NORTHWESTERN UNIVERSITY

"Hey Google, Do Unicorns Exist?": Children Finding Answers Through Conversational Agents

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Media, Technology & Society

By

Silvia B. Lovato

EVANSTON, ILLINOIS

June 2019

Abstract

Children are known to be curious and persistent question-askers. The pervasiveness of voice interfaces represents an opportunity for children who are still learning to read and write to independently search the Internet by directing questions at conversational agents such as Amazon's Alexa, Apple's Siri, and the Google Assistant. However, little research exists on how children use these agents and on children's conceptual understanding of them. Through a two-week, in-home deployment involving 40 families (children aged 5-8 and their parents), this study offers a description of how early elementary school-aged children ask questions of a conversational agent (the Google Assistant). Through analysis of the usage logs associated with the smart speakers, I report on how often children's questions are transcribed correctly by the technology, how often they receive an answer addressing their question and which questions children choose to ask the conversational agent. Based on log data supplemented by interviews before and after the deployment, I discuss challenges in use and children's perceptions of the technology. Although about 90% of children's questions are transcribed correctly, the system only offers answers that address the questions about half of the time. Most of children's questions are about topics related to culture and science, demonstrating that children choose to use this technology to ask about the world around them. I offer directions for future research and considerations for the design of voice-based conversational agents that aim to support young children's question-asking behavior and subsequent development.

Keywords: children, question-asking, conversational agents, voice assistants

Acknowledgements

Many people contributed to the ideas and the work presented in this dissertation. I start by thanking my advisor and the chair of my dissertation committee, Ellen Wartella, for her unwavering support and encouragement throughout my graduate school career and for inspiring me with her commitment to children and her big ideas. I thank her also for keeping me well-fed, for always being "the good guy," and for being a walking encyclopedia of children's media research that I could consult at will.

I am lucky to have had the opportunity to work with the members of my dissertation committee, who each offered key insights that helped to shape this work. My sincerest thanks to Alexis Lauricella, who met with me countless times over the last five years, helping me organize my ideas, offering advice and kindly challenging my assumptions. I am deeply grateful to Anne Marie Piper for working with me as a co-author on three publications and helping me develop many of the ideas contained in this dissertation. Thank you to Mike Horn for meeting with me several times over the past two years and letting me talk through my research questions and coding scheme, for asking great questions and offering me support and confidence. I thank Lisa Anthony for generously sharing her expertise and advice, for the great conversations we had at IDC' 2016 and for the on-going collaboration they led to, generating many Skype calls from which this work has benefited.

This dissertation would not have been possible without the willingness of the 40 participating families. I thank them for opening their homes to my project, allowing me to record their interactions with the smart speakers and to interview their children.

To my cohort mate Jeremy Foote, thank you for helping me recruit participants and especially for donating hours of your time to help me reorganize data in a way that made part of my analysis possible.

Special thanks go to CMHD research assistants, particularly Su Min You, for her dedication and enthusiasm for the project over the past two years, and for all of our thoughtful conversations about the descriptive codes used in the data analysis. Thank you to Ashlee Mitchell, Allyson Snyder and Johnny Wang for helping recruit, transcribe, code and organize data.

To all of my CMHD friends, past and present, thank you for your support and camaraderie. I could not have completed this work without them. I am especially grateful to Bri Hightower who, during an early workshop session about my dissertation idea, suggested I leave the smart speaker in participants' homes, and to Sarah Pila, for helping me recruit participants and keeping me accountable through the summer. Thank you both for your incredible friendship and support during this dissertation journey.

Finally, I thank my family and friends for their love and support. To the children in my life, especially Luísa, Joana, Julia, Sammy, Anna, Maddie, Ben and Alex, thank you for your inspiring questions and creativity. To my late grandparents, Wanda and Egberto, thank you for your support of my education and for funding the first stage of my American graduate school dream twenty years ago. To my partner, Dhruv Sharma, thank you for all the trips, all the dinners and all the times when you pretended to have to work just so I would keep going. To my whole family, especially my mother, thank you for teaching me to ask questions, think critically and be persistent – all essential tools for writing a dissertation.

Table of Contents

List of Tables and Figures

Tables

Figures

Chapter 1: Introduction

In *The Children's Machine*, Seymour Papert (1993) writes about a 4-year-old girl named Jennifer who was curious about how giraffes, with their long necks, lay down to sleep. She thought he might know, since he was from Africa. Papert describes a Knowledge Machine that a future Jennifer might use, which would allow children to answer their own questions:

> Such a system would enable a Jennifer of the future to explore a world significantly richer than what I was offered by my printed books. Using speech, touch or gestures, she would steer the machine to the topic of interest, quickly navigating through a knowledge space much broader than the contents of any printed encyclopedia (Papert, 1993, p. 8).

A recent report by Common Sense Media (Rideout, 2017) revealed that almost all children in the United States (98%) now have access to a machine similar to the one described by Papert, in the form of a tablet computer or smartphone, a trend that includes low-income families (Kabali et al., 2015). Portable and less expensive than traditional computers, mobile devices also have the potential to afford children more choices in what content they consume and more independence finding information. Because they do not require the operation of a keyboard and mouse or the ability to type, or even read, these devices are more child-friendly. Touchscreens and microphones (along with natural language processing software) allow children to navigate and choose content by swiping and touching images, and search using their voices, by pressing a microphone icon present in the operating system keyboards and search bars of mobile applications such as Google and YouTube, or by speaking to a conversational agent such as Apple's Siri, Amazon's Alexa or the Google Assistant.

While the touchscreen alone allows even very young children to swipe and tap on icons or thumbnail images, choosing from content visible on the device's screen or within applications such as Netflix or YouTube, these choices are still limited to some degree by what apps have been installed on devices – often the prerogative of adults – or, in the case of video apps, by what is chosen by algorithms or human content managers to be featured. The use of voice search – the act of tapping a microphone icon and speaking search terms, which are then transcribed openly into the search field – broadens the realm of accessible content significantly. However, making choices from a list of search results can be less straight-forward to children. Voice search may work well to find content that generally matches children's interests, such as videos about airplanes, cute baby animals or *Minecraft* castle building. Finding the information needed to answer specific questions can be more complicated, requiring scanning web pages or clicking on multiple videos before finding one that contains the desired answer. To this end, conversational agents may be able to assist children more effectively, as they are built to respond verbally to information requests.

Young children are persistent and prolific questions-askers. They are known to ask factual and causal questions about the world around them when they perceive a gap in their understanding (e.g., Callanan & Oakes, 1992; Chouinard, Harris & Maratsos, 2007; Tizard & Hughes, 1984). Children, before they are able to read and write, have historically depended on others for answers to questions they are unable to answer on their own, through logic, observation, or experimentation. The recent proliferation of intelligent conversational agents—particularly digital assistants such as Apple's Siri, Amazon's Alexa and the Google Assistant—make it possible for children who are old enough to articulate a question clearly, but not yet

fluent in reading or writing, to independently search for answers on the Internet by pressing a button, or saying a wake word (e.g. “Hey, Google”), and speaking their questions directly to a device. Such a source of information could also support self-directed learning, which has been shown to promote better retention of information in adults (Markant, DuBrow, Davachi & Gureckis, 2014) and younger children (Partridge, McGovern, Young & Kidd, 2015) than learning from a program designed by someone else.

Conversational agents are available in smartphones and tablets, as well as through home-based smart speakers. Unlike most mainstream tablet computers and smartphones, smart speakers have been marketed to families, with television ads that feature children resolving disputes by asking the speaker factual questions. Home-based smart speakers in particular have experienced explosive growth recently: according to industry reports (e.g., RBC Capital Markets, 2019; Abramovitch, 2018; Nielsen, 2018), 41% of American households had one at the end of 2018, and more than 50% will have one between 2020 and 2022. A Common Sense Media and SurveyMonkey survey of parents of children ages 2 to 8 (2019) found that 59% said their children interacted with a voice-activated assistant, including those on smart speakers and on smartphones and tablets. Additionally, the companies behind conversational agents have announced partnerships with makers of products such as cars, household appliances, and other consumer electronics, indicating that the likes of Alexa and the Google Assistant are about to become even more pervasive.

Despite their growing prevalence, we know relatively little about how children use conversational agents in general and even less about how they might use them to seek answers to children’s questions in particular. While research has recently begun to study children’s use of

conversational agents (e.g. Druga et al., 2017, Yarosh et al., 2018), these studies have so far been based on observations of children using the technology during short sessions outside the home. Further, there is little research on the kinds of questions children choose to ask such interfaces. In spite of their anecdotal appeal to pre- and emerging readers, conversational agents were not developed with children as their primary users and may not be able to transcribe child speech accurately or to understand children's requests. They may also respond with language that is too difficult for children to understand. Missing from the literature is a detailed understanding of in-home use of conversational agents, particularly as children use them together with their parents for an extended period of time, allowing the opportunity to freely experiment with and explore the agent.

This dissertation presents the results of a field study in which 40 families with a child between the ages of 5 and 8 years old used a smart speaker (i.e., a Google Home Mini) for about two weeks. Chapter 2 covers the research literature that informs my project. I start with studies on question-asking from the field of developmental psychology, including children's source selection and cultural differences in question-asking behavior. Next, I discuss the interdisciplinary literature that investigates children's understandings of connected electronic devices and, finally, the human-computer interaction (HCI) and library and information sciences research that studies children's interactions with search agents and, more recently, voice-based conversational agents. Chapter 3 discusses the method used to study children's questions to a conversational agent, including a parent survey, the use of an in-home smart speaker and pre- and post- semi-structured interviews. Additionally, I describe the data analysis approach. Chapters 4 through 6 present the findings. In chapter 4, I describe the participating families' use

of technology and their current information-seeking approaches, both from children's and parents' perspectives. Chapter 5 presents the analysis of the usage logs obtained from the Google accounts associated with the smart speakers used by families, including a quantitative analysis of usage, transcription and answer quality, the subject matter of the children's questions and the challenges faced when trying to ask questions of the conversational agent. Chapter 6 presents an analysis of children's mental models of the devices used, as well as their understanding of the source of the information they hear through the speaker, based on interview data and Likert-scale ratings of the device. In chapter 7, I discuss the implications of this work for system design and information literacy curricula. I conclude with a word about limitations and future research directions for this work.

The analysis presented here focused on questions asked by children to the Google Assistant, and it reveals that, even though children are understood correctly the majority of the time, only about half of the answers they hear address their questions directly. This dissertation contributes one of the first naturalistic studies of how young children use an in-home voice-based conversational agent to ask questions and find information.

Chapter 2: Related Work

While research has just begun addressing the interaction between children and voice-based conversational agents, with even less work concerned specifically with conversational agents and children's questions, the research literature in Developmental Psychology, Library and Information Sciences and Human-Computer Interaction can help predict how such an interaction might unfold. Here I review relevant work on children's question-asking behavior, children's conceptual models of computers and the Internet, earlier work on children's information-seeking behavior and Internet search, and finally, recent studies addressing children's interactions with voice-based conversational agents.

Children's Question-Asking Behavior

Intense question-asking is a common childhood behavior that has long interested the field of developmental psychology. Children ask questions when they perceive an inconsistency, or a gap, in their understanding of the world that cannot be resolved by their own first-hand observation or experimentation (Piaget, 1929). In order to learn about abstract ideas or invisible scientific phenomena, for example, children must rely on information they can obtain from others. Chouinard, Harris and Maratsos (2007) found that children of preschool age asked an average of 76 questions per hour, remarking that they knew that children were really looking for information (as opposed to adult attention, for example) because unanswered questions continued to be asked in subsequent occasions, hours or days later (Chouinard, Harris, & Maratsos, 2007).

Callanan and Oakes (1992) asked 30 parents of children aged 3-5 (ten of each age) to fill out diary entries of children's causal questions (questions about "how things work" or "why

things happen"), including details such as the setting, the parent's response and any follow-up questions by children. They obtained 188 entries containing 479 conversational turns by children (including questions and statements). A detailed analysis of the children's questions and parents' answers showed that the questions reflected the children's curiosity and motivation to learn. One of their hypotheses – that "why" questions would be mostly tied to negative words (not, no, can't, don't, etc.) and used when expectations were violated – was not supported by the study. They concluded that children's causal questions are an important theory-building mechanism that supports conceptual change in childhood. Not unexpectedly, children's causal questions in their study often seemed to be just the beginning of a multi-turn conversation.

Chouinard, Harris and Maratsos (2007) also identify questions as a mechanism for cognitive development. Questions are likely to generate information at moments when children are particularly open to it, which results in better learning (as opposed to information provided when adults decide to provide it). They investigated young children's question-asking in four studies. The first was an analysis of the questions asked by four native English-speaking children between 1 and 5 years of age, based on recordings from the CHILDES database (MacWhinney, 2000). The data revealed that about 25% of the questions asked by the children were explanation-seeking (why and how questions).

