

## MIT Open Access Articles

*"What color are the fish's scales?" Exploring parents' and children's natural interaction with a child-friendly virtual agent during reading*

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

**Citation:** Lin, Grace, Schoenfeld, Ilana, Thompson, Meredith, Xia, Yiting, Uz-Bilgin, Cigdem et al. 2022. ""What color are the fish's scales?" Exploring parents' and children's natural interaction with a child-friendly virtual agent during reading."

**Published Version:** <https://doi.org/10.1145/3501712.3529734>

**Publisher:** ACM|Interaction Design and Children

**Permanent Link:** <https://hdl.handle.net/1721.1/146454>

**Version:** Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

**Terms of use:** <https://creativecommons.org/licenses/by/4.0/>



# "What color are the fish's scales?" Exploring parents' and children's natural interactions with a child-friendly virtual agent during storybook reading

Grace C. Lin  
Ilana Schoenfeld  
Meredith Thompson  
Yiting Xia  
Cigdem Uz-Bilgin  
gcl@mit.edu  
ilanasch@mit.edu  
meredith@mit.edu  
yitingxia23@gmail.com  
cigdemuz@mit.edu

Massachusetts Institute of Technology  
Cambridge, Massachusetts, USA

Kathryn Leech  
The University of North Carolina at Chapel Hill  
Chapel Hill, North Carolina, USA  
leechk@unc.edu

## ABSTRACT

With increasing integration of AI-powered agents into educational technologies available to families with young children, the landscape of how caregivers and children interact together or separately with these technologies is underexplored. Understanding the nature of these interactions could critically inform the design of educational technologies to facilitate children's learning, research methods for use in future studies involving adult-child dyad technology use, and policy decisions regarding the use of educational technology with young children. In this study, we explored the natural interactions among parent, child, and a child-friendly virtual rabbit character named Floppy. Floppy is a virtual agent in a Smart Speaker app that models adult dialogic reading and conversational strategies for use with young children (ages 4-6 years). Over a span of four to six weeks, 18 parent-child dyads read *The Rainbow Fish*, a classic children's book by Marcus Pfister, during 24 at-home, remote sessions with Floppy. Of the 189 conversations generated during this time, 125 were initiated by a prompt spoken by Floppy. Though there were some variations among the dyads, across all conversations, parent-driven interactions made up 63% of the conversations, followed by child-driven conversations at 15.3%, Floppy-driven at 14.3%, and Floppy-and-parent-driven at 7.4%. A select few parents were more comfortable having their children interact directly with Floppy, whereas the majority of the parents would direct children's attention back to themselves or help children understand the questions by repeating or reformulating Floppy's prompts. More than half of the parents reported that their children formed emotional connections with the virtual character. These findings point to a

need to clearly define the role of virtual agents, even ones with limited AI, in this type of triadic interaction.

## CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**; • **Social and professional topics** → **Children**.

## KEYWORDS

triadic interaction; early literacy; virtual agent

### ACM Reference Format:

Grace C. Lin, Ilana Schoenfeld, Meredith Thompson, Yiting Xia, Cigdem Uz-Bilgin, and Kathryn Leech. 2022. "What color are the fish's scales?" Exploring parents' and children's natural interactions with a child-friendly virtual agent during storybook reading. In *Interaction Design and Children (IDC '22)*, June 27–30, 2022, Braga, Portugal. ACM, New York, NY, USA, 11 pages. <https://doi.org/10.1145/3501712.3529734>

## 1 INTRODUCTION

The ubiquitous nature of smartphones has affected family time and parent-child interactions. Many researchers warn against the distractions and negative consequences of parental smartphone use. For example, an observational study of caregiver-child dyads in natural settings (e.g., parks and shopping malls) found that caregivers engaged in less joint attention with their children when they were absorbed with mobile devices [27]. The study echoes the findings from an experimental design study, where parents randomly assigned to use their phones as much as possible at a museum felt less connected to their children [22]. Indeed, a recent review of studies on parental phone use suggests that it may negatively affect parenting quality by resulting in fewer interactions and decreased sensitivity and accuracy to child needs and cues [24].

However, when used by parents and children together, smartphone devices have the potential to promote learning in children and improve parent-child interaction. By capitalizing on the qualities of smart devices that excite children (e.g. bright colors, lights, sounds, interactivity), researchers at Harvard have developed and



This work is licensed under a Creative Commons Attribution International 4.0 License.

*IDC '22*, June 27–30, 2022, Braga, Portugal  
© 2022 Copyright held by the owner/author(s).  
ACM ISBN 978-1-4503-9197-9/22/06.  
<https://doi.org/10.1145/3501712.3529734>

released a series of three apps intended to improve the joint media experiences between children and caregivers [34]. One app, for example, allows dyads to record their own voices for characters in different story scenarios. Greater use of these apps by caregiver-child dyads between study sessions resulted in greater improvements in children's mean length of utterance, an early literacy measure of language complexity.

Beyond the consequences to children of smart device use, academics and general audiences alike are also interested in how children perceive, treat, and use smart technology and devices. The interactions that have emerged from the use of these devices have been the prime focus of many recent reports (e.g., [1, 5, 43]). Children as young as three, for instance, are able to give Alexa, a personal assistant, commands such as "Alexa, set the timer for one minute" to hold their parents to a promise to play with them in this amount of time [1]. In the area of early literacy, attention has been focused on conversational agents (CA), where the purpose of the child's interaction with the non-living agents is to improve their language outcome. Thus far, early literacy studies show that the CA approach has been somewhat effective [16, 45].

The nature of the human dyads' interactions with smart technology also extends to robots, or physical agents. A number of studies have investigated how children interact with robots and how social robots, for example, are able to enhance the interaction of children with adults or other peers [20, 36, 39]. Furthermore, interviews with caregivers indicated that they perceived themselves being in a supportive role when their children interacted with a physical agent, yet they still mostly considered the activity to be a caregiver-child one rather than a child-robot one [13]. Such studies open up further questions regarding the nature of the caregiver-child-agent triadic interactions.

Rather than a smart device serving as a potential source of distraction or as a device for coordinated joint activities, the non-living agent—whether a virtual CA or a physical social robot—can help focus attention by functioning as an active member in the dyad's interaction. However, while the studies with robots may have focused on the role parents play when their children are interacting with the agent or how the child interacts with the adult in the presence of the agent, less is explored about the role the agent should be playing. Most of the time, the agent has a set and defined role, and how the humans act toward it was assumed to align with its design or simply not reported in studies. More research is therefore needed to investigate whether human dyads naturally hold different attitudes toward and interaction styles with non-living agents and if the styles may affect the intended objective the agent is designed to perform. Understanding such tendencies and their effect will inform designers and researchers to better tailor the functions and roles of the agent to fit the needs of the dyads.

We are thus particularly interested in exploring the nature of the triadic interaction in which a member of the triad is a non-living intelligent agent. In this paper we seek to address the following research questions:

- (1) How do caregiver-child dyads interact with an embedded virtual agent while reading?
- (2) Is the quality of conversations associated with different interaction styles or patterns?

- (3) Relatedly, how do parents perceive and describe the interaction and facilitation of the virtual agent?

To investigate these questions, we analyzed in-app audio recordings generated by 18 parent-child dyads as they read a children's book aloud with Floppy, a virtual cartoon rabbit with limited artificial intelligence capabilities who interjects with spoken questions along the way. We also analyzed parents' interviews to examine their perceptions of their interactions with the virtual agent.

