2006). Relations among oral language, reading, and writing development. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 171–183). Guilford Press.

Şimşek, Z. C., & Işıkoğlu Erdoğan, N. (2015). Effects of the dialogic and traditional reading techniques on children's language development. *Procedia—social and Behavioral Sciences*, *197*, 754–758. https://doi.org/10.1016/j.sbspro.2015.07.172

Tewari, A., & Canny, J. (2014, April). What did spot hide? A question-answering game for preschool children. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1807–1816). https://doi.org/10.1145/2556288.2557205

Towson, J. A., Fettig, A., Fleury, V. P., & Abarca, D. L. (2017). Dialogic reading in early childhood settings: A summary of the evidence base. *Topics in Early Childhood Special Education*, *37*, 132–146. https://doi.org/10.1177/0271121417724875

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. (2012). *Thought and language*. MIT Press.

Westlund, J. M. K., Dickens, L., Jeong, S., Harris, P. L., DeSteno, D., & Breazeal, C. L. (2017). Children use non-verbal cues to learn new words from robots as well as people. *International Journal of Child-Computer Interaction*, *13*, 1–9. https://doi.org/10.1016/j.ijcci.2017.04.001

Whitehurst, G. J. (1992). *Dialogic reading: An effective way to read to preschoolers*. Reading Rockets. http://www.readingrockets.org/article/dialogic-reading-effective-way-read-preschoolers

Willoughby, D., Evans, M. A., & Nowak, S. (2015). Do ABC eBooks boost engagement and learning in preschoolers? An experimental study comparing eBooks with paper ABC and storybook controls. *Computers & Education*, *82*, 107–117. https://doi.org/10.1016/j.compedu.2014.11.008

Xu, Y., & Warschauer, M. (2020a). What are you talking to?: Understanding children's perceptions of conversational agents. In *Proceedings of the 2020 CHI conference on human factors in computing systems*, April 25–30, 2020, ACM, Honolulu, HI. https://doi.org/10.1145/3313831.3376416

Xu, Y., & Warschauer, M. (2020b). Wonder with Elinor: Designing a socially contingent video viewing experience. In *Proceedings of the 19th ACM international conference on interaction design and children (IDC '20)*, Association for Computing Machinery, New York, NY. https://doi.org/10.1145/3397617.3398024

Xu, Y., & Warschauer, M. (2020c). Exploring young children's engagement in joint reading with a conversational agent. In *Proceedings of the interaction design and children conference* (pp. 216–228). https://doi.org/10.1145/3392063.3394417

Xu, Y., & Warschauer, M. (2021). Conversational agents for dialogic reading. *Dryad*, Dataset. https://doi.org/10.7280/D10Q3P

Xu, Y., Yau, J. C., & Reich, S. M. (2020). Press, swipe and read: Do interactive features facilitate engagement and learning with e-Books? *Journal of Computer Assisted Learning*, 1–14. https://doi.org/10.1111/jcal.12480

Yuan, Y., Thompson, S., Watson, K., Chase, A., Senthilkumar, A., Brush, A. B., & Yarosh, S. (2019). Speech interface reformulations and voice assistant personification preferences of children and parents. *International Journal of Child-Computer Interaction*, *21*, 77–88. https://doi.org/10.1016/j.ijcci.2019.04.005

Zaga, C., Lohse, M., Truong, K. P., & Evers, V. (2015, October). *The effect of a robot's social character on children's task engagement: Peer versus tutor*. In International conference on social robotics (pp. 704–713), Springer, Cham. https://doi.org/10.1007/978-3-319-25554-5_70

Zevenbergen, A. A., & Whitehurst, G. J. (2003). Dialogic reading: A shared picture book reading intervention for preschoolers. In A. van Kleeck, S. A. Stahl, & E. B. Bauer (Eds.), *Center for Improvement of Early Reading Achievement, CIERA. On reading books to children: Parents and teachers* (pp. 177–200). Lawrence Erlbaum Associates Publishers.

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*, 1523–1545. https://doi.org/10.1007/s11423-017-9533-2

How to cite this article: 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, e149–e167. https://doi.org/10.1111/cdev.13708