The second of Chouinard, Harris and Maratsos' (2007) studies was a diary study of 68 children between the ages of 1 and 5 recruited from a preschool with a diverse SES population. The results revealed that, as with the CHILDES database sample, most of the questions asked by children were information-seeking questions. As children aged, the proportion of explanatory questions grew. Pre-verbal inquiries, not present in the CHILDES recordings, were mostly fact-

oriented (such as “what is this?” or “what is this called?”). In agreement with Callanan & Oakes (1992), findings in this study also revealed that explanatory questions, more common in 4 and 5-year-olds, were “build-up” questions, or questions that occurred after one or more questions had been answered, as part of a conversation.

Taking a different angle on the study of question-asking, Mills, Legare, Bills and Mejias (2010) conducted an experiment to investigate preschool (ages 3-5) children’s ability to use questions to solve problems. They identified three skills needed in order to use questions as a problem-solving strategy: (1) children must be able to determine sources of reliable, useful information, (2) they must be able to formulate effective questions to elicit the required information and (3) they must be able to use the information received to solve the problem. The researchers presented 50 children with sets of cards (‘blickets”) featuring a solid background color and a familiar shape (star, heart, circle, etc.). Two expert puppets were in the room: one was an expert in colors and the other in shapes. Children had to figure out which blicket would open a slot in a box by asking questions of the expert puppets. In the simple trial, cards within a set varied only in either color or shape (i.e. each of the shapes presented on an orange background or only the heart shape presented on each of the background colors). In the complex trial, both colors and shapes varied. Children’s questions were coded as either effective (on-task and able to elicit useful information) or ineffective (off-task, vague, incomplete, repeated, etc.). Subsequently, questions were rated on whether they were directed at the appropriate puppet. The results showed that 3-year-olds had difficulty figuring out who to ask their questions, and they also asked ineffective questions more than half of the time. Four-year-old children were successful at directing their questions to the right expert most of the time, but they were not as

good at formulating effective questions, asking about the same amount of ineffective and effective questions. Only 5-year-olds were successful both at knowing who to ask and at asking effective questions more often than ineffective ones. Although the authors admit this study has limitations (the task was novel; the children were unfamiliar with the "experts" before the experiment, and expertise in "shape" or "color" is an artificial construct, unlike real life experts such as "doctor" or "fireman"), its framing and results are relevant to the issue of whether children could effectively use intelligent conversational agents to find answers to their questions. In order to use such systems successfully, children would need skills similar to those identified by Mills and her colleagues: they would need to know (1) what kinds of questions devices can answer satisfactorily, i.e. what the device's expertise is; (2) how to formulate efficient questions to obtain the desired information and (3) how to use the information obtained to fulfill their needs. The study results suggest that even though younger children could attempt to use voice interfaces, they might face difficulties when trying to formulate questions that can be answered by technology.

In one of the most widely-cited studies about children's questions, Tizard & Hughes (1984) recorded the conversations of 30 four-year-old girls for one afternoon at their homes and two mornings at a nursery school in the United Kingdom. Half of the children were from working-class families and the other half were from a middle-class background. The researchers noticed a marked difference in the number of questions asked by children between the home and school environments: children who appeared lively and engaged at home, constantly talking with their mothers and asking challenging questions that stretched them intellectually, appeared subdued while in school, their interactions with staff limited mostly to monosyllabic utterances.

This difference was even more significant for the girls from working-class families, who could have been intimidated by a classroom environment (i.e., style of furnishings, amount and kind of books and toys available) that was different from that of their own homes. The researchers attributed some of the difference between the amount of questions asked at home and in school not to differences in IQ or ability to ask questions, but to the quality of answers received from the nursery teachers, who were watching many children at once, compared to the more concentrated attention received from the mothers. These results, albeit limited in scope, suggest that encountering a gap in knowledge and having the language skills needed to ask questions is not all that is required for children to ask questions: children's perceptions of the availability of a source to whom they might direct a question also impact their questioning behavior.

Therefore, a constantly-available device with access to endless information such as a smart speaker with a conversational agent, can be an attractive proposition. However, in the context of technology-based sources, it is important to understand whether devices are perceived by children as reliable.

Deciding what information sources to trust. Even though children are thought by classic philosophers to be indiscriminately trusting of those around them (Corriveau & Harris, 2009), recent research shows that, although children often readily trust information from those they have formed attachments with, they can be selective when facing lesser-known sources, basing their choices on prior experience. Building on the experiment described above (Mills, Legare, Bills & Mejias, 2010), Mills, Legare, Grant & Landrum (2011) asked children aged 3 to 5 years to find out which of two cards corresponded to the object inside a closed box. First, children had to choose between a knowledgeable puppet, an ignorant puppet (who would just

say, “I don’t know”) and an inaccurate puppet, who gave absurd, implausible information (such as “people wear coats to stay cool when it’s warm outside”). Then, children had to choose between the knowledgeable puppet, a “guesser” puppet who told them “I’m not sure, but…” then gave them plausible, but inaccurate information, and a puppet who simply gave plausible inaccurate information. Not surprisingly, children had more trouble eliminating the puppets who simply gave plausible but inaccurate information than they had eliminating puppets who pledged ignorance or said they were not sure before answering. However, besides choosing their sources incorrectly, the other main reason why children chose incorrect answers was not gathering enough information before making a choice: researchers concluded that to be successful, children had to decide who to ask, know what questions to ask and then keep asking questions until they had all the information needed to solve the problem. As expected, older children were better at choosing the correct source puppet, but age was not significantly correlated to how effective children’s questions were. Researchers hypothesize that, for young children, generating effective questions is difficult and, therefore, there might be fewer cognitive resources left available to choose the correct source and ask sufficient questions.

Danovitch & Alzahabi (2013) conducted a series of experiments investigating how 3 to 5-year-olds decide which technology-based informants to trust. In each experiment, they showed children videos of an adult user sitting in front of two computers, each with a solid color frame around the screen (red and blue). The adult user would say what he was looking for (e.g., a picture of a ball) and each computer screen would show an image, one corresponding to the query and the other not. In the first experiment, during the training phase, the adult on the video stated he was looking for a picture of a familiar object (e.g., a ball), so the children could identify

the reliable computer. During the trial, the user would say he was looking for a picture of a novel object, like a “toma”. Children were subsequently asked to identify the “toma” from live objects on a table. In experiment 2, during the training phase, the adult was looking for the answer to a simple query (e.g., “What is the color of grass?”) that the children knew the answer to. During the trial, he asked for unknown information (e.g., “What is the weather in Daxtown?”). In both phases, queries could be answered with an image (i.e. a patch of color or pictures of sun or rain). At the end of experiment 2, children were asked which computer had the right answer.

In experiment 1, most children were able to point to the computer that was correct during the training phase, and they were also more likely to choose the object displayed by the reliable computer during the trial, showing that they took previous performance into account, as they would with social agents. In experiment 2, however, while most 4- and 5-year-olds chose the reliable computer during the trial phase, 3-year-olds performed at chance. After each of these first two experiments, researchers asked the children which computer they themselves would use to look up information in the future. In experiment 1, despite the having shown a preference for the reliable computer during test trials, children chose the computer for future use at about chance, which the researchers attributed to children not seeing the reliability of the computer as a stable trait, or the children having a preference for one of the colors in the frames (red or blue). In the second experiment, 88% of all children chose the reliable computer for future use, with only a slight majority of 3-year-olds, most 4-year-olds and 100% of 5-year-olds choosing the reliable computer. This result showed an increase in the use of prior information to choose a source as children aged.

After both experiments, children were asked why they thought the “wrong” computer had been wrong. They mostly said, “I don’t know”; some blamed the user, others mentioned a technical issue such as “mixed up cables”. To further understand children’s explanations for the computer’s errors, researchers performed a third experiment. They used the same videos from experiment 1 (image-based queries) but omitted the test trials. To circumvent the color preference issue, children were asked, right after the training video, which computer they would use in the future and also which computer they thought their teacher should use. Then they were asked why they thought the wrong computer was wrong as a forced choice question, blaming either the user or the computer. While 4- and 5-year-olds showed a strong preference for the reliable computer for future use, 3-year-olds were still at chance. As for the reason for the computer’s errors, there was an order effect in 3-year-olds’ responses: they tended to choose the explanation that was presented last. Four and 5-year-olds were more likely to blame the computer for the error, saying “it didn't know”.

The experiments suggest that children tend to think of and place trust in computers similarly to the way they choose to place trust in people (or other social agents, like puppets). This is convergent with the media equation proposed by Reeves & Nass (1996) and might be even stronger in conversational agents, where the device (or agent) is made to sound like a human. Based on these findings, the still imperfect ability of voice interfaces to understand children’s speech is likely to have an impact on how much children trust voice agents as sources: if Siri, for example, misunderstands a child and responds with an answer that doesn’t make sense, the child could lose trust in it as a source of answers.

Although Danovitch & Alzahabi's (2013) experiments offer relevant information – they are unique in that they attempt to measure preschoolers' trust in a technology-based information source – there are severe limitations. It is difficult to make claims about children's trust in computers or other technology by relying on their observation of a video of someone else using the technology, and it is unclear whether these results would stand should the children look up the information on the computers themselves in a more realistic fashion.

A Word about Theory of Mind. Mastery of classic false-belief tasks (understanding that the information people hold in their heads may not correspond to reality) usually happens around 5 years of age (Wellman & Liu, 2004). However, understanding that people can have different interpretations of the same situation (Interpretive Theory of Mind, or IToM) is an ability that begins to develop a little later, around 6 or 7 years (Lagattuta et al., 2015) and continues to improve into adulthood. Children in middle childhood are in the process of developing an understanding of how someone's background and prior knowledge impacts how they interpret and report information. This ability should lead to increased sophistication in selecting sources of information.

Cultural differences in question-asking. More than graphic interfaces, which make use of images, voice interfaces depend heavily on language use. Because natural language processing depends on a corpus of user utterances to build reliability, the longer a system or agent has been available in a particular language, the better it tends to work. The conversational agents widely available in the market today (e.g., Google Assistant, Apple's Siri, Amazon's Alexa and Microsoft's Cortana) have been developed in English first, therefore there is reason to believe that they work better in English than in other languages, when other languages are available.

With time, and as the technology spreads more widely around the globe, other languages may gradually improve. However, language is not the only culture-related aspect of question-asking behavior in children.

Research has shown that children ask questions at different rates and to different sources in different cultures. Based on the finding that Mexican-American first-graders tended to ask fewer questions than white children in the same classrooms, Henderson & Garcia (1973) devised an intervention to facilitate the development of question-asking skills, focusing specifically on causal questions. They trained half of the children's mothers in behaviors that would better elicit questions from children, such as modeling and cueing (the other half served as a control group). They also modeled question-asking to all children by showing them pictures and giving examples of questions that could be asked about them. Post-treatment testing revealed that all children benefited from the experimenters' modeling of question-asking. The treatment group, whose mothers were trained, asked significantly more questions of the target question-type (causal questions) following the treatment. At the end of the study, mothers and children were taken on a visit to a facility that housed local wildlife. Mothers reported that children asked many more questions about the wildlife, some of which the mothers didn't know the answer to, demonstrating the transfer of the new-found question-asking ability to a new environment. Mothers noted down the questions and a meeting with museum staff was organized to answer them. This study shows that parental behavior can influence children's question-asking habits and that question-asking skills can be modified in first graders.

If children can be encouraged to ask questions of adults by an intervention, it seems possible that a future conversational agent could be programmed to offer encouragement to children, or to behave in such a way as to be seen as an approachable source.

Demonstrating other aspects of cultural differences in question-asking, Gauvain, Munroe & Bebee (2013) analyzed language data collected in 1978 and 1979 from 96 children from four non-Western cultures: Garifuna in Belize, Logoli in Kenya, Newars in Nepal and Samoans in American Samoa. There were 24 children from each culture included in the study, half were 3 years old and the other half were 5. At the time of the recordings, all four cultures could be described as "small-scale traditional" societies, which despite being in the process of abandoning their subsistence economy, still retained some of their indigenous customs. A total of 30-35 observations per child were collected over a 6-week period. The study compared the children's question-asking behavior amongst the four cultures and also Western children (based on the American sample used by Chouinard et al. (2007) which was found in the CHILDES database and also recorded in the 1970s). The main difference found in the nature of the questions was the proportion of information-seeking questions that were of an explanatory, or causal, nature (e.g. why-questions). While in Chouinard et al.'s (2007) sample about 25% of the information-seeking questions were of an explanatory nature, in the non-Western cultures, only about 4.5% of the information-seeking questions fell into that category. Another important difference was that while U.S. children were studied mostly at home with their mothers, to whom they directed almost all of the questions in Chouinard's data, the non-Western children directed about 75% of their questions to peers, since their default environment was peer-dominated. In the comparisons

among the four non-Western groups, the Logoli children of Kenya asked significantly fewer questions than children from the other three cultures, and no explanatory questions at all.

Of the cultural differences highlighted in this study, children's usual target for questions might be the most relevant when considering how children from different cultures might approach a technology-based information source. Since peers are not authority figures, children who are more used to directing questions at peers rather than at parents might be more used to a collaborative style of finding and assessing information and perhaps less likely to receive information without questioning its validity. These children might perhaps attempt to probe or argue with a conversational agent in the same way they do with their usual targets, which is to say, other children. Another possibility is that children used to finding answers collaboratively might try asking questions at the same time, perhaps talking over each other. The cultural setting reported by the authors has, in all likelihood, changed significantly in the past 35 to 40 years, but this comparison draws attention to an important aspect of children's behavior when looking for information. Whether or not children tend to work in pairs or groups when asking questions of a device could impact how well they are able to use technology to find answers.