## 2 RELATED WORK

Our study marks a first step in investigating the nature of triadic interaction where a member of the triad is a somewhat intelligent virtual agent. We draw from previous work on joint media engagement between caregivers and children as well as human-computer interaction studies involving non-player characters (NPCs), conversational agents (CAs), or robots.

### 2.1 Joint Media Engagement

Joint media engagement (JME) refers to people's shared media experiences, especially for meaning-making and learning [40, 41]. JME can occur between parents and children, peers, siblings, or even teachers and students [41]. Earlier studies on JME examined the nature and effects of co-viewing (e.g., watching television programs such as *Dora the Explorer* together) or game-play experiences [41]. For instance, an ethnographic study where 10 families were observed for over 60 hours showed that JME, particularly co-viewing, was most often initiated by the children, that children answered more prompts when co-viewing a show with others, and that they imitated each other's responses [25]. The findings helped unearth design principles for JME, such as mutual engagement and dialogic inquiry. In the specific case of reading, dialogic inquiry may involve the more experienced reader (i.e., the caregiver) prompting the new reader (i.e., the child) with questions about the story or content [41]. In addition to extracting design principles, researchers also noted potential challenges in JME. For instance, while playing an intergenerational video game, parents needed role clarification and took time to figure out strategies to carry out their roles [3]. Moreover, the desirable forms of interaction (e.g., dialogic inquiry) may not always happen naturally [41] as parents may not be well versed in how to guide their children or may be too busy to do so.

Regardless of the challenges, the research around the potential benefits of JME is compelling enough that the American Academy of Pediatrics has recommended JME to parents with 2- to 5-year-old children [15]. However, shifts in the technological landscape also signal potential shifts and updates in research and policy recommendations. Unlike the research a decade ago, where JME tended to be limited to co-viewing of television programs, a recent review of 27 papers was focused on JME around newer digital media such as smartphones or tablets [8]. In this new landscape, the authors first investigated whether JME occurred around smart devices. For the most part, the answer was affirmative, though some studies reported that some parents still preferred to engage in JME with nonmobile media [38, 42]. Unfortunately, when JME did occur, the parent-child interactions were more limited, and many key aspects of language quality were reduced in joint e-book reading compared with joint print reading [8]. The authors cautioned against jumping

to conclusions, as many results were mixed, and called for more research on parent-child interaction during JME.

## 2.2 Interactions with Media Characters and Robots

A prerequisite of examining joint media engagement is understanding how individuals may interact with media or media characters. In games studies, a whole class of characters has been identified based, in part, on its interaction types with players. Characters that assist and guide players as part of the game without engaging in the game action itself are called “non-player characters,” or NPCs [44]. NPCs can fulfill different functions within games. Warpefelt [44] identifies three common functions: vendors who buy and sell things, service providers who provide different services, and quest-givers who provide information to the player(s) about how to complete different actions in the game. He further classifies NPCs as adversarial, such as an enemy or opponent, or friendly, including sidekicks, allies, companions, pets, or minions [44]. Johansson et al. [19] discusses three types of roles that NPCs can play during games. While single agents can either be acting agents, who are goal-directed and use language and tools, or reacting agents, who can interrupt gameplay, adapt, and acquire information from players, multiple agents have an additional capability of interacting with game players. NPCs can also make educational games more appealing to the users [48], and can be used to gather information from players through surveys [10]. The appearance of NPCs is important and can impact how players interact with them during the game [14]. Even though NPCs are not active players, they can provide essential information to players, and can become pivotal to players’ experiences. Players may even develop strong attachments to them during the course of a game [4].

There have been similar findings related to children’s media characters. In co-viewing studies, for example, some children’s responses were prompted by TV characters such as Dora from *Dora the Explorer* [25]. Research has also shown that children who have stronger parasocial relationships with media characters may learn more effectively from those characters [18]. Some virtual agents have been created as peer-like learning companions for children [35]. It is also very common for virtual characters to be situated as teachers or tutors in multimedia learning environments [37].

**2.2.1 Conversational Agents.** Conversational agents (CAs) are yet another form of media character. Whether virtual or robotic, CAs are designed to carry on conversations with humans. Consumer-oriented CAs like Amazon’s Alexa or Apple’s Siri are increasingly becoming part of households [33]. Children use Alexa in similar ways as adults: primarily for retrieving information such as the time, the weather, and searching for a topic on the internet [23]. In a study of 26 participants between the ages of three and ten, researchers found that children interacted with CAs in four main ways: gauging the level of intelligence of the CA, exploring what the CA could do as if it were a person, engaging the CA in a playful way, and attempting to establish understanding with the CA [6]. They also found that younger children (aged 3-4) had difficulty engaging with the audio-only CAs, but enjoyed interacting with a physical robot. A study of 80 children between six and ten found that a majority of children said a voice assistant could be their friend [12],

suggesting that children can make emotional connections with CAs. The younger children were also more likely than older children to view CAs as “social counterparts” that could be sources of reliable information.

Another class of CAs are designed to help children learn. Educational CAs are often developed with the goal of improving children’s learning in literacy, math, history, or science [16, 17, 28, 45–47], in which case the conversation flow may be more guided (question-answer-feedback). Other researchers develop CAs with the goal of applying them to industry—designing agents that may one day be capable of conducting, for example, interviews with personalized questions based on their conversational partners’ personality [49]. Typically, CAs are powered by AI, using techniques ranging from decision trees (e.g., [46]) to topic analysis by a human coder followed by deep learning [49]. Though the end-goal may be designing an AI-powered agent, studies in human-computer interaction sometimes involve a “fake” CA controlled by a human researcher behind the scenes (e.g., [16, 17]). This is known as the Wizard-of-Oz protocol. It is widely used by researchers to understand how potential users might react to a future technology should it be built. For instance, in examining how children learn from their interactions with CAs, researchers used a program that would convert the human wizard’s voiced responses to a robotic voice, creating the illusion that the child was responding to a CA on a phone [16]. The design is powerful in terms of illustrating children’s ability to pick up linguistic routines or words when the CAs are carrying on human-like conversations. It is less clear whether the connection between children and a virtual agent can be as readily established when the agent might possess fewer human-like qualities, e.g., initiating questions without the ability to respond.

**2.2.2 Social Robots.** Studies of triadic (2 humans and 1 agent) interactions have mostly surfaced in work related to social robots with varying functionalities. Using a robot meant to teach English as a second language to toddlers, researchers were able to observe triadic interactions and verify such interactions during semi-structured interviews with parents. Even though the robot gave the instructions to the toddler, half of the parents in the study saw the game as a parent-toddler activity. Indeed, they saw themselves playing a support role, mediating the novel experience for their child [13]. In another study with the humanoid robot NAO, researchers found that 4-year-old children were able to interact with NAO without help. However, parents sometimes offered unsolicited support, repeating the robot’s questions and scaffolding their children [31]. These recent studies provide evidence that the focal point of the investigation in a caregiver-child-robot triadic interaction has been on the child-robot dyad; they meant to explore how children can interact with the robot and see the role the caregivers would play to enhance that interaction.