Children's Understandings of Digital Devices

Beyond question-asking behavior, how children approach a technological source of information can also depend on how they think of the technology itself. Recent reports about how children aged 8 and under use technology (e.g. Rideout, 2017, Common Sense Media & SurveryMonkey, 2019) show that their primary use is for entertainment, in the form of games and videos for touchscreens and music or stories for smart speakers. Eisen and Lillard (2016, 2017) found that children aged 3 to 6 years tended to attribute fewer uses to devices than adults,

that is, they might think of a mobile phone as a music player and of a tablet as a game device but fail to see that both devices could serve both purposes, although this tendency decreased with age. Therefore, if children use technology mostly for entertainment, they may not readily think of the devices as sources of information.

Smart speakers, as well as mobile phones and tablets, are essentially computers connected to the Internet, though the streamlined shape and ease of use of more recent devices may make this identity less obvious to children.

Historically, research has found that older children can, but do not necessarily think of computers as entities able to provide information. Rücker and Pinkwart (2016) present a systematic, interdisciplinary review of studies of children's conceptions of computers. They identify five main ideas children have expressed in studies over the years, between 1968 and 2012: (1) intelligent machines; (2) omniscient databases; (3) mechanical devices; (4) wire networks and (5) programmable machines. While the first two concepts imply the ability to provide information, the others do not. Examples of computers as omniscient, or intelligent agents in the review include a study (Van Duuren, Dossett & Robinson, 1998) that found that children aged 8 and 11 (but not 5-year-olds) believed computers had the results of all possible mathematical calculations already stored in their memory. Two other studies found that 12-to-16-year-olds believed that the entire Internet was stored in one single computer, either the user's own or another accessible through the network (Diethelm, Wilken, & Zumbrägel, 2012; Papastergiou, 2005).

More recent research with younger children still present a mixed picture. A study of Dutch 4- to 7-year-olds' perceptions of their own computer use, most of whom had daily access

to computers both in and out of school, found that the overwhelming majority of young children used computers to play games and that using the computer for a creative or communicative activity or to search the Internet was far less common (McKenney & Voogt, 2010).

One important difference between children's interactions with computing devices via touchscreens when compared to voice-based conversational agents is that the conversational agents may more easily blur the line between machine and intelligent, and perhaps living, being, which could impact how children approach such agents, for information-seeking or other uses. In the opening chapter of her book *The Second Life,* Sherry Turkle (1984) describes the reaction of a group of children to an electronic game of tic-tac-toe. The game was programmed to occasionally make a mistake and "allow" the player to win. Children figured this out but could not replicate it at will. This led to the belief that the game could "cheat", which to the children meant that it was alive – they thought that machines do not cheat. Research on how children understand and interact with robots and with other media provides some insight into how machines that behave contingently (or as humans, to some degree) are perceived by children. Kahn, Gary and Shen (2013) argue that social robots are establishing a new ontological category, distinct from humans, animals or simple artifacts. As children interact with a social robot, they tend to believe that it has rights and feelings (Kahn et al., 2012). At the same time, they are aware of the robot's machine status.

Through a number of experiments, Reeves and Nass (1996) found that people tended to respond to computers and other media as they would to humans, a finding they termed "the media equation." The set of interactions that are specific to computers, whose responses, unlike those of television, are contingent on user input, are studied under an area of research called

CASA (Computers as Social Actors). But is the tendency the same in children? Chiasson & Gutwin (2005) predicted that children would be even more affected by the media equation than adults, since they are more likely to anthropomorphize objects and accept fictional characters as real. They also predicted that providing social cues in interfaces that made interactions more similar to those with people would help children stay engaged in educational activities. To test this, the researchers replicated two classic Reeves & Nass (1996) CASA experiments comparing groups of adults to children aged 10 to 12. In both experiments, they measured the impact of social language – praise in one case and treating the participant as part of the computer's team in another – on user's assessments of their own experiences playing simple games. Surprisingly, they found that, while social language had a positive impact on adults, it had no impact on children. They proposed two explanations for this: one is that children are so affected by the media equation that this overwhelms any difference between experimental conditions (i.e., they would have had a positive experience regardless of the social language in the game). The other explanation is that people who have grown up with computers, as was the case of the child participants, are less susceptible to the media equation than those who learned to use computers later in life, as was the case of the adult participants. One might question the validity of the media equation today, since more adults and most children now grow-up with technology and live immersed in it, which according to Chiasson & Gutwin (2005) would make them less susceptible to the effect. However, computing devices have also developed new ways to behave more like humans, by speaking to us in human-sounding voices, which research has shown to accentuate our inclination to react to them as we would to people (Nass & Brave, 2005).

Conversational agents have existed since before speech recognition, in the form of command line interfaces such as ELIZA (Weizenbaum, 1966), a simple program that simulated a psychotherapist by "mirroring" the user's statements to elicit further input. For example, if the user typed "I am feeling well", it would respond with something like, "Why do you say that you are feeling well?" Turkle (1994) reports that, even though most users who had access to ELIZA were educated enough to understand how it worked, some tried to "talk" to it avoiding phrases or subjects that would result in non-sensical responses from the software, thus preserving the illusion that they were communicating with a person. Given children's penchant for fantasy, it is easy to predict that they would be even more likely to play along with a conversational agent.

If young children approach conversational agents for information as they would approach an adult human, that does not mean they would place unqualified trust on such a system as a source of information. Children's trust in technology sources has been found to be largely based on previous experience, as it is with people. This behavior evolves with age, with 4- and 5-year-olds being more likely to use past experience as a reference than 3-year-olds (Danovitch & Alzahabi, 2013). Because of this, the imperfect ability of voice agents to understand children's speech could have an impact on how much children trust conversational agents as sources: if Alexa or the Google Assistant misunderstands a child and responds with an answer that doesn't make sense, the child could lose trust in it as a source of answers. In fact, anecdotal evidence exists that children do not take Siri seriously (e.g. Shulevitz, 2014). Being the first modern voice-based conversational agent to be widely accessible to the public, Siri was probably the least sophisticated at launch, especially when it comes to understanding children, which can be more challenging than understanding adults.

Understanding children's perceptions of connected devices is only one side of understanding how they use these technologies. Next, I discuss the research that investigates the mechanisms through which children search for information.

Children, Information-Seeking and Internet Search

The context in which children's information-seeking behavior is studied has changed dramatically over the past three decades. Early studies of information-seeking behavior in children (and adults) were carried out in the context of library catalogues, systems designed and managed by humans, and where one's success in obtaining answers depended heavily on understanding the system itself, or speaking its language (Kuhlthau, 1991). Studies of how children search for information in digital interfaces began with the CD-ROM encyclopedias and digital libraries of the 1980s and 1990s, where the realm of information available was static and limited, but the possibility of a full-text search emerged (e.g. Marchionini, 1989), meaning that it became less important to know exactly how others had catalogued texts. Even then, elementary-aged children showed a tendency to use natural language in search fields (Marchionini, 1989), but in a system that was designed to find keywords, this strategy failed, generating no results (Solomon, 1993). Asking a question of Siri, Alexa, or Google Assistant is essentially an Internet search. Even though the experience may be different, what is currently known about how children use search interfaces offers insight into the kinds of information children think can be obtained from the technology and how they attempt to communicate what they want to know to a computer interface.

In a series of studies of seventh graders using the web directory Yahooligans, a child-focused resource managed by Yahoo, Inc. from 1996 to 2006, Bilal (2002) found that older

children (middle-schoolers) consistently preferred to browse the directory than to use the search functionality, largely due to search difficulties related to incorrect syntax and misspellings. Only 50% of the students succeeded at finding answers to specific, fact-based queries given by a science teacher, while 69% partially succeeded at researching a topic more generally using their own queries and 73% succeeded at finding answers to an undirected, self-generated query. Bilal also reported that 13% of children used natural language instead of keywords, a habit seen as a liability at the time, which led to the conclusion that students should receive better web search training. Natural language processing, or the ability of computers to understand human language, has evolved enough as to probably render such practical web search training unnecessary. However, the need to teach children how to use technology provides an opportunity to teach them how it works and where information comes from, themes that may be discussed less often in the absence of such training.

In a slightly more recent study about how children used keyword interfaces (Druin et al., 2009), the researchers found that children ages 7, 9 and 11 had trouble typing, spelling and deciding which words to use as search terms. Specifically, children tended to look at the keyboard while typing, making it difficult to catch typos until the entire word or phrase had been entered or to see the predictive terms offered by the search engine. Parents in their study suggested voice-input as a solution to children's typing and spelling problems. Voice input could certainly solve the typing and spelling issues, however, the study also found other difficulties that would not be eliminated: for example, children had difficulty choosing what words to use when searching and struggled when breaking down a complex query into multiple steps when needed. When asked to find what day of the week the vice-president's birthday would fall on the

following year, none of the children were able to find the answer; the youngest children, age 7, did not even try – possibly due to not knowing how to get the information from a search engine, or simply not understanding the complex, multi-part question themselves. This suggests that even if mechanical barriers are removed, children who are capable of initiating a query may not be able to get to the answers they seek without some form of assistance.

In a larger follow-up study including 83 children (again aged 7, 9 and 11) and their parents (Druin, Foss, Hutchingson, Golub & Hatley, 2010) Druin et al.'s (2009) findings were confirmed and expanded. The researchers identified seven distinct search "roles", or search behavior patterns, displayed by the children, in isolation or combined with one or more other roles. Each of these roles is associated with specific behaviors, motivations for using search, obstacles (such as typing, spelling and reading difficulties, lack of motivation and self-imposed limiting rules) and influencer roles (demonstrator, fixer, mentor). The most common role was that of a *developing searcher*, displayed by 58 of the children. Developing searchers were found to be willing to search but possess a limited command of search tools and, again, a tendency to use natural language. The developing searcher role was most often displayed at the same time as that of *domain-specific searcher*, in which children are comfortable with a few "tried-and-true" resources, usually related to personal interests, and tend to return to those resources repeatedly, even when searching for unrelated information. For example, children attempted to find information about dolphins and about the vice-president of the United States at a gaming website and on spongebob.com, simply because those were resources that they were familiar with. Other, self-explanatory roles identified were power searcher, distracted searcher, non-motivated searcher, visual searcher and rule-based searcher.

The growing ubiquity of voice search raises new questions of how children formulate natural language queries. While Druin and colleagues (2009, 2010) found that children's use of natural language in search engine was problematic, Kammerer and Bohnacker (2012), only a few years later, compared natural language to keyword searches performed by 21 8-to-10-year-olds using Google in German and found that natural language users were more successful than those using keywords. The researchers gave children a two-part task in which the first part was a simple yes/no question (do all kangaroos have pouches?) and the second required a more complex strategy and answer (how do baby kangaroos stay in pouches?). Tasks were given orally, and children could choose what to enter in the search field. Of the 13 natural language users, eight were able to answer both parts of the task correctly, four were able to answer only the first and one was unable to answer either. The eight keyword users, who had to come up with the keywords on their own, fared far worse, with only three being able to answer both queries correctly, three answering only the first and two being unable to answer either.

As we consider younger children using voice interfaces to search, some of the obstacles identified by prior work may lose importance or change in nature as new ones appear. For example, typing and spelling do not apply to voice input, however, voice interfaces may still not always be able to understand younger children's speech or choice of words correctly. Reading search results, when results are not available through sound or video, is likely to pose an even more significant barrier to younger children than it did to the children in Druin et al.'s (200 9, 2010) studies, who were aged 7 and older. In an exploratory analysis of YouTube videos depicting children using Apple's Siri (Lovato & Piper, 2015), we found that young children who succeeded in having their queries understood sometimes were stumped when the device

displayed a screen full of text. Druin et al.'s (2009) domain-specific searchers might become app-specific in this generation. Young children who become comfortable searching inside an application such as YouTube Kids could attempt to use it for queries that would be better served by a different tool.

Children and conversational agents. While earlier research (e.g., McKenney & Voogt, 2010) found that children mostly used computers and tablets to play games and watch videos, seldom attempting to search for information; the proliferation of conversational agents has the potential to change that understanding. Voice interfaces that provide spoken answers make information accessible to pre- and emerging readers, offering an opportunity for younger children to find answers independently, which could also lead to them thinking of connected devices as sources of information.