Flipping the roles a little, research on using social robots to help improve the social skills of children places the emphasis on how the social robot can potentially affect children’s interaction with other people. For example, in a study with the social robot Jibo, 12 participants with autism received the Jibo intervention for 30 days [36]. During the intervention, Jibo modeled social skills such as making eye contact and sharing attention. Compared with their pretest

behavior, children made more eye contact, initiated more conversations, and responded more when others called them after the intervention [36]. Similarly, research using a dinosaur-shaped social robot named Pleo also showed that children with autism spoke more to the human interaction partner in the condition with Pleo than in other conditions (with another human or with a tablet) [20]. In a study with typically developing kindergartners, researchers demonstrated the effectiveness of a mediator robot named Keepon in resolving conflicts that naturally occur when children play with each other [39]. Keepon, operated by a human using the Wizard of Oz protocol, would sound a whistle to capture children’s attention when a conflict had been spotted. The ensuing prompt helped children to alter their interaction, leading to conflict resolution [39]. Taken together, these studies suggest that a physical agent can affect behavioral changes in human interactions.

However, less is known about how the interaction of human dyads would play out with a virtual agent with a limited physical presence, especially in a naturalistic home environment, and whether different interaction styles (if any) would affect the quality of the interactions.

### 3 METHOD

#### 3.1 Participants

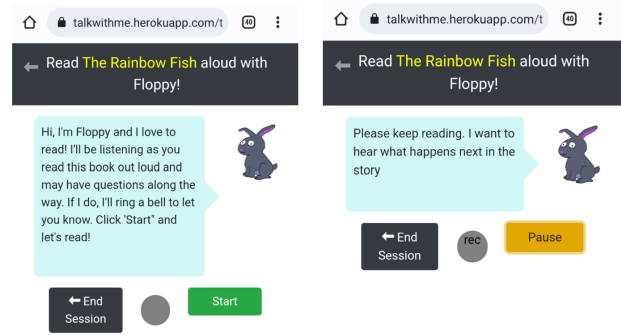
Twenty parent-child dyads (18 mothers, 2 fathers) from across the United States participated in a pilot study on the usability of a Smart Speaker app with a virtual agent named Floppy. (For details on the recruitment process, please see selection and participation of children.) For the purpose of this paper, we focus on the interactions that took place at home while the parent-child dyads were reading *The Rainbow Fish*, a children’s picture book by Marcus Pfister, using the app. One parent did not read *The Rainbow Fish* at home with their child, and another sped through only the first page of the book. Therefore, 18 dyads remain in our analysis sample. The 18 participating children were between the ages of four and six, and evenly split between boys and girls. Tables S1-S5 in the supplementary materials display the demographic characteristics of our sample.

#### 3.2 Floppy Design and Functionality

Floppy’s function in the triadic interaction (between parent, child, and virtual agent) is two-fold: to create a rapport with the child; and to support the parent by modeling dialogic reading practices as hard copy books are read aloud by the parent to the child. We designed Floppy (see Figure 1), the Smart Speaker app’s child-friendly, virtual, cartoon rabbit character with input from a team of community-recruited parent-codesigners early in the design process, and in consultation with experts in the field (a children’s educational game designer, librarians, and pre-school care providers).

The developers pre-programmed the app with questions that match the contents of particular pages of the books. Floppy uses AI technology to “listen in” and recognize the words spoken by the human dyads. When it “hears” a marker word indicating a page with a pre-programmed question, a chime sounds and Floppy interjects with the question.

In more detail, at the beginning of a reading session, parents activate a green “Start” button, as shown in the image on the left in Figure 1. At this point, Floppy, who is present on the phone screen,



**Figure 1: Screenshot of the Smart Speaker App. Floppy is the rabbit character displayed on the right side of the screen. Before starting to read, there is the “Start” button (left image). Once the dyad presses “Start,” the screen changes to the display on the right. The circular “rec” symbol will blink to indicate the app is recording. Dyads may also press “Pause” at any time (see image on the right).**

records and transcribes the dyads’ spoken words. A circle icon on the screen blinks red to indicate to the parents that Floppy is properly “listening” and recording. When specific marker-words are read by the parent and recognized by the app, a chime sounds and Floppy interjects with a question that is tailored to the specific content of the book at that location in the narrative. Although Floppy does respond to the caregiver’s reading with pre-programmed dialogic reading questions, Floppy is not designed to give dynamic feedback to the dyad that a more advanced conversational agent might provide. Instead, Floppy’s primary goal is to support the parent in the use of dialogic reading practices by providing (and modeling) questions to encourage back and forth conversation about the book between the parent and child.

At any point during the in-app reading, the parent can pause the recording by pressing the “Pause” button and resume recording again by pressing the green “Start” button one more time. The decision to add the pause functionality to the app was based on feedback from participants during early informal usability trials. Parents requested a way to pause and control the app’s recording function as they anticipated there would likely be interruptions during the in-app reading of books at home (e.g. other siblings interrupting, loud background noises, etc.) and times that they might prefer more privacy. While recording, the Smart Speaker app picks up any sound in the vicinity of the phone. The app stops recording when the dyad taps the “Pause” or “End Session” button, as shown on the right panel in Figure 1.

#### 3.3 Procedure

Before the start of their first study session, each participating dyad was mailed the children’s storybooks they would need to complete the study (the books were gifted to the dyads after the study), as well as a smartphone with the Smart Speaker app (the phones were returned to the research lab after the study). Each dyad joined an online session via Zoom where they read one storybook together without using the Smart Speaker app before the researchers

introduced them to the app and dialogic reading practices. After the parents watched videos explaining and modeling several dialogic reading strategies, the parent-child dyads read another book while using the Smart Speaker app. It is worth noting here that the videos the parents viewed did not provide any instructions or recommendations as to how to interact with the app's virtual Floppy character. At the end of the session, the researchers interviewed the parents about their experience using the app and asked them to read three books (*The Rainbow Fish*, *Corduroy*, and *Ladybug Girl and Bumblebee Boy*) at home on their own using the app. After three to six weeks, the dyads came back for a final Zoom session that followed an almost-identical format to the initial session. The only difference was that they were no longer required to watch the introductory videos during this session.

### 3.4 Coding Framework and Data Analysis

The data we analyzed are drawn from the at-home readings of *The Rainbow Fish* as well as interviews from participants' second Zoom sessions. Interviews were recorded both through Zoom and the researchers' own notes and memos. All of the interview notes and memos were reviewed for common themes. A single researcher then condensed the interview notes and direct quotes from caregivers into an Excel table to document themes that extended across the interviews. The extracted themes, together with the patterns that emerged from the audio files, uncover parents' approaches and impressions of Floppy during the reading sessions. The research team transcribed the recordings from the app. Each transcription was checked for accuracy by at least one other researcher. Any discrepancy was discussed. Because of the nature of children's speech and occasional glitches in the audio files, some utterances were marked as unintelligible and therefore not coded.

A coding scheme was developed by the research team to capture the types of interactions that took place during these reading sessions. Specifically, our codes sought to capture parents' and children's responses to Floppy (adapted from dialogic strategy codes [26, 29]). We also developed a broader set of interaction style codes (see Table 1) inspired by previous research on parent-child interactions in digital co-play situations (see [43]). Drawing from developmental science research on the triadic interaction among two caregivers and a child (see Lausanne Trilogue Play or Family Alliance Assessment Scales; [9, 30]), we were interested in whether the human dyads saw Floppy as another partner in their conversation and included Floppy by mentioning it or speaking directly to it. A primary coder coded all the transcripts, and a secondary coder coded 20% of the transcripts. Interrater reliability was above .85 for all codes. All discrepancies were resolved through discussion.