An emerging literature has begun to examine children's interactions with voice interfaces and conversational agents in particular. Early studies of children interacting with voice user interfaces focused on children aged 7 and older using specific applications, such as homework tutors (e.g., Ward et al., 2011), or tools meant to diagnose speech impairments (e.g. Maier et al., 2011). More recent work has begun to examine the practices of younger children. A study based on an analysis of YouTube videos (Lovato & Piper, 2015) found that children's interactions with Apple's Siri were mostly exploratory, where children attempt to understand the agent, followed closely by information-seeking questions. Druga, Williams, Breazeal and Resnick (2017) carried out an exploratory study focused specifically on children's perceptions of what the authors term intelligent agents (conversational agents enabled by artificial intelligence). Children aged 3-10 years engaged with four devices, including two home speakers (Amazon Echo and Google

Home), a toy robot (Cozmo) and an Android app (Julie Chatbot). Findings show that children perceived the agents as friendly and trustworthy. With regards to intelligence, while preschoolers offered mixed responses, the older children in the group, ages 6 to 10, mostly thought the agents were smarter than the children themselves, and that they largely related smarts to access to information, or the ability to answer questions about topics that were familiar to the children. More recently, preschoolers showed persistence when interacting with a non-responsive voice interface inside a tablet game (Chen et al., 2018). Although a bug prevented the application from receiving sound input from the children, most of them insisted, repeating the utterance, or varying tone and pronunciation, until encouraged by an adult to give up. Yarosh and colleagues (2018) used a Wizard of Oz study (where researchers manually operate a system to make it appear as if it is working a certain way) to investigate children's (5 to 12 years old) ability to ask questions that required reformulation, as well as children's preferences with regards to personalization (being addressed by the agent by name) and personification (the agent identifying itself by name). They found that many children needed assistance with queries that required reformulation, such as those involving comparisons. Also, while children preferred personified interfaces to non-personified ones, personalization made no difference. Beneteau and colleagues (2019, forthcoming) studied how families (N=10 families) with at least one child between the ages of 4 and 17 repaired communication break-downs with Alexa during a 4-week deployment. They found that family members used a variety of strategies to repair communication, including prosodic changes (adjustments to the rhythm of speech), over articulation (exaggerating sounds), and repetition, among others. They note that the burden of repairing communication falls almost completely on users, with Alexa offering little guidance.

While Yarosh and colleagues (2018) focused on the important issue of reformulation using a custom Wizard of Oz interface, the present study looks at question-asking behavior more broadly, analyzing the types of questions children choose to ask when a smart speaker is available in their home for an extended time period, including the challenges they may find besides questions that require reformulation. The current study builds on this growing body of work by examining the naturalistic use of a commercially available conversational agent as part of children's question-asking behavior, a core contribution of this dissertation. More specifically, I aim to answer the following questions:

Research Questions

RQ1: How successful are children in obtaining information from conversational agents?

RQ1a: Is success predicted by age, sex, reading ability or parental education?

RQ2: What difficulties do children encounter when asking questions of a conversational agent?

RQ3: What subjects do children ask about?

RQ4: What are children's mental models of the conversational agent after two weeks of use?

Chapter 3: Method

To understand how children use a conversational agent to find information, I designed a mixed-methods study that included a parent survey, records of actual usage of a smart speaker-based conversational agent and interviews with parents and children before and after using the speaker. Although the usage logs from the smart speaker alone can answer three of the four research questions described above, the parent survey and initial interviews allow us to understand factors that may impact usage, such as prior technology experience and information-seeking habits. Debrief interviews shed light on children's mental models of the conversational agent. To gain additional insight into children's mental models of the technology, I also asked them to rate the smart speaker along 6 dimensions.

Participants

Forty child and parent dyads were recruited[1] from a university database of families who had agreed to be contacted for research, through snowball sampling, social media calls and through fliers posted in public libraries, coffee shops and other community spaces. Half of the participating children identified as female (n=20). Mean child age was 6.99 years (SD=1.22) and there were equal numbers of children (n=10) in each age group (5, 6, 7, 8 years). Children aged 5 to 8 were the target age group for this study because this range allowed me to compare the use of

[1] A word about recruitment: Smart speaker adoption is growing quickly. At less than $50 per unit for the "mini" or "dot" versions, the technology is not inaccessible to most people in the highly-educated, mid- to high-income community around our campus. However, because of privacy concerns (some are afraid the device will record more than what is said right after the wake words), smart speakers are not a universally accepted technology. Since we were only recruiting families who had not yet had a smart speaker in their home, we expected the number of eligible families to decline during the study period as more and more people bought speakers. Additionally, there were several cases when participants initially responded to the recruitment call only to later cancel because their spouse had objected to having the research smart speaker in the home for two weeks. This resulted in an extended recruitment and study period (from March 2018 to March 2019).

this technology among pre- and emerging readers (ages 5 and 6), for whom voice interfaces could increase independence in information access more dramatically, and children who are already able to read. This age range covers the period during which literacy emerges in most children, 5 and 6 years (Teale & Sulzby, 1986), as well as the period right after, during which children tend to be more fluent readers. Additionally, while younger children may be able to articulate questions to the speaker, the developmental psychology literature identifies 5 years as the age when Theory of Mind has developed sufficiently for children to understand what someone might know (Wellman, Cross & Watson, 2001). Five of the children were in (or registered for, in the case of summer participants) preschool, 10 were in kindergarten, 6 were in first grade, 11 were in second grade and 8 were in third grade.

In order to participate, families had to be able to speak English, not have a smart speaker installed in their home at the time of data collection or any time prior, have a working home Internet connection (Wi-Fi), and be willing to have a smart speaker placed in their home for a period of two weeks. Of the 40 families recruited, 29 identified as White, five as Asian, one as Native American, two as African American, two as other or mixed race and one refused to identify. Additionally, three families identified as Latino. Average annual family income was about $135,000 (*SD* =$64,500) and average family size was 4.1 members. Parents reported an average of 5.33 years of post-secondary education. Seven families reported speaking a language other than English at home 50% or more of the time.

Procedure

I conducted a naturalistic study that included two home visits with each parent and child dyad. During the first visit, after obtaining parental permission and consent and child consent,

parents answered a questionnaire and children and parents were interviewed about their question-asking and information-seeking habits.

Parent questionnaire. First, parents were asked to fill out a paper questionnaire while a researcher interviewed the child. The questionnaire took about 25 minutes to complete and contained sections about family technology use, demographic information and measures of three aspects of the child's personality: shyness, impulsivity and curiosity.

The questions about family technology use were partly adapted from the 2017 Zero to Eight Common Sense Media report (Rideout, 2017). The remaining questions in that section asked about technology use specifically as it relates to finding answers to questions, focusing on open-ended questions about search and voice assistant use.

The rationale for choosing to measure the shyness, impulsivity and curiosity dimensions of children's temperaments was that shy children could be more reluctant to speak to a machine in front of their family and therefore experiment less with the smart speaker, resulting in fewer opportunities to learn how to successfully use it and, consequently, lower success rates. Conversely, more impulsive children might be more willing to "jump right in" and experiment more with the device, resulting in higher success rates. Curiosity could also drive use and experience. The shyness and impulsivity measures came from the Children's Behavior Questionnaire (Rothbart, Ahadi, Hershey & Fisher, 2001), a validated caregiver report assessment of temperament at ages 3 to 7 years. The curiosity scale was developed and validated by Piotrowski, Litman & Valkenburg (2014) as a parent-report measure of epistemic curiosity (EC) in children ages 3 to 8.

The complete questionnaire can be found in Appendix A.

First interview. I conducted a semi-structured interview with the child while the parent filled out the questionnaire. The interview started with general interest questions (e.g., "What do you like to do when you come home from school?") that were meant to build rapport and inform subsequent questions. The child's interests gave me a starting point to ask how they ask questions and find information about topics they are curious about and interested in. If children discussed using technology to find answers, I asked them to explain how the technology worked. I also asked them to explain what the Internet is, if they brought it up. Children were reassured that there were no right or wrong answers and that the researcher was interested in how they thought.

After completing the questionnaire, parents joined their children in the conversation with the researcher and were asked to discuss what they usually do when they do not know off-hand the answer to their child's questions. They were also asked to recall the child's past questions and to describe any evolution they could remember in the child's questioning behavior. A list of interview-guiding questions can be found in Appendix B.

At the end of the first interview, children were given a basic reading assessment (i.e., the Word Reading Test, form A (Chall, Roswell, Curtis, Strucker, 2003)), which consists of five lists of ten words each, in increasing difficulty. It estimates reading grade level (varying from 0 to 10) based on word recognition. Average grade estimated was 3.55 (*SD*=3, range: 0-10, median 3).

Smart speaker deployment. At the end of the first home visit, a Google Home Mini device was installed in the home. It was tied to a unique Google account accessible only to the research team. I selected the Google Home Mini because the device provides log files that include raw sound files (the Amazon Echo did not at the time the study began) and because I wanted to avoid additional complexity associated with accidental purchase requests (i.e., Alexa might misunderstand a query and initiate a series of questions about making a purchase). Additionally, popular press reviews at the time (e.g., McGregor, 2017, November) rated the Google Assistant as better at answering questions than Amazon's Alexa.

Eight identical speakers were used in the study. Each device was marked with a sticker containing an ID number and a telephone number that families could call if they encountered technical problems. The family was allowed to choose where in the house the device should live but were asked to avoid spots near equipment that produced noise (e.g., TV) or areas that were not regularly frequented by the child. Of 40 families, 24 chose to place the smart speaker in the living room, 12 in the kitchen, 3 in a playroom and 1 in the child's bedroom. Both the device and the Google account were left at their default settings (female voice, American accent). No third-party content accounts (e.g. Spotify, Pandora, Netflix) were made available, but devices could play music using Google's own free music service, which does not play specific tracks or artists, but does play related music stations. Families were told that the researchers would have access to a log containing their interactions. A researcher demonstrated basic functions of the device, such as volume controls, and conducted

Figure 1: A Google Home Mini device such as the ones used in the study.

voice training with the child, which involved having the child follow prompts to say the device's wake words a few times, in order to "teach" the device to recognize the child's voice. Participants were told that anyone in the household could use the device without restriction, provided they were told that researchers would have access to their interactions. The child was then encouraged to ask any question they wished. If the child did not come up with a question of their own within a couple of minutes, they were encouraged to ask about the weather.

The research team monitored usage remotely, and families were contacted if no use was detected in more than three days, to check for technical problems. I allowed families to keep the device for 2 to 3 three weeks, depending on their availability during that time period (i.e., to accommodate summer travel). When known, travel days were noted in the records of usage. Participants kept the device for an average of 15.4 days (min=14, max=22, median=14.5).

Debrief Interviews. At the end of that period, child and parent pairs were interviewed again in their homes about their experience with the smart speaker. Additionally, similar to Druga et al. (2017), I collected Likert-style ratings from the child using a six-point smiley-o-meter scale (three negative frowns and three positive smileys, representing -3 to +3) in order to determine children's conceptual models of the smart speaker. Children were first given a positive/negative forced choice (i.e., is it friendly or not friendly?), and then asked to choose the smiley or frowny face that best represented the degree to which they agreed or disagreed. I used this approach to ask whether the child thought the device was friendly, smart, alive, trustworthy, safe, and funny.

Finally, parent and child pairs were compensated with a $50 gift card. This amount was equivalent to the market value of the Google Home Mini and of the Amazon Echo Dot at the

time, allowing the family to replace the research unit with one of their own if they desired. A different Google account was created for each device. Devices were factory reset between families, and account settings were reset to default values. All interviews were video and audio-recorded.

Data Analysis

Data from the parent questionnaire was entered into a spreadsheet and imported into JMP Pro 14 (statistical software) for analysis.

Initial interviews with parents and children were transcribed and analyzed using a thematic analysis approach (Braun & Clarke, 2006). Since the interviews turned out to be more structured than initially planned (all, or nearly all, participants answered the majority of the questions), answers were compiled in a spreadsheet. For example, participant's descriptions of their existing practices for finding information were extracted from interview transcripts and transferred to a column in the spreadsheet. Next to this column, additional columns were created for coding. Each set of answers was coded by two researchers. Discrepancies were discussed and reconciled.

Given the naturalistic focus of this study, my main source of data were device logfiles capturing in-home usage. Usage logs from the devices were downloaded and entered into a spreadsheet for analysis. A total of 12,622 interactions were recorded among all the families, with an average of 286 (*SD*=218) interactions per family (range: 58-1,506). Each interaction included the date and time, transcription, raw audio file and response provided by the conversational agent. Given the study's focus on question-asking behavior, the research team then analyzed all utterances to identify those that were questions (a total of 5,068 interactions).

Other utterances (e.g., commands such as volume controls, requests for music or news; interactions with *actions* (Google's name for third-party software such as game and story applications) and attempts at social interaction, such as "Good morning" or "I'm home!") were not included in the analysis. Although I had conducted voice training with the focal child in each family hoping that this would help identify which utterances were theirs, these were often mislabeled or unidentified. Therefore, the research assistants and I listened to each interaction and manually identified the speaker in order to identify the focal child's questions. At the same time, we coded whether the child's question was transcribed correctly. We then coded the subject of the focal child's questions. We used an iterative process where two researchers first assigned descriptive codes to questions (e.g. animals, sports, public figures) and then created larger categories to contain the original descriptions (see list of larger categories and descriptive code examples in Table 6). This resulted in 47 initial question subject areas (see Appendix C for a complete list of descriptive codes), which we further refined and grouped for a total of 8 subject areas. Each question was also coded based on whether the resulting answer addressed the question fully, partly, not at all, or was refused (i.e., the agent responded that it could not help with the question). For this coding, we were as literal as possible. If the response was vague, or meant to be humorous, but did not provide the information requested, it was coded as not addressing the question. Each usage log was double coded by two researchers. Discrepancies were discussed and reconciled.