**3.4.1 Quantitative Metrics of Interest and Analytic Plan.** In addition to obtaining the frequency of the interaction styles, the quantitative measures derived from the coding of the transcripts include the following:

- Utterance: A unit of speech made by the parent or child. This excludes pure reading of the text in the book.
- Conversation: A unit of utterance exchanges initiated by Floppy, the parent, or the child that ends when the parent returns to reading or when the book or recording ends.

- Conversational turns: number of times the speaker changes in a conversation.
- Expansion: An occurrence when the topic expands beyond the original one during the same conversation.

For more description and definition of the utterance codes, see Table S6 in the supplemental materials. In our analysis, we note that Floppy may initiate a conversation with a prompt. However, with its limited intelligence, Floppy does not engage in the conversations; thus, its prompt is not counted as an utterance. As is conventional in dialogic reading research, we use conversational turns and expansion as our indicators of conversational quality [26, 29, 32].

We examined all quantitative measures at multiple levels (e.g., by the reading sessions, by the dyads) to obtain a more complete picture of the interaction and to ensure that the results are not biased by particularly active participants. In the following section, we collapse the reporting to total frequencies or proportions and averages across reading sessions as "by the dyads" and "by reading sessions" provide similar information (i.e., the results are not being pulled or driven by more active participants who read more times). We only report multiple levels when conflicting findings arise from the different levels.

## 4 RESULTS

Across our sample of 18 dyads, 28 audio files totaling 110 minutes were generated in 24 reading sessions of *The Rainbow Fish*; in four reading sessions, the dyad had pressed "Pause" and a new file was generated when they resumed reading. Thirteen dyads read the book with Floppy once, four read it twice, and one read it three times. Overall, the dyads generated 1,022 coded utterances, 189 conversations, and 792 conversational turns in the conversations. Of the 189 conversations, 125 were started with a Floppy prompt.

### 4.1 Interaction Type and Style

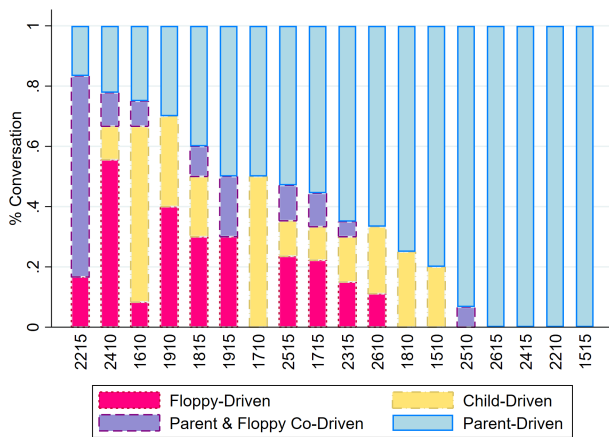
All of the interactions took place in the caregiver's homes. On average, every at-home reading session of *The Rainbow Fish* generated an average of 10.49 (SD = 4.57) conversations, but there was a wide range of only 1 conversation to 20 conversations. The number of conversational turns also varied widely, ranging from 1 to 90, with an average of 48.71 (SD = 22.28). The most common type of utterances from the parents and children were responses to their child (total frequency = 371, Mean per reading session = 23.06, SD = 11.76) and responses to their parents (total = 372, M per reading session = 23.15, SD = 11.44), respectively. Out of 125 conversations initiated by a Floppy prompt, children responded directly to Floppy 38 times.

Floppy was only mentioned by the dyads a total of 11 times. The majority of the dyads (12 out of 18) did not mention Floppy, though one dyad had a conversation about Floppy's functionality, and another caregiver actively encouraged his child to "answer [Floppy]" when the child was silent after Floppy asked a question. There were far fewer instances of the dyads speaking directly to Floppy than instances of the dyads speaking to each other or speaking about Floppy; only two dyads spoke directly to Floppy for a total of six times. Only under two circumstances did the human dyads speak to Floppy: when Floppy's prompt timing was off (e.g., "we're not even there yet") or when they were ending the session (e.g., "goodbye").

Interaction Style	Definition
Parent-Driven	- The parent initiated the conversation OR - Floppy initiated with a question, but the parent repeated the prompt before the child answered
Floppy-Driven	- Floppy initiated with a question, - The child responded directly to Floppy’s prompt, AND - The parent did not follow up with additional questions
Parent- and Floppy-Co-driven	- Floppy initiated with a question, - The child responded directly to Floppy’s prompt, AND - The parent followed up with additional questions
Child-Driven	- The child initiated the conversation OR - Floppy initiated with a question, the child answered Floppy’s prompt directly, and the child asked their parent additional questions

**Table 1: Main Codes and Definitions**

Across all the conversations generated by the participants, parent-driven style was the most common, accounting for 119 occurrences out of 189 conversations (63.0%), followed by child-driven (29 out of 189; 15.3%), Floppy-driven (27 out of 189; 14.3%), and parent-and-Floppy jointly driven (14 out of 189; 7.4%). However, examining the pattern of interactions by the dyads, this interaction style was not dominant across the board (see Figure 2. For descriptive statistics, see Table S7 in the supplemental materials). Although on average, 61.2% of the dyads’ conversations were parent-directed, this proportion ranged from 16.7% (Dyad 2215) to 100% (Dyads 2615, 2415, 2210, 1515).



**Figure 2: Proportion of Interaction Styles by the Dyads.**

When we examined the interaction style of the triads, we saw only about half (10 of the 18) of the dyads had a predominantly parent-driven interaction style. The remainder had a mix of styles (see Figure 2 and Table S8 in supplemental materials). In the instance of parent-driven interaction, the parents would repeat the Floppy prompt in some form and expect their children to respond directly to them. Interestingly, a couple of dyads exhibited a true triadic interaction where the conversations were equally led by the child, the parent, and Floppy (e.g., 1815). To illustrate this mix of interaction styles, we present three examples from 1815 in Table 2.

## 4.2 Quality of Conversation

**4.2.1 By Interaction Style.** Although there is a distinct pattern among the dyads in terms of how they prefer to interact with Floppy, the interaction style was not associated with common measures of conversation quality such as the number of conversational turns or expansion of conversations. Specifically, the number of turns taken by the dyads was uncorrelated with the percentage of their conversations that were driven by parents ( $\tau_b = .04, p = .85$ ), by Floppy ( $\tau_b = .10, p = .60$ ), by the child ( $\tau_b = .16, p = .39$ ), and jointly by parents and Floppy ( $\tau_b = .02, p = .93$ ). Similarly, the number of expansions was also uncorrelated with the percentage of dyads’ conversations that were driven by parents ( $\tau_b = .08, p = .71$ ), by Floppy ( $\tau_b = -.06, p = .80$ ), by the child ( $\tau_b = -.02, p = .93$ ), and jointly by parents and Floppy ( $\tau_b = .00, p = .99$ ).

**4.2.2 By Conversation Starter: Floppy-Initiated vs. Parent/Child-Initiated.** In general, Floppy started 125 of the total conversations. In contrast, parents started 40 conversations and children initiated the remaining 24 conversations without Floppy prompts. Of the 24 reading sessions that were included in this data analysis, a prompt from Floppy was the first conversation starter for 16 (66.7%) of the 24 sessions. Parents initiated the first conversation without a prompt from Floppy seven (29.2%) times, and children only started the first conversation once (4.2%).