To understand whether children's success rates at obtaining full answers from the smart speakers are predicted by age, sex, parental education level, reading scores, degrees of shyness, impulsivity or curiosity, regression analyses were conducted. Additionally, I compared the rate

of correct transcriptions and full answers received by children during week 1 to those from the remainder of the time of deployment, to understand whether children's ability to obtain full answers improved with practice.

Debrief interviews with each parent and child pair were transcribed and analyzed interview data for emergent themes following a thematic analysis approach (Boyatziz, 1998, Braun & Clarke, 2006), which involved a process of inductive coding in order to identify patterns in parents and children's descriptions of their experiences. Children's ratings of the device using the smiley-o-meter were entered into a spreadsheet for analysis. Finally, I conducted linear regression analyzes to determine whether children's ratings of the smart speaker could be predicted by sex or age.

Chapter 4: Existing Household Technology and Answer-Seeking Practices

In order to understand how the smart speaker may fit into parents' and children's information-seeking habits, I begin by reporting on the technology environment that our participants live in, using data from parent questionnaires and participant interviews conducted before the smart speaker was installed in the home. This includes (1) the technologies that are currently available in homes – a question adapted from the nationally representative 2017 Common Sense Media Census (Rideout, 2017); (2) parents' descriptions of their children's use of technology, including the activities for which the children use the devices they have access to, the children's use of search functions and of phone- and tablet-based conversational agents (e.g. Siri, the Google Assistant); (3) how parents mediate technology use, including how independently children are allowed to choose what they do on connected devices (computers, tablets, and smartphones), how often children are allowed to have mobile devices in their bedrooms overnight and whether parents use software to monitor or block unwanted content; and (4) children's and parents' explanations of how they search for answers to children's questions.

Technology in Homes

I first examined technology access in participants' homes. Results show that these families have technology access comparable to that found in the Common Sense Media (CSM) Census (Rideout, 2017). Table 1 shows the percent of families in my sample (N=40) that have each technology in their home, side-by-side with nationally representative data from Rideout (2017). Ours being a mostly wealthy and highly educated group, it would not have been surprising if they had more technology in their homes than the general population. However, that is not the case.

Table 1. Technology in Participants' Homes

Percent of children that live in homes with:	**Participants**	**CSM 2017**
Television	92.5%	98%
An antenna (to watch over-the-air TV)	25%	29%
Cable or satellite TV	40%	65%
DVR (a device to record TV shows to watch later)	27.5%	44%
A device that allows you to watch online video on your TV (such as Roku, Apple TV, Chromecast, etc.)	77.5%	75%
Video subscription service (such as Netflix, Amazon Prime, Hulu, etc.)	95%	72%
DVD Player	65%	86%
Computer (laptop or desktop)	92.5%	91%
Video game console (such as X-box, PlayStation, Wii)	55%	69%
A handheld video game (such as Game Boy, PSP, Nintendo DS)	15%	31%
Tablet (such as iPad, Samsung Galaxy Tab, Microsoft Surface or Kindle Fire)	85%	78%
iPod Touch or similar	12.5%	21%
E-reader (such as a Nook or Kindle)	30%	29%
Virtual Reality headset (such as Oculus Rift, Google Cardboard, PlayStation VR)	5%	11%

There are two important considerations that can be speculated to explain the differences between our sample and the general population. The first is that the study required that participants have home Internet access and not have a smart speaker. The Internet requirement could have driven some low-income families away and also meant that even if a family was low-income, they perhaps prioritize Internet access more than other low-income families. Given the growing penetration rate of smart speakers and the overall high SES of our participants, the no smart speaker requirement might have resulted in a group that is perhaps more cautious than average towards technology. This might explain the lower rates of ownership than the general population in several items (e.g., videogame consoles, VR headsets).

The second consideration is that technology use and attitudes toward it are constantly changing, so differences between our participants and the general public could be attributed to

the two years that have passed since the 2017 Common Sense Media Census (Rideout, 2017) data were collected. For example, while only 40% of our participants have a cable or satellite TV subscription, compared to 65% of the population in Rideout's study (2017), 95% of our families have a streaming subscription, compared to 72% in Rideout's (2017) sample, suggesting that consumers may be migrating from cable to services such as Netflix and Hulu.

The questionnaire also asked which technologies children owned personally. Tablets were by far the technology most often owned by children among our participants, with 22 children (55%) owning one. Next came educational games (e.g., the Leap Pad), which 6 children (15%) own, followed by iPod Touch and handheld video games (3 children, 7.5% each), and smartphones, owned by 2 children (5%). During interviews, parents clarified that these smartphones did not have a SIM card or a data plan and were meant for use over Wi-Fi only.

Children's Use of Connected Devices

To understand what children currently do with Internet-connected devices, we asked parents how often children used tablets, smartphones and computers to do each of a number of common activities. Six frequency options were offered: "has never done" (0), "less than once a week (1), "once a week" (2), "several times a week" (3), "once a day" (4) and "several times a day" (5). Table 2 shows the means and standard deviations for each activity frequency, sorted by decreasing frequency for all participants.

Table 2: Activity Frequency on Connected Devices

Activity	All Participants	5-year-olds	6-year-olds	7-year-olds	8-year-olds
Watch videos	M=3.08 (SD=1.34)	M=3.3 (SD=1.64)	M=3 (SD=1)	M=3.22 (SD=1.39)	M=2.8 (SD=1.4)
Play games	M=2.82 (SD=1.48)	M=2.9 (SD=1.96)	M=2.8 (SD=1.31)	M=2.55 (SD=1.5)	M=3 (SD=1.25)

Watch TV or movies	M=2.79 (SD=1.34)	M=3.22 (SD=1.72)	M=2.6 (SD=1.07)	M=2.7 (SD=1.64)	M=2.77 (SD=0.95)
Listen to music	M=2.73 (SD=1.48)	M=2.55 (SD=1.87)	M=2.4 (SD=1.26)	M=2.77 (SD=1.78)	M=3.2 (SD=1.03)
Look at photos or videos	M=2.65 (SD=1.12)	M=3.66 (SD=1)	M=2.5 (SD=0.85)	M=2.11 (SD=0.78)	M=2.4 (SD=1.26)
Read books	M=2.42 (SD=1.98)	M=2.4 (SD=2.41)	M=1.8 (SD=1.81)	M=3.2 (SD=1.75)	M=2.3 (SD=1.94)
Draw or write	M=2.35 (SD=1.98)	M=2.33 (SD=2.34)	M=2.3 (SD=1.7)	M=2.6 (SD=2.01)	M=2.2 (SD=2.15)
Take photos or videos	M=2.16 (SD=1.25)	M=3.11 (SD=1.53)	M=2.1 (SD=1.19)	M=1.44 (SD=0.52)	M=2 (SD=1.12)
Look something up/search	M=1.89 (SD=1.39)	M=2 (SD=1.87)	M=1.3 (SD=0.95)	M=2 (SD=1.56)	M=2.3 (SD=1.06)
Talk to Siri, Google Assistant or similar	M=1.65 (SD=1.35)	M=1.9 (SD=1.66)	M=1.5 (SD=1.18)	M=1.8 (SD=1.40)	M=1.4 (SD=1.26)

Most activities average between “once a week” (2) and “several times a week” (3). However, looking for information (*M*=1.89) and interacting with a conversational agent (*M*=1.65) happen, on average, less than once a week. The most frequent activities are what one might call close-ended, or activities that do not require any creative input, such as watching videos (*M*=3.08), playing games (*M*=2.82) – save for the instances of game-playing that refer to Minecraft and similar games –, watching TV shows or movies (*M*=2.79) and listening to music (*M*=2.73). Open-ended activities such as drawing and writing (*M*=2.35), taking photos (*M*=2.16), typing search terms (*M*=1.89) or speaking to a conversational agent (*M*=1.65), are reported by parents as undertaken less frequently by this group. When asked what their children’s favorite activity was to do on devices, only 3 (7.5%) parents mentioned looking up pictures or information. However, 25 (67.5%) parents mentioned “YouTube”, “YouTube Kids” or “videos” (excluding those who mentioned movies, TV shows, Netflix, etc.), which could also

include videos that were found when in search for answers to questions. One parent mentioned using the Google Assistant on a smartphone as a favorite activity.

Children's Use of Search. Parents were asked whether their children had ever used connected devices to search for information, either by typing or by using voice input. Additionally, they were asked to describe their children's use of phone and tablet-based conversational agents. Over half of the parents (26, 65%) reported that their children used search, either by voice or by typing. Half of those children (13, 32.5%) were reported to use voice input to search for pictures or videos. For example, the mother of participant (age 7, female) reported that her daughter uses the Google Assistant on a parent's smartphone to say "show me pictures of" to look up animals so she can draw them. The father of another participant (age 6, male) reported that his son "will use voice-to-text to get answers to various math problems, ask general information questions, and search out specific YouTube videos (e.g. for specific Minecraft videos or creepy spider videos)." Parents also reported that 18 children (45%) used Siri or the Google Assistant. The uses mentioned most often included asking information-seeking questions, asking it to play music and also asking it "silly questions." For example, the mother of a participant (age 8, female) wrote, "She was looking up the temperature outside. She was able to find what she was looking for without any help, using Siri." Another mother reported that her son (age 8) "asks Siri to play music quite often. He asks Siri to play *Imagine Dragons* and *Adele*." A father said his daughter (age 7) "used Siri to be silly and ask it silly questions." A mother reported that her son (age 8) "randomly asks [Siri] math questions, or random joke questions such as "what color is your hair?"."

It is worth noting that parents' and children's descriptions of children's technology use do not always coincide. For example, the mother of a participant (age 5, female) marked in the questionnaire that her child had never used a voice assistant like Siri, Alexa or the Google Assistant. However, her daughter demonstrated to me how she uses Siri on an iPad during our interview.

Parents' Mediation of Technology Use

How independently children are allowed to use the technology available to them is relevant to our understanding of their information-seeking habits. Children who are able to choose what they do on devices may be able to experiment more than those who are more closely supervised. Such experimentation may result in more flexibility in how children think about the devices. The questionnaire asked parents how independently they allowed children to choose what to do on each connected device used (tablet, smartphone, computer). Table 3 shows the mean and standard deviation for each device, with 1 being "independently," 2 being "mostly independently," 3 being "under some supervision," and 4 being "under strict supervision". The data are examined by age of the child.

Table 3. How independently do children decide what to do on the device?

Device	All participants	5-year-olds	6-year-olds	7-year-olds	8-year-olds
Tablet	M=1.93 (SD=0.82)	M=1.83 (SD=1.06)	M=2.4 (SD=0.67)	M=1.89 (SD=0.78)	M=1.55 (SD=0.52)
Smartphone	M=2.75 (SD=1.06)	M=2.28 (SD=1.15)	M=3.11 (SD=0.93)	M=3.2 (SD=1.03)	M=2.4 (SD=0.97)
Computer	M=3.13 (SD=0.83)	M=3.5 (SD=0.75)	M=3.5 (SD=0.53)	M=3 (SD=1.05)	M=2.7 (SD=0.67)

Tablets are the devices children are allowed to use most independently, which corresponds to the fact that more than half of participant children have their own tablet.

Smartphones and computers, conversely, tend to belong to parents. It is not surprising, then, that their use would be more closely supervised. Parents may be concerned that independent use by children could carry the risk of the children finding inappropriate content and also possibly cause harm to the device itself or to parents' data and communications (i.e., accidentally deleting files or sending messages to work contacts).

When asked how often a tablet, smartphone or computer were in the child's bedroom overnight, 27 parents (67.5%) checked "never" and 6 parents (15%) checked "hardly ever." "Every night," "most nights," and "some nights" were checked by 2 parents (5%) each. Six (15%) parents reported using software or a device to monitor or limit their children's access to the Internet.

Parents were asked whether their child had ever found content they wished the child had not seen. Twenty-five parents (62.5%) reported that their child had found objectionable content. The most common case was YouTube videos that played after a video the child had been watching, through the auto-play feature. About one-third of the responses were not specific to the nature of the content that was considered inappropriate. Other reports most often included bad language and violence. Parents also mentioned commercial content that resulted in purchase requests, exposure to children or child characters that were mean to others, and exposure to scary content (Halloween, horror movie). When discussing what action was taken after the objectionable content was found, the parents most often reported closer monitoring of the child's device use and family discussion about why that content was not appropriate. For example, a mother reported that her son (age 8) found a 'scary video' containing violence. "It came up as a YouTube suggestion in the side panel. We discussed it as a family once he told us he'd seen

something scary (we couldn't retrace the steps and see it ourselves). We encourage our children to tell us about inappropriate or scary content they come across, we make it clear that they will not get in trouble, and that this is one of the dangers of the internet we have to learn to navigate as best we can."