The quality of conversation that ensued differed according to who the conversation starter was. Averaged across the dyads, the conversations that started with a Floppy prompt had 36.3 turns (SD = 30.1, median = 31), whereas those initiated by parents or children had only 7.72 turns (SD = 9.00, median = 4.5). The conversations initiated with a Floppy prompt on average had 1.61 expansions (SD = 3.11, median = 0.5), whereas those initiated by parents or children had 0.22 expansion on average (SD = 0.73, median = 0). See Figure 3 for boxplots of conversational turns and expansions with different conversation starters.

As the distributions of conversational turns and expansion were skewed, we conducted Wilcoxon signed-rank tests on the number of turns and expansions in conversations that were initiated by a Floppy prompt versus not. The difference between the medians for conversational turns by different conversation starters was significant,  $z = 3.38, p = .0007$ , as is the difference between the

Floppy-driven	Parent-driven	Child-driven
	Floppy: Can you remember a time when you shared something with your friend? What was it?	Child: Mama?
Floppy: If you have special glittering scales like Rainbow Fish, would you share them with your friends?	Parent: [Unintelligible] new car with your sister?	Parent: Yes, (child’s name)?
Child: Never.	Child: It was my new car.	Child: Um. My teacher had this book too.
Parent: No? All right.	Parent: Yeah. Remember sharing your new car with your sister, right?	Parent: She did? I didn’t know that. All right. You ready?
	Child: Um (Affirmative)	Child: Um (Affirmative)
	Parent: She shared with you? It was fun.	

Table 2: Examples from Dyad 1815

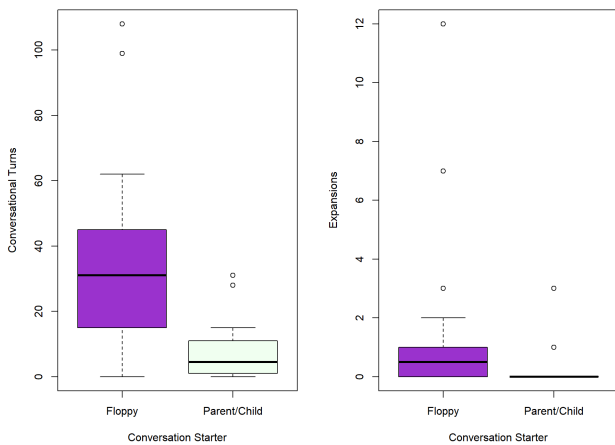


Figure 3: Boxplots of the number of turns (left) and expansions (right) when the conversation is started by Floppy or by the Parent/Child.

medians for the number of expansions by the conversation starter,  $z = 2.798, p = .0051$ .

In addition to the quantitative difference, the conversations were also qualitatively different based on the conversation starters. Conversations that began with a Floppy prompt centered around the topic of the story, and parents and children often built upon it, sometimes expanding the conversation further. In contrast, conversations that were initiated by the parents or children without Floppy tended to be about the contextual elements depicted in the book. Table 3 displays three examples of conversations, one initiated by Floppy, two by the parent, and two examples that were initiated by the child.

### 4.3 Parents’ Perception

Although we did not specifically ask parents for their perceptions of their interactions with Floppy, parents provided unsolicited insights into the connections they and their children made with Floppy in their responses on how Floppy influenced their reading.

*4.3.1 Perception of Floppy’s Role and Its Influence.* In terms of Floppy’s overall role, parents appeared to see Floppy as either a helper providing questions or as a behavioral buffer.

Even though Floppy could not interact dynamically with the dyad, the presence of the virtual character did have an influence on both the caregiver and the child. Caregivers noted that Floppy “makes me think to ask more questions while we’re reading” (1915), and that “sometimes it asks questions I didn’t think of” (2210). A couple of caregivers mentioned that using the app while reading started a question-asking habit that extended even when Floppy was not there. One caregiver noted that “When we aren’t using Floppy, I kind of miss it. I pause it to be my own Floppy. Like, remember when we did this? Instead of going through it, we’re pausing to tie it back in” (1815).

Data from the interviews also revealed that in one particular case, the dyad was able to engage in conversations around Floppy-prompted topics days after they read the book. Occasionally, the child would even chastise Floppy (“I already answered it”) when Floppy asked the same question the second time they read the book (1815).

As a behavioral buffer, Floppy provided relief to the parents because they could ask their children to listen in and pay attention to Floppy. For example, a parent reported that she could tell her child, “Wait wait wait, listen to it. Ding ding ding. Let’s stop.” when the child’s attention may have been waning (2215). Another parent reported that her child was more engaged with the story when reading with Floppy because the child “isn’t trying to tell me a story about whatever, but more like paying attention because she knows she will be listening to and answering questions” (2510).

*4.3.2 Perception of Emotional Connection.* Twelve of the 18 parents revealed that their children had grown attached to Floppy. In fact,

Floppy-initiated (Parent- and Floppy- driven)	Parent-initiated (Parent-driven)	Child-initiated (Child-driven)
Floppy: Can you remember a time when you shared something with your friend? What was it?	Parent: Look at the scales. See how shiny they are?	
Child: um. It was a piece of chocolate.	Child: um hum. (affirmative)	Child: Hey, what are you doing? What's he doing?
Parent: Oh. So you shared a piece of chocolate with your friend. How does that make you feel?	Parent: All the different colors. Do you want to feel... the book?	Parent: See, he is giving out his scale. Maybe.
Child: Great.	Child: um hum. (affirmative)	
Parent: What about... How does that make your friend feel?	Parent: Ohh~ Child: Ohh~ Parent: He's very pretty, right?	Child: Is that an octopus?
Child: Great.	Child: All the tiles there?	Parent: um hum (Affirmative)
Parent: Yeah?	Parent: Right?	

**Table 3: Examples of Conversations Started by Different Members of the Triad**

at the start of one interview with a parent, the child had wanted to continue reading with Floppy so much that the child’s mother let her take the phone with the app to her father. As the child was leaving, she announced that she would “come back and let you know that Floppy didn’t do anything [bad]” (2510):

**Child:** No. I wanna read it here xxxx and with Floppy.

**Mom:** here, take it, and show Daddy how to work Floppy. (to researchers) Is that ok if she takes this? (Both researchers nodded.) Here, take this and show Daddy how to work it.

**Child:** Ok. my baby. I don’t need to keep the [computer].

**Mom:** I know you don’t need the computer. Go, so I can talk to the girls [the researchers] okay?

**Child:** Ok and you’ll see.. And I’ll tell you how Floppy didn’t do anything.

**Mom:** Ok, you’ll tell me how Floppy didn’t do anything?

**Child:** Yep!

The mom later explained that the child meant that Floppy did not do anything bad, suggesting that the child saw Floppy as a peer whose behavior was something she would need to report to her mother.

This observation of a developed connection with Floppy was not an isolated instance. A number of parents also revealed how their children appeared to have formed a bond with Floppy. A father explicitly used the phrase “attachment to Floppy” when describing his daughter’s connection to the character (2315). Multiple parents reported that their children were asking for Floppy (e.g., “two days later, he asked if we could use Floppy again” (2410), “he wanted to do all the books with Floppy” (1815), “she wanted to read any book we have with Floppy” (2510)). Furthermore, the presence of Floppy also affected the family dynamics such that the siblings who were not part of the study also wanted to be involved (e.g., “My oldest

child would . . . grab the phone and read it [the book] to Floppy” (1710)).

Even though the researchers did not specify how the triadic interaction should take place, many caregivers shared the view that Floppy should establish a rapport with their children. They recommended visual or audio changes to Floppy to make it more colorful (five parents) and appealing to the kids. One parent even explained that the changes would be essential as she would actively bring Floppy into the reading activity: “I think Mr. Floppy can be more colorful because I introduce him like ‘Mr. Floppy has a question’” (2210).

## 5 DISCUSSION

Our first research question was focused on how human dyads naturally interact with an embedded virtual agent while reading. Relatedly, we also explored how parents described the interactions and connections with the virtual agent (RQ3). Without any explicit recommendations or guidance on how to use or treat the virtual agent Floppy, the 18 dyads responded to the virtual agent in four distinct interaction patterns while reading *The Rainbow Fish*: parent-driven, Floppy-driven, parent- and Floppy-driven, and child-driven. Although parents drove the majority of the conversations, the extent that parent-driven interaction happened varied greatly among the dyads.

The differences in style may reflect the parents’ varying levels of comfort with the presence of a virtual agent and is important as a potential indicator of how parents perceive the role of a virtual character in a traditionally dyadic exercise of reading. Some parents treated Floppy as an assistant meant to scaffold them, and would take Floppy’s questions, rephrase them, and expect the child to answer directly back to them. The “stars” of the reading session

were still the human dyads: caregiver and child. This take was further supported with parents' interview data, where many of them enjoyed how Floppy provided them with the questions. Others saw Floppy as a co-facilitator or even a buffer or a relief, and were comfortable letting their child interact directly and even bond with Floppy. Some parents even made suggestions for visual and audio changes to Floppy to make it more kid-friendly and appealing. The interaction styles that surfaced can provide insights into future designs of similar reading technologies, differentiating and tailoring the characters to suit the caregivers' natural preferences.

The interaction style differences may also indicate a natural ambiguity of the roles as the dyads experienced this new, emerging type of reading interaction with a virtual agent. As mentioned earlier, a father-and-son dyad (2410) had an exchange where the father had prompted the child to "answer it [Floppy]" when the child was silent after Floppy asked a question. A closer inspection revealed that immediately prior to that exchange, the father had partially repeated a Floppy question. In ambiguous situations, children often use social referencing and look to the adult for cues on how to react (see [7] for a review). Thus, perhaps the change in style made it confusing for the child to adapt in the novel reading situation with a virtual agent. As the child was able to promptly answer the question Floppy asked earlier after his father's nudge (note that the father did not repeat the question), the original silence was likely due to the child waiting to respond to his father rather than to inattention. This kind of confusion/attention-split was experienced by the parents as well. One mom mentioned during her last session interview that "it [Floppy] required more attention from me" in order to hear the question so that she had to split her focus between Floppy and her child.

On a more granular level, it should be noted that children answered Floppy directly in 38 of the 125 conversations started by a Floppy prompt. Given previous research showing that preschool-aged children would treat robots or CAs as informants or social partners and seek answers from them [2, 12], it might not have been surprising that some of our participants had developed a connection to Floppy, answering its questions directly. Indeed, the emotional connection the children formed with Floppy was confirmed by the parents during the interviews. This finding is a contribution to the field, as previous research mostly focused on the connection younger children could create with a physical agent such as a robot [6]. As a virtual agent, Floppy cannot establish connections via gestures or other social contingency cues the way physical robots can. Moreover, as an agent with limited artificial intelligence (unlike a true conversational agent), Floppy could only "listen in" and provide unidirectional prompts but could not respond back. It is therefore interesting that children were able to make connections with a virtual character that was primarily meant to provide prompting and modeling to the adult. These connections suggest that the character Floppy was engaging enough to prompt an emotional connection, similar to an emotional connection that might be made with a CA [12] or an NPC [4]. In the interviews, parents shared evidence that their child went beyond engaging with Floppy during one reading session, requesting Floppy in future reading sessions, another indication that the children were able to make a connection with a CA with limited responding power. Future research should

further examine the potential differences and the ways in which children make connections with virtual and physical agents.

In contrast to Floppy-sparked conversations, we observed far fewer instances of the human dyads speaking to Floppy directly (though it did happen). One of the two circumstances when Floppy was actively included in the conversation occurred at the end of reading sessions when a member—typically the child—said "goodbye" to Floppy. As a reminder, pressing the button to "End Session" as shown in Figure 1 would prompt Floppy to say "goodbye" to the dyad. The cases of the children saying "goodbye" were perfect examples of their emulating Floppy's speech (even the cadence of the echoed "goodbye" was sometimes similar to Floppy's), a behavior that has been observed in previous research with more animated and, arguably, more responsive, wizard-controlled conversational agents or robots [16, 21]. These rare circumstances again suggest the potential of virtual, cartoon agents with limited AI to establish rapport and inspire emotional connections with children. Additionally, because children who establish rapport, trust, and emulate their conversational agents appear to have greater learning gains [11, 21], future studies should also investigate whether the children who are more prone to answer Floppy (or any conversational agent), include Floppy, or speak to Floppy have higher language outcomes resulting from engagement with the Floppy app.

We found mixed evidence for our secondary research question of whether the quality of conversations is associated with the virtual agent's prompts. Our study's finding that Floppy was the first conversation starter for 67% of the sessions (16 out of 24 reading sessions) suggests that Floppy is influential in catalyzing conversations that may not have happened without Floppy's reminder. The pilot study did not include a comparison group, so we are not currently able to comment about what would have happened had Floppy not been present. We were, however, able to compare the conversations that were initiated by Floppy prompts with conversations that were not initiated by Floppy prompts, and saw both quantitative differences in the number of turn-takes and expansions in the conversations as well as qualitative differences in the content and depth of the conversations. However, if we look at the interaction styles, we see no difference in conversational turns based on whether a conversation was driven by the parents, by Floppy, by the child, or jointly by the parent and Floppy. Taken together, this finding suggests that while Floppy is an influential catalyst in the conversations, the ways the human dyads choose to interact with Floppy do not seem to influence conversational quality. Future research should investigate whether the interaction patterns have any influence on child learning outcomes.

## 6 CONCLUSIONS AND FUTURE DIRECTIONS

We saw the emergence of different styles of triadic interactions when a semi-intelligent virtual agent was present during parent-child storybook reading, but the styles were, for the most part, not associated with the number of conversational turns or expansions in the conversations by the parent-child dyads. The different interaction style patterns and the reactions to Floppy by the dyads during the triadic interactions illustrate the complexity of adding a virtual agent to the typically dyadic activity of parent-child reading. Floppy, a virtual cartoon agent with limited artificial intelligence,

was influential in sparking conversations, and there is evidence suggesting that children were able to connect with the virtual agent despite its limited functionality.

Future studies should investigate whether the various interaction styles and patterns may be associated with additional outcomes in both usability and human learning as well as whether different patterns of triadic interactions may emerge with a more intelligent, truly conversational agent. Additionally, more studies should examine potential differences between virtual and physical agents as well as the level of AI necessary to engage children and foster their learning. There should also be further investigation into the designed role of the agent, and whether the changes in the role could elicit different types of interactions and foster learning. In our study, Floppy was designed to model dialogic reading strategies for the caregivers, and the caregivers indeed saw Floppy as a support and/or buffer. More work could be done to highlight the agent's support role or its buffer role. It would also be interesting to see, for instance, how the caregivers and children would interact with an agent whose role was specified to be the child's peer, modeling for the child how to answer certain questions. The ensuing triadic interactions and the relationship of rapport (established between the child and the non-human agent) could shed further insight on how the design of these roles might be leveraged to foster children's learning.