How Children and Parents Find Answers to Children's Questions

In the following paragraphs, as well as in additional interview data reported in chapters 5 and 6, I use ranges rather than specific numbers of participants. "Most participants," or "the majority" means more than half of the participants (i.e. more than 20). "Many participants" means between 25% and 50% (i.e. 10 to 20). "Several" means between 5 and 10. "Some" or "a few participants" means fewer than 5.

A variety of strategies were used by children to find answers, including varying degrees of technology use.

Choosing the right source. Most children in the study reported asking a grown-up as their first method to find an answer when they have questions. However, even the youngest children (age 5) articulated a strategy, however simple, for choosing which grown-up to ask. This included parents, who sometimes have different specialties, grandparents, teachers and other community members. For example, one of the youngest participants (age 5, male), explained that, when he has questions about soccer, he asks his grandfather. For other questions, he may ask one of his parents, or a teacher. Another participant (age 8, male), said that he gets "all those nature answers" from his father, who is a scientist. Teachers came up as the source of choice for several children because, for example, "they can read us a book about it" (age 5, female) and because "teachers are supposed to be really smart" (age 6, male). Other children

brought up people in occupations related to their questions. One participant (age 6, female), who had a question about how flour is made, said she would go to a bakery to ask people there. Similarly, she would ask a zoo-keeper to find out why zebras have stripes. Children also mentioned asking other children who have expertise in what they want to know. For example, a child (age 7, male) who had questions about roller coasters at a local amusement park said he would ask friends who have been to the park. These findings agree with the research on source selection by children, which shows that children as young as 5 are able to choose sources based on expertise fairly reliably (e.g., Harris, 2012; Mills et al., 2011, Mills & Landrum, 2014).

Where do children think grown-ups find information? When children said they would ask their parents to answer a question they had, I asked the children to think about how parents might find the answer if the parent didn't know. Some of the younger children thought parents would ask other adults, such as "a friend" (of her mother's; age 5, female) or "the robot teacher," – referring to the instructor of an extracurricular activity she had taken part in – in the case of a participant (age 6, female) who wanted to know "why robots imitate people", or "the engineer," to find out how to build structures that are "not wobbly". However, most children thought their parents, when faced with a question they did not know the answer to, would resort to technology, especially smartphones. For example:

Researcher: "When grown-ups don't know something, how do you think they find out?

Child (age 5, male): "By asking Siri."

Choosing between people and technology. Older children were more likely to report choosing adults for some questions and technology for others. Popular culture questions, for example, are often directed straight to technology. The same participant (age 8, male) who said

that he would ask his father about nature, explained that he would ask Siri about *Yu-Gi-Oh* cards, because his parents would not know. Another child (age 8, male) reported searching on the computer for instructions on how to build a shed, or how to do things on *Minecraft*. His mother later told us that he asks her or his father questions about people's behaviors and motivations, such as why a child in his class acts in a disruptive way. However, in many cases, parents control devices, either physically or with passcodes, so children must ask for permission before using technology to find answers.

Parents and siblings as proxies to technology. The children reported sometimes asking parents or older siblings directly to look up information, rather than asking them a question and letting them decide how to find the answer. This happens mostly due to either access restrictions (the parent controls the device or holds the password to the device) or because the child needs help entering search terms or speaking to a conversational agent or a voice-to-text search interface. The questions most often directed to technology were about popular culture and instructions for various crafts or science activities (e.g., how to make slime). For example:

Father of a 6-year-old female: "Last night, we were watching High School Musical 2, and [she] wanted to know if the main characters Troy and Gabriela had ever kissed, right? […] And so [she] was like, 'Dad! Google it! Look it up.' So we did, and we found a video out there, with my iPhone, and we found a video from High School Musical 3 that they had kissed."

Independent search. Many children reported searching for answers themselves, either by typing or using voice input into search fields, or by speaking to a conversational agent. Some of the time, children had to ask permission or obtain a passcode before searching, but not always. Children were more likely to report searching independently for instructions or for more generic

interests rather than specific questions. At least four 5-year-old children demonstrated their use of voice search or spoke with Siri during the interview. In one case, the parent seemed surprised that his child (age 5, male) was able to find dinosaur videos on YouTube by using the voice search functionality. The mother of a participant (age 5, female), showed us her daughter's search history on the YouTube app. It contained terms like "a bug", "big kids songs", and "how people exercise" – the child demonstrated this last search by using the microphone icon to enter terms into the YouTube search field during the interview. Another parent told us that her son (age 8) has used her phone extensively to speak to Siri and to search the Amazon app to find out how much certain toys cost.

At the same time, several 7- and 8-year-old children had heard of Siri but not used it and had only used Google occasionally on the computer and with parental assistance, possibly a reflection of parents' approaches to managing children's use of technology.

Parental approaches to searching for answers. Based on questionnaire and initial interview data, I identified three parental approaches to searching for answers to children's questions using technology. Additionally, a few children tend to ask parents mostly questions that are not fact-based. For fact-based questions, they are able to undertake Internet searches independently.

First, a few of the parent participants explicitly said that they try to keep their children away from technology and prefer to show them how to find information in books or other sources. Two families pointed to print encyclopedias they keep in the living room. Another mother said that, when she does not know the answer to one of her child's questions, she tries to write the question down so they can look up the answer together at the library. In these families,

when parents do resort to technology to find an answer for children, they do so privately and then relay the information to the child. For example:

Mother a 6-year-old female: "…if I don't know, I mean, I'll say 'I don't know.' That happens a lot. I can't think of what [the questions] are but I'll say, 'I don't know what that answer is, what do you think?' But as you probably picked up, we don't usually then go investigate it. […] We do have an old encyclopedia, but I try, and I can speak for myself, to not say 'let's just go straight to the Internet.' So sometimes that leaves the question unanswered, because I don't want to just go straight to ... We do Google things, of course, but sometimes maybe the question gets dropped because I don't pursue it that way."

In a second approach, used by several of the families, parents do the searching themselves, choose the result on which to click or tap, and then show the result – usually a photo or video – to the child. The mother of a participant (age 5, male) explained that often, she finds that showing a video works well when answering children's questions, but she wants to ensure that the video is appropriate, and that's why she looks it up privately first.

Thirdly, in most of the families I interviewed, parents let children see what the parent is doing as the parent searches – or search collaboratively, helping the child spell the search words or talk to Siri or the Google Assistant – and then explore the results together. This sometimes evolves into more independent use of search. For example:

Mother of a 6-year-old male: "We started to look up some videos about scuba diving because he was asking a lot of questions about that, like, how it's done and who does it and can kids do it… So we looked up some videos—some instructional videos—and now he knows how to find videos on YouTube, on my phone, so occasionally he will use the phone, and he will use

the Google search box, just to voice-search it, or he'll ask me sometimes how to type in a word to search for."

Finally, a few of the parents reported that children (8-year-olds) are able to find answers online independently, directing to parents mostly questions that are not factual, but philosophical or about behaviors and relationships.

In summary, participating parents report that children use tablets and computers mostly to consume streaming media and play games, with more open-ended activities such as drawing, taking photos or searching for information happening less frequently. Albeit not frequently, at least 80% of the children in the study do speak with a conversational agent at least occasionally. Parents allow children to use tablets fairly independently; less so computers and smartphones. With the exception of the independent searchers mentioned above, children approach adults and other children with questions, but are discriminating when choosing sources, often directing questions to people the children perceive as knowledgeable. Most parents search online for information to answer questions that go beyond their existing knowledge, and they mostly allow children to participate in the process, raising the children's awareness of technology as a source of information.

Next, I will report on how children – and their families – used the Google Assistant that was left in their homes to find answers to questions.

Chapter 5: Children Using Smart Speakers to Find Answers

To understand how children used the smart speakers to find information, I downloaded the usage logs for the periods each family had the Google Home Mini in their homes and entered them into spreadsheets for analysis. Logs from two of the families were excluded, one due to a technical problem that caused most of the audio files to be missing[2] and the other due to an outlier usage pattern that generated a log 5.7 standard deviations longer than the mean. Therefore, the analysis of the usage logs includes N=38 families.

Throughout the in-home deployment, families interacted with the smart speakers 10,865 times. An interaction here is defined as any time a person says the wake words ("Hey Google" or "OK, Google") and generates an entry in the usage log. This includes incomplete interactions. It does include those for which no audio recording is available (labeled simply as "used assistant" or "unrecognized voice command" in the logs). Focal children were responsible for 5,439 interactions, or 50% of all interactions (M=143.13, SD=112.5, see Figure 2 for distribution).

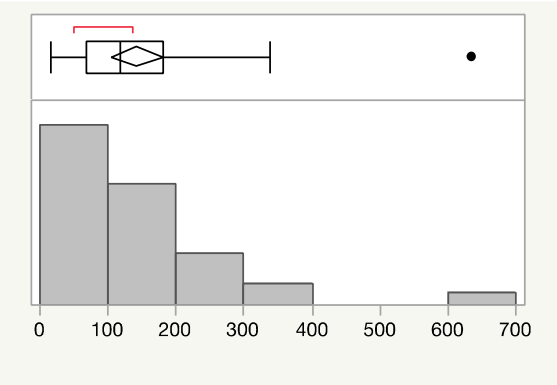

Figure 2: Distribution of Child Interactions

The remainder of the interactions were attributed to siblings, parents and occasional visitors. Of all interactions, 4,461 (41%) were questions. The remainder were commands, interactions with third-party *actions* (stories, games and other content) or

[2] There seems to have been a change in how audio files are logged by Google: unlike what happened at the beginning of data collection, now the usage log only contains audio files for speakers whose voices are recognized by the system. Because Google's Voice Match has proven imprecise with children in this study, this resulted in less than 10% of the audio files for this family being available.

social interaction (e.g., “Hey Google, I’m home!”). On average, each of the 38 families asked 117.39 questions during the deployment (*SD*=76.9; range=20-327 questions). Of children’s interactions, 47% were questions (2,587 out of 5,439 interactions). In total, focal children were responsible for 58% of all the questions asked – a total of 2,587 questions out of 4,461 questions asked across all families. On average these children asked 67.95 questions each (SD=43.74, range 8-187 questions, see Figure 3 for distribution.). Table 4 below shows the total number of questions and the mean number of questions asked by children of each age.

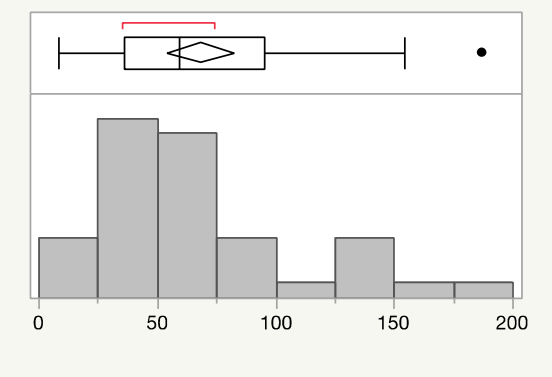

Figure 3: Distribution of Children's Questions (raw count)

Table 4. Total Questions and Mean Number of Questions by Age

	All Focal Children	**5-year-olds (n=10)**	**6-year-olds (n=10)**	**7-year-olds (n=9)**	**8-year-olds (n=9)**
interactions	5,439 (*M=143.1)*	797 (*M*=79.7)	1097 (*M*=109.7)	1684 (*M*=187.1)	1861 (*M*=206.8)
questions	2,587 (*M*=67.9)	535 (*M*=53.5)	523 (*M*=52.3)	683 (*M*=75.9)	841 (*M*=93.4)
% questions	47%	67%	48%	40%	45%

Pearson’s correlation shows a significant positive association between age and overall interactions ($r(37)=0.375$, $p=0.02$), but not between age and the proportion of questions asked ($r(37)=0.267$, $p=0.1045$). Questions make up a higher proportion of 5-year-old’s interactions with the smart speaker, with two-thirds of all interactions being questions, while fewer than half of interactions are questions in the other age groups.

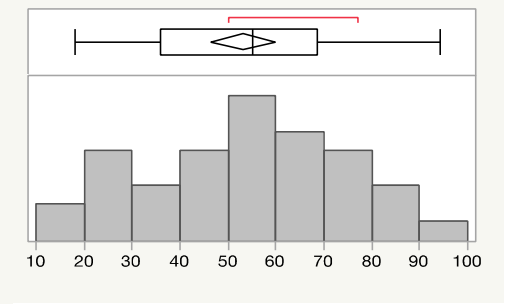

Figure 4: Distribution of the Percentage of Children's Interactions that Are Questions

To understand whether household composition was related to question-asking behavior, such as through the need to share the device or modeling behavior between siblings, I analyzed family size and birth-order of the focal children. Families ranged in size from 3 to 6 members (excluding infants). The majority of families had 4 members (n=25), a few families had 5 members (n=6), 3 members (n=6) and 6 members (n=1).

Family size seemed to influence the number of questions asked by the family overall, with larger families asking fewer questions, but only up to five members (*M*=143 questions for 3 members; *M*=116.4 questions for 4 members; *M*=69.66 questions for 5 members, *M*=275 for 6 members). The mean number of questions asked by the focal child (*M*=104.33 questions asked by children in 3-member families; *M*=64.52 in 4-member families; *M*=44 in 5-member families and *M*=79 in a 6-member family) is lower in larger families again, with the exception of the focal child in the one family of six. The percent of total family questions that were asked by the focal child (78% in 3-member families, 57% in 4-member families, 67% in 5-member families and 29% in a 6-member family) does not appear to be related to family size.