## 7 SELECTION AND PARTICIPATION OF CHILDREN

This study was reviewed and approved by the Institutional Review Board of Massachusetts Institute of Technology. The initial screening and recruitment of parent-child dyads (with children ages 4-6 years of age) was conducted by Touchstone Research\* in consultation with the research team. The study procedure was explained in detail to potential participants by the researcher-coached Touchstone contact and any questions were answered. Interested parent-participants that consented to being screened were asked to complete an online demographic screening survey and a follow-up phone interview to validate the information captured in the online survey. Out of 131 parent-child dyads that completed the online screening, 25 parent-child dyads were recruited, with a goal of 20 active study dyads. The parent-guardians of each parent-child dyad that met the study criteria (e.g., being residents of the U.S.; having children from 3-5 years old; having a household income below \$75,000; users of mobile phones with internet connection at home) were asked by the researchers to sign a consent form to participate in the research study. In both the written consent form and orally during the research study, parents and children were informed that they could discontinue the study at any time if they wished to do so. All of the researcher-led remote study sessions were carried out with the children's parents present.

\*Touchstone Research's security program is based on the Massachusetts standards for the Protection of Personal Information and incorporates COPPA, CCPA, GDPR and Privacy Shield. They are COPPA (Children's Online Privacy Protection Act) Certified for research with children as well as Privacy Shield Certified, and California Consumer Privacy Act (CCPA) and General Data Protection Regulation (GDPR) ready.

## ACKNOWLEDGMENTS

This research was funded by the Chan Zuckerberg Foundation as part of the Reach Every Reader project. We would like to thank all the participating families for devoting their time to read with us and test out the Smart Speaker app. The feedback we received has been invaluable. We would also like to thank all of the research assistants who have helped sort, clean, transcribe, and code the data. We would also like to give special shout-outs to Scot Osterweil, who designed and created Floppy, Melissa Callaghan, who, among other research and design work, voiced Floppy, and Brandon Hanks, without whose tech wizardry none of this work would have been possible.

## REFERENCES

- [1] Caroline Bologna. [n. d.]. 40 Tweets About Parenting With Today's Technology. *The Huffington Post* ([n. d.]). [https://www.huffingtonpost.ca/entry/parenting-alexa-siri-tweets\\_1\\_6008e784c5b62c0057c35100](https://www.huffingtonpost.ca/entry/parenting-alexa-siri-tweets_1_6008e784c5b62c0057c35100)
- [2] Cynthia Breazeal, Paul L Harris, David DeSteno, Jacqueline M Kory Westlund, Leah Dickens, and Sooyeon Jeong. 2016. Young children treat robots as informants. *Topics in Cognitive Science* 8, 2 (2016), 481–491.
- [3] Mindy Brooks, Ashley Fenwick-Naditch, and Erica Branch-Ridley. 2011. Electric racer: An intergenerational gaming experience designed to promote literacy. In *The New Coviewing: Designing for Learning through Joint Media Engagement*, Lori Takeuchi and Reed Stevens (Eds.). The Joan Ganz Cooney Center at Sesame Workshop, 30–33.
- [4] Jacqueline Burgess and Christian Jones. 2020. I harbour strong feelings for Tali despite her being a fictional character": Investigating videogame players' emotional attachments to non-player characters. *Game Studies* 20, 1 (2020).
- [5] Daria Bylieva, Zafer Bekirogullari, Victoria Lobatyuk, and Tatiana Nam. 2020. How virtual personal assistants influence children's communication. In *Knowledge in the Information Society*. Springer, 112–124.
- [6] Stefania Druga, Randi Williams, Cynthia Breazeal, and Mitchel Resnick. 2017. "Hey Google is it ok if I eat you?" Initial explorations in child-agent interaction. In *Proceedings of the 2017 Conference on Interaction Design and Children*. 595–600.
- [7] Samantha Ehli, Julia Wolf, Albert Newen, Silvia Schneider, and Babet Voigt. 2020. Determining the function of social referencing: the role of familiarity and situational threat. *Frontiers in Psychology* 11 (2020), 3474.
- [8] Carrie A Ewin, Andrea E Reupert, Louise A McLean, and Christopher J Ewin. 2021. The impact of joint media engagement on parent-child interactions: a systematic review. *Human Behavior and Emerging Technologies* 3, 2 (2021), 230–254.
- [9] Nicolas Favez, Chloé Lavanchy Scaiola, Hervé Tissot, Joëlle Darwiche, and France Frascarolo. 2011. The Family Alliance Assessment Scales: Steps toward validity and reliability of an observational assessment tool for early family interactions. *Journal of Child and Family Studies* 20, 1 (2011), 23–37.
- [10] Julian Frommel, Cody Phillips, and Regan L Mandryk. 2021. Gathering Self-Report Data in Games through NPC Dialogues: Effects on Data Quality, Data Quantity, Player Experience, and Information Intimacy. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–12.
- [11] Lauren N Girouard-Hallam and Judith H Danovitch. 2022. Children's trust in and learning from voice assistants. *Developmental Psychology* 58, 4 (2022), 646.
- [12] Lauren N Girouard-Hallam, Hailey M Streble, and Judith H Danovitch. 2021. Children's mental, social, and moral attributions toward a familiar digital voice assistant. *Human Behavior and Emerging Technologies* (2021).
- [13] Omer Gvirsman, Yaacov Koren, Tal Norman, and Goren Gordon. 2020. Patric: A platform for triadic interaction with changeable characters. In *2020 15th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 399–407.
- [14] D Fox Harrell, Dominic Kao, Chong-U Lim, Jason Lipshin, Ainsley Sutherland, Julia Makivic, and Danielle Olson. 2014. Authoring conversational narratives in games with the Chimeria platform. In *Proceedings of the 9th International Conference on the Foundations of Digital Games (FDG)*.
- [15] David Hill, Nusheen Ameenuddin, Yolanda Linda Reid Chassiakos, Corinn Cross, Jeffrey Hutchinson, Alanna Levine, Rhea Boyd, Robert Mendelson, Megan Moreno, Wendy Sue Swanson, et al. 2016. Media and young minds. *Pediatrics* 138, 5 (2016).
- [16] Alexis Hiniker, Amelia Wang, Jonathan Tran, Mingrui Ray Zhang, Jenny Radesky, Kiley Sobel, and Sungsoo Ray Hong. 2021. Can Conversational Agents Change the Way Children Talk to People?. In *Interaction Design and Children*. 338–349.
- [17] Hui-Ru Ho, Bengisu Cagiltay, Nathan Thomas White, Edward M Hubbard, and Bilge Mutlu. 2021. RoboMath: Designing a Learning Companion Robot to Support Children's Numerical Skills. In *Interaction Design and Children*. 283–293.

- [18] Alice Ann Howard Gola, Melissa N Richards, Alexis R Lauricella, and Sandra L Calvert. 2013. Building meaningful parasocial relationships between toddlers and media characters to teach early mathematical skills. *Media Psychology* 16, 4 (2013), 390–411.
- [19] Magnus Johansson, Björn Strååt, Henrik Warpefelt, and Harko Verhagen. 2013. Analyzing the social dynamics of non-player characters. In *International Simulation and Gaming Association Conference*. Springer, 173–187.
- [20] Elizabeth S Kim, Lauren D Berkovits, Emily P Bernier, Dan Leyzberg, Frederick Shic, Rhea Paul, and Brian Scassellati. 2013. Social robots as embedded reinforcers of social behavior in children with autism. *Journal of Autism and Developmental Disorders* 43, 5 (2013), 1038–1049.
- [21] Jacqueline M Kory-Westlund and Cynthia Breazeal. 2019. A long-term study of young children’s rapport, social emulation, and language learning with a peer-like robot playmate in preschool. *Frontiers in Robotics and AI* (2019), 81.
- [22] Kostadin Kushlev and Elizabeth W Dunn. 2019. Smartphones distract parents from cultivating feelings of connection when spending time with their children. *Journal of Social and Personal Relationships* 36, 6 (2019), 1619–1639.
- [23] Irene Lopatovska, Katrina Rink, Ian Knight, Kieran Raines, Kevin Cosenza, Harriet Williams, Perachya Sorsche, David Hirsch, Qi Li, and Adrianna Martinez. 2019. Talk to me: Exploring user interactions with the Amazon Alexa. *Journal of Librarianship and Information Science* 51, 4 (2019), 984–997.
- [24] Brandon T McDaniel. 2019. Parent distraction with phones, reasons for use, and impacts on parenting and child outcomes: A review of the emerging research. *Human Behavior and Emerging Technologies* 1, 2 (2019), 72–80.
- [25] Siri Mehus and Reed Stevens. 2011. Ethnographic field studies of joint media engagement. In *The New Coviewing: Designing for Learning through Joint Media Engagement*, Lori Takeuchi and Reed Stevens (Eds.). The Joan Ganz Cooney Center at Sesame Workshop, 19–22.
- [26] Gigliana Melzi and Margaret Caspe. 2005. Variations in maternal narrative styles during book reading interactions. *Narrative Inquiry* 15, 1 (2005), 101–125.
- [27] Wendy Ochoa, Stephanie M Reich, and George Farkas. 2021. The observed quality of caregiver-child interactions with and without a mobile screen device. *Academic Pediatrics* 21, 4 (2021), 620–628.
- [28] Dimitra Petousi, Akrivi Katifori, Sierra McKinney, Sara Perry, Maria Roussou, and Yannis Ioannidis. 2021. Social bots of conviction as dialogue facilitators for history education: Promoting historical empathy in teens through dialogue. In *Interaction Design and Children*. 326–337.
- [29] Elaine Reese, Catherine A Haden, and Robyn Fivush. 1993. Mother-child conversations about the past: Relationships of style and memory over time. *Cognitive Development* 8, 4 (1993), 403–430.
- [30] Jérôme Rime, Hervé Tissot, Nicolas Favez, Michael Watson, and Werner Stadlmayr. 2018. The diaper change play: validation of a new observational assessment tool for early triadic family interactions in the first month postpartum. *Frontiers in Psychology* 9 (2018), 497.
- [31] Katharina J Rohlfing, Angela Grimminger, and Britta Wrede. 2020. The caregiver’s role in keeping a child-robot interaction going. In *International Perspectives on Digital Media and Early Literacy*. Routledge, 73–89.
- [32] Rachel R Romeo, Julia A Leonard, Sydney T Robinson, Martin R West, Allyson P Mackey, Meredith L Rowe, and John DE Gabrieli. 2018. Beyond the 30-million-word gap: Children’s conversational exposure is associated with language-related brain function. *Psychological Science* 29, 5 (2018), 700–710.
- [33] Michael S Rosenwald. 2017. How millions of kids are being shaped by know-it-all voice assistants. *The Washington Post* (2017), 185e198.
- [34] Meredith L Rowe, Rosa Guzman Turco, and Joseph H Blatt. 2021. Can interactive apps promote parent-child conversations? *Journal of Applied Developmental Psychology* 76 (2021), 101326.
- [35] Kimiko Ryokai, Cati Vaucelle, and Justine Cassell. 2003. Virtual peers as partners in storytelling and literacy learning. *Journal of Computer Assisted Learning* 19, 2 (2003), 195–208.
- [36] Brian Scassellati, Laura Boccanfuso, Chien-Ming Huang, Marilena Mademtzi, Meiyang Qin, Nicole Salomons, Pamela Ventola, and Frederick Shic. 2018. Improving social skills in children with ASD using a long-term, in-home social robot. *Science Robotics* 3, 21 (2018), eaat7544.
- [37] Noah L Schroeder and Chad M Gotch. 2015. Persisting issues in pedagogical agent research. *Journal of Educational Computing Research* 53, 2 (2015), 183–204.
- [38] Hogeun Seo and Claire Shinhe Lee. 2017. Emotion matters: What happens between young children and parents in a touch screen world. *International Journal of Communication* 11 (2017), 20.
- [39] Solace Shen, Petr Slovak, and Malte F Jung. 2018. "Stop. I See a Conflict Happening." A Robot Mediator for Young Children’s Interpersonal Conflict Resolution. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. 69–77.
- [40] Reed Stevens and William R Penuel. 2010. Studying and fostering learning through joint media engagement. In *Principal Investigators Meeting of the National Science Foundation’s Science of Learning Centers*. USA Arlington, VA, 1–75.
- [41] Reed Stevens and L Takeuchi. 2011. The new coviewing: Designing for learning through joint media engagement. (2011).
- [42] Hui-Yun Sung. 2019. Adult mediation of preschool children’s use of mobile technologies in public libraries in Taiwan: A socio-cultural perspective. *Journal of Librarianship and Information Science* 51, 1 (2019), 196–207.
- [43] Weimin Toh and Fei Victor Lim. 2021. Let’s play together: ways of parent-child digital co-play for learning. *Interactive Learning Environments* (2021), 1–11.
- [44] Henrik Warpefelt. 2016. *The Non-Player Character: Exploring the believability of NPC presentation and behavior*. Ph.D. Dissertation. Department of Computer and Systems Sciences, Stockholm University.
- [45] Ying Xu, Dakuo Wang, Penelope Collins, Hyelim Lee, and Mark Warschauer. 2021. Same benefits, different communication patterns: Comparing Children’s reading with a conversational agent vs. a human partner. *Computers & Education* 161 (2021), 104059.
- [46] Ying Xu and Mark Warschauer. 2019. Young children’s reading and learning with conversational agents. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–8.
- [47] Ying Xu and Mark Warschauer. 2020. Exploring young children’s engagement in joint reading with a conversational agent. In *Proceedings of the Interaction Design and Children Conference*. 216–228.
- [48] Andhik Ampuh Yunanto, Darlis Herumurti, Siti Rochimah, and Siska Arifiani. 2021. A Literature Review for Non-player Character Existence in Educational Game. *Cyber Physical, Computer and Automation System* (2021), 235–244.
- [49] Michelle X Zhou, Weixi Chen, Ziang Xiao, Huahai Yang, Tracy Chi, and Ransom Williams. 2019. Getting virtually personal: Chatbots who actively listen to you and infer your personality. In *Proceedings of the 24th International Conference on Intelligent User Interfaces*. ACM, 123–124.