Of the 38 focal children, 22 had older siblings, 11 were the oldest and 5 were only children. Children with older siblings asked an average of 61.5 questions, or on average 54.5% of their family's questions. Children who are the oldest in their family asked an average of 66.6 questions, or 69% of the questions asked by their families. Only children asked an average of 99.2 questions, which amounted to 81% of the questions asked by their families. These numbers show that, not surprisingly, only children seem to be able to ask more questions, probably because they don't have to share the device with other children.

Transcription and Answer Quality

For children to achieve success in finding information with the help of a conversational agent, they have to be understood correctly by the system. There are two parts to this: the first is the software's ability to transcribe what children say correctly. The second is the software's ability to interpret what children mean in order to provide a response. To measure how well the software transcribes children, two research assistants and I listened to each utterance and compared it to the transcript provided in the log. To measure how well it understood what children wanted to know, we rated the responses as to how well they answered the child's questions (fully, partly, not at all or refused, which is when the assistant says "Sorry, I don't know how to help with that" or another such answer).

To understand how well the Google Assistant transcribes what children say, we coded each question as being transcribed correctly or not. Our analysis shows that focal children's questions were transcribed correctly about 90% (range 65-100%) of the time. Split by age, average transcription accuracy was M=82% (SD=14%) for 5-year-olds, M=95% (SD=6%) for 6-year-olds, M=90% (SD=8%) for 7-year-olds and M=93% (SD=5%) for 8-year-olds. Not surprisingly, accuracy appears to improve with age. A linear regression predicting question transcription accuracy based on age of the child (calculated at the end of the smart speaker deployment) was significant ($F(1,37)=5.63$, $p=0.0231$, $R^2=0.135$). Sex, grade in school, and personality traits (shyness, impulsivity and curiosity) were not significantly related to transcription accuracy. Reading score was mildly significant as a predictor of the rate of questions transcribed correctly ($F(1,37)=3.55$, $p=0.0676$, $R^2=0.08$), suggesting a possible

relationship between the ability to articulate questions and reading ability. However, reading score is also strongly correlated with age ($r(37)=0.7364$, $p<0.0001$).

The transcription accuracy found in our study appears to be an improvement over results reported by Kennedy and colleagues (2017), which found that the speech of children of 5-6 years of age was transcribed correctly about 50% of the time when repeating after a researcher, and about 18% for spontaneous speech (Kennedy et al., 2017).

However, even with mostly correct transcripts, only about 50% of children's questions received a full answer, revealing that their difficulties in obtaining answers are not related to how well they can pronounce questions. Table 5 below shows the percent of the time that children's questions were transcribed correctly and answered fully, partly, not at all, or refused (regardless of transcription accuracy).

Table 5. Answer Quality by Age

Questions transcribed correctly		**Answered**[3]			
		Fully	*Partly*	*Not at all*	*Refused*
5-year-olds	439 (82%)	227 (42%)	48 (9%)	87 (16%)	145 (27%)
6-year-olds	494 (95%)	284 (54%)	34 (6%)	63 (12%)	107 (21%)
7-year-olds	620 (90%)	341 (51%)	63 (8%)	97 (14%)	135 (19%)
8-year-olds	790 (93%)	432 (51%)	69 (7%)	137 (16%)	174 (22%)
All children	2342 (90%)	1284 (50%)	214 (8%)	384 (14%)	561 (22%)

Since we did not code answers for accuracy or how accessible they would be for children, the best measure I have of children's success in obtaining answers from the smart speaker is the rate at which their questions received a full answer. Not surprisingly, the rate of questions

[3] To a very small portion of children's questions, the conversational agent offered alternative questions to which it could provide answers (e.g. "I am not sure, but I can tell you the answer to a similar question."), coded as "alternate".

transcribed correctly by the technology was correlated with the rate of full answers received ($r(37)=0.5476$, $p=0.0004$). Simple linear regressions of the rate of full answers on age, sex, reading score, grade in school, and personality traits (shyness, impulsivity and curiosity) were not significant.

Parental education level, computed as years of post-secondary education of the participating parent, was a negative predictor of the rate of correct transcriptions children obtained ($F(1,37)=6.39$, $p=0.0160$, $R^2=0.15$). The level of education of the participating parent also appears to be a negative predictor of the rate of full answers received by the child ($F(1,37)=6.73$, $p=0.0136$, $R^2=0.15$), but when controlling for the rate of correct transcripts, the effect disappears. Parental education level was not a significant predictor of smart speaker usage by any measure (number of family interactions, family questions, total child interactions, or total child questions). Parental education level was a significant predictor of the child's search use frequency prior to the smart speaker deployment ($F(1,37)=4.30$, $p=0.0450$, $R^2=0.10$), however, no significant correlations were found between the child's prior search use and volume of use of the speaker or the children's rate of correct transcriptions or full answers.

To look at this finding from a different perspective, I asked whether the rate of questions refused by the speaker (rather than given a full, partial or mismatched answer) was predicted by parental education level. The result was statistically significant, with children of highly educated parents having a higher proportion of their questions refused ($F(1,37)=11.95$, $p=0.0014$, $R^2=0.24$). Age is not a significant predictor of the rate of questions refused. The rate of questions refused, not surprisingly, is strongly predicted by the rate of correct transcripts ($F(1,37)=15.31$, $p=0.0004$, $R^2=0.30$).

Eleven parents in the sample reported speaking a foreign language at home, seven of them 50% or more of the time. Since the data were collected in an academic community, there was a possibility that a high level of parental education might be correlated to a foreign language being spoken in the home, accounting for the high rate of international scholars on our campus. Speaking a foreign language at home, however, was not found to be correlated to parental education level in our sample.

To test whether speaking a foreign language at home was a predictor of the rate of correct transcriptions, I created a group of English-only homes (n=27), one of homes where a foreign language was spoken less than 50% of the time (n=4) and another where a foreign language was spoken 50% or more of the time (n=7). Children in foreign language families were slightly less likely to be transcribed correctly, but the difference was not statistically significant (p=0.12).

The education level of both parents, computed as the average between the years of post-secondary education of each parent, was not a significant predictor of the percentage of correct transcriptions of child questions, child questions answered in full or any usage measures.

Additionally, none of the usage measures are predictors of either the rate of full answers or the rate of correct transcripts for focal children's questions.

Evolution of Use

I was interested in whether the rates at which children's questions are transcribed correctly by the software improve over time, as well as in whether the rate at which they receive full answers improves, through adaptation to the technology. To test for this, I compared the rate of correct transcripts and the rate of full answers received between the first week of use and the remainder of the time (roughly another week) using paired t-tests. There was no significant

difference in the rate of correct transcriptions between week 1 (M=0.91, SD=0.12) and week 2 (M=0.86, SD=0.2) of the deployment; $t(36)$= -1.219, p=0.2306. Similarly, there was no significant difference in the rate of full answers in week 1 (M=0.48, SD=0.17) and week 2 (M=0.52, SD=0.25) of the deployment; $t(36)$= -1.137, p=0.1313.

One caveat to this approach is that about 75% of questions asked by focal children were asked during the first 7 days of use, accounting for a more intense exploratory period following placement of the device. This renders the comparison between week 1 and week 2 uneven. Therefore, the mid-point of the deployment may not be the right inflection point for comparison. However, usage was not evenly distributed over the deployment or over different children, and children may adapt to the technology at different rates (i.e., for some, adaptation might have taken place over a short period and for others not at all), so it was difficult to identify another point at which to divide the period.

To further investigate how usage evolved over the period, I created daily counts for each child's total number of interactions and total number of questions. Figure 5 below shows the evolution of the average number of focal child interactions per day, broken down by age group.

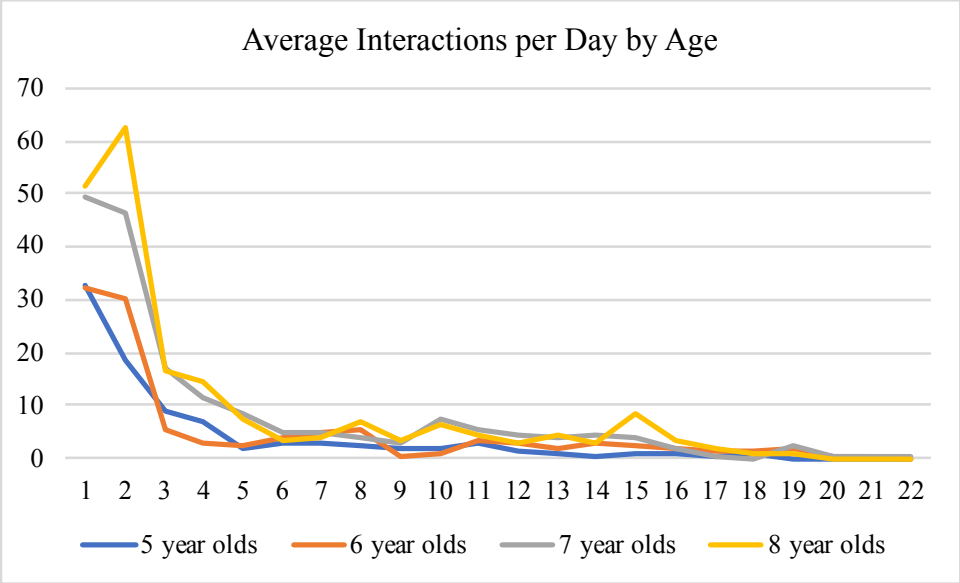

Figure 5: Average number of interactions per day since deployment, examined by age group.

Note that interactions include anything the child said to the smart speaker after the wake words: commands (volume, requests for stories and music, etc.), interactions with third party software (*actions*) as well as social language and questions. Figure 6 shows the number of questions asked per day, again analyzed by age group.

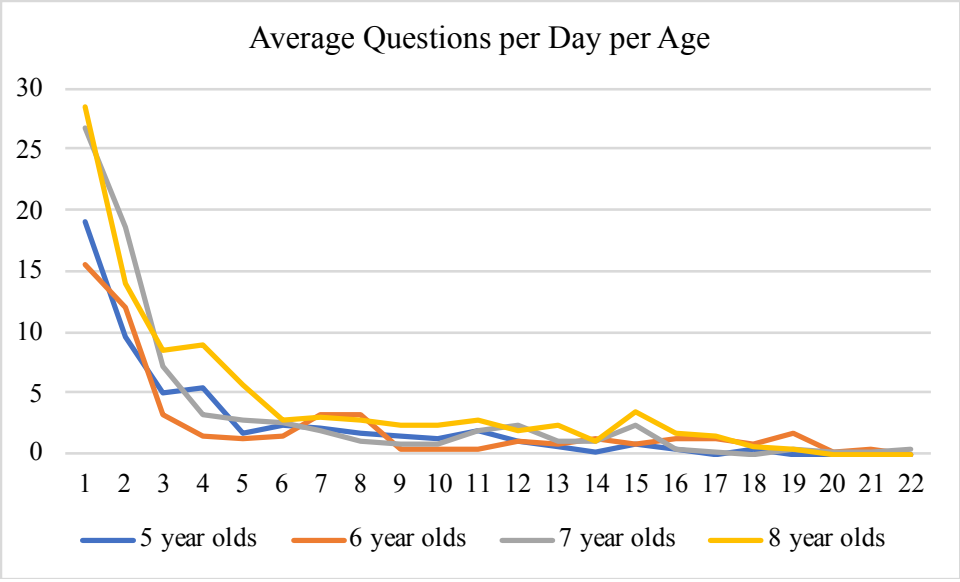

Figure 6: Average number of questions asked by focal children per day since deployment, analyzed by age group.

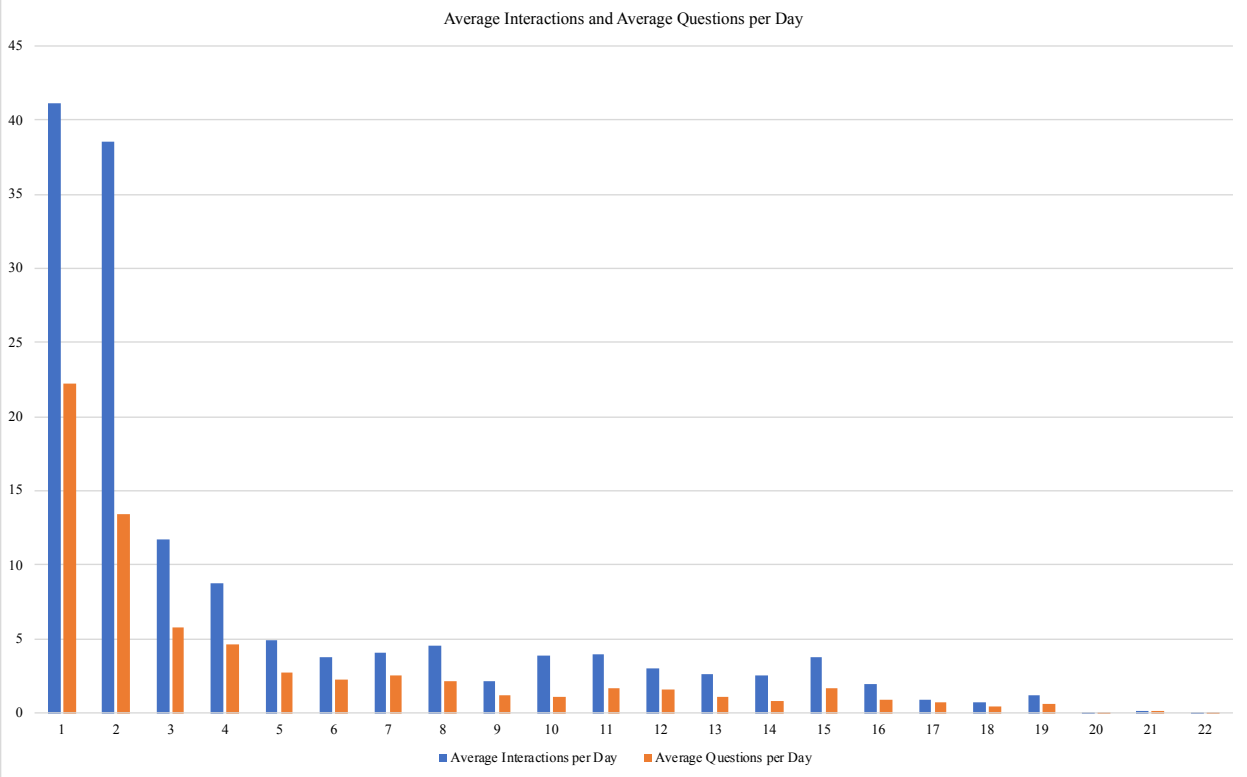

Figure 7: Focal Children's Average Interactions and Questions per Day

Figure 7 shows a comparison between average total interactions and average questions, per day of deployment, for all children.

As the preceding figures show, there was no large variation between age groups in how the use of the smart speaker evolved over the two-to-three-week period. The slight bump on day 15 corresponds to the last day of deployment for most families, which led some of the children to ask a few last questions because the speaker was going away.

Subject Matter of Questions

To understand the topics and domains of children's question-asking behavior, we coded all questions with descriptive codes that were subsequently grouped in eight main subject areas (see Appendix C for the full list of codes). Table 6 below shows the eight main subject areas, along with examples of descriptive codes and sample utterances.

Table 6. Subject Areas Focal Children Asked About

Subject Area	**Example Descriptive Codes (Subcodes)**	**Example Utterances**	**Total Instances**
Culture	history, television shows, fantasy, sports	"How old is Leo Messi?" "Do unicorns exist?"	574 (22%)
Science & Technology	plants, animals, engineering	"How are pillows made?" "What's the fastest animal in the world?"	497 (19%)
Agent	agent, other agents (Siri, Alexa)	"How old are you?" "What's your favorite color?"	468 (18%)
Practical	weather, recipes, directions	"What's the weather today?" "Where's the closest park?"	457 (18%)
Language	meaning, spelling, translations	"What does sassy mean?" "How do you spell Arizona?"	212 (8%)
Math	calculations	"What is 8 divided by 2?"	152 (6%)
Personal Information	user's name, age, family information	"What's my name?" "How old is (sibling's name)?"	132 (5%)
Jokes	joke questions	"Why did the chicken cross the road?	35 (1%)

Culture, which includes pop culture and celebrity questions, accounted for 574 of focal children's questions (22%), the most of any category. It was followed by the science & technology category, including questions about animals, plants, space and the human body, among others, accounting for 497 questions (19%). Questions that were of a practical nature, such as questions about the local weather, local resources or how to prepare food accounted for 457 of all questions (18%). The remaining subject areas that children asked about represented a smaller percentage of their total questions (language: 212 questions, 8%; math: 152 questions, 6%; jokes: 35 questions, 1%).

Two question categories are specifically about the technology: questions coded as "agent", about the conversational agent itself, and questions coded as "personal information", which are about the child and his or her surroundings and family and attempt to explore how much the agent knows about them (e.g., "What is my brother's name?"). Of all the focal

children's queries, 468 questions (18%) asked the conversational agent questions as if it was a person – the "agent" in our code (e.g., "What's your favorite ice cream flavor?"). This was not, however, evenly distributed. One child alone accounted for one sixth of those questions (78) and more than half (*n*=24) asked 8 or fewer questions about the agent. In the category personal information (132 questions, or 5% of the total) one child, a 6-year-old girl, was responsible for 66 out of the 132 questions. She succeeded in introducing herself to the agent early in the 2-week period and taught the system the names of her family members. She then asked multiple times a day about those names (i.e. "What's my name? What's my dog's name? etc.), as if to check if the device still remembered them. (Note that this table does not include questions for which coders could not identify a subject area).

Further analysis of focal children's questions revealed that the younger children in the study, those aged 5 and 6 years, appear to have slightly different interests than those age 7 and 8. Figure 8 below shows the average percentage of each subject in children's overall questions.

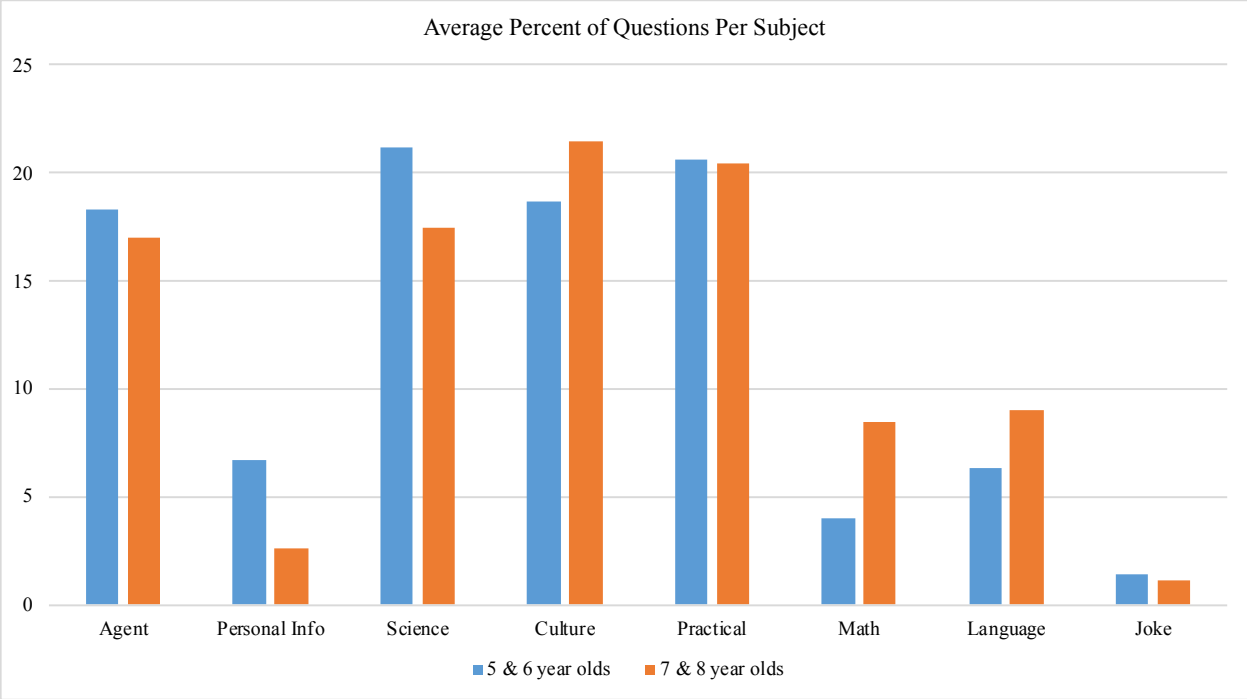

Figure 8: Average Percent of Questions Asked Per Subject

Even though older children asked the conversational agent fewer questions about science and personal information, and more about culture, math and language, these differences are not statistically significant.

Table 7 (below) shows the break-down of questions by subject and by family member category (i.e. focal child, parents or siblings). Note that this count does not include visitors such as friends and relatives, whose questions added to the total count of questions asked by the family.

Table 7. Questions by Subject and Family Member Category

Subject Area	**Focal Child**	**Parents**	**Siblings**	**Total**
Culture	574 (M=15.10, SD=18.06)	157 (M=4.13, SD=6.27)	406 (M=10.68, SD=29.81)	1137
Practical	457 (M=12.03, SD=12.31)	242 (M=6.37, SD=7.62)	124 (M=3.26, SD=6.23)	823
Science & Technology	497 (M=13.07, SD=14.76)	81 (M=2.13, SD=2.88)	156 (M=4.10, SD=7.24)	734
Agent	468	51	178	697

	(M=12.31, SD=15.62)	(M=1.34, SD=2.07)	(M=4.68, SD=8.24)	
Language	212 (M=5.58, SD=8.20)	53 (M=1.39, SD=1.91)	47 (M=1.23, SD=2.29)	312
Math	152 (M=4, SD=9.33)	8 (M=0.21, SD=0.58)	55 (M=1.45, SD=3.67)	215
Personal Information	132 (M=3.47, SD=10.67)	22 (M=0.58, SD=1.44)	35 (M=0.92, SD=3.13)	189
Jokes	35 (M=0.92, SD=1.84)	7 (M=0.18, SD=0.73)	30 (M=0.79, SD=2.65)	72
Total	2,527	621	1031	4179

Parents, perhaps not surprisingly, asked more practical questions (e.g., "What is the weather?") than any other category, followed by culture questions (which include questions about sports, entertainment, and public figures). It is important to note that adults, knowing that researchers would have access to usage logs, might have been intimidated and asked fewer curiosity questions than they would have otherwise. Like focal children, siblings asked more culture questions than other categories; this was followed by agent and science questions. Most children in our sample had older siblings (n=22), who were responsible for most sibling questions. However, siblings as young as 18 months asked questions.

Challenges

The majority of questions that were refused (i.e., the agent responded that it was unable to answer the question) presented fluency, wording, or content challenges. For example, one child (age 5, female) asked, "How does the how do people who know what the weather is like how do they know?" Another (age 6, female) asked, "What the flowers inside of them?" Other refused questions resulted from wording or phrasing mistakes, such as "When was the nearest snow storm?" (age 6, male). Following this, an adult rephrased it as "the most recent snowstorm." Another child (age 5, female) asked, "How far can English count?" as part of a

sequence of questions about counting, possibly meaning to ask, “how far can one count in English?” Yet another child (age 7, male) asked, “What's the weather in Lake Michigan in the water?” possibly meaning to ask about the water temperature.

Other questions were refused due to lack of context or details, such as *“*What the heck is that?” (age 5, male), “What’s the best-selling children’s author?” (age 6, male), “How many Iowa games were there in the past?” (age 6, male) and “What’s the thing on the side?” (age 7, male). Similar to lacking context, other questions sought advice or opinion from the conversational agent, to which it was unable to respond. Examples of these questions include: “When is a good time to play?” (age 5, male), “What should we bring to the beach today?” (age 5, female), “How do you run away from homework when your dad tells you to do homework?” (age 7, male) and “How hard is it to build a house?” (age 8, male). Other refused questions were broad in nature, such as “Where do people come from?” (age 6, female), “How is everything in the world made?” (age 5, female), or too detailed, such as “Can you return the case that originally came with the Amazon Fire 7 kids edition and trade it in for a different color?” (age 8, male). The Google Assistant was also unable to respond to fantasy-related questions such as “Can ghosts fly?” (age 5, male), and “Why do zombies eat people?” (age 5, female).

Although we did not code for question format, I noticed that “why” and “how” questions, called explanatory by Chouinard and colleagues (2007) and considered an important theory-building mechanism by developmental psychologists (e.g., Callanan & Oakes, 1992), seemed particularly hard for the system to answer. Consider the following exchanges between the conversational agent (CA) and a 6-year-old girl (C):

C: “Do squirrels like pumpkins?”

CA: “Here's a summary from the website chicagotribune.com: Pumpkins can attract snacking squirrels. A squirrel stands on pumpkins carved into Halloween jack-o'-lanterns Oct. 25, 2004, in Washington. There's this: Squirrels don't even particularly like pumpkins.

C: “Why do squirrels eat pumpkins if they don’t like them?”

CA: “Sorry, I don’t know how to help with that.”

The same participant tried to engage the Google Assistant in an exchange about the lifespan of dog breeds, connecting a previous answer to a request for further explanation:

C: “How long does a golden retriever live?”

CA: “The lifespan of the Golden Retriever breed is typically between 10 and 12 years.”

C (5 minutes later): “How long does a Labradoodle live?”

CA: “On the website coppercanyonlabradoodles.com, they say: Labradoodles are relatively long lived as a breed and an average life expectancy is between 12 and15 years.”

C: “Why do labradoodles live longer than golden retrievers?”

CA: “Sorry, I don’t understand.”

Another participant (age 5, female) asked, “how is paper made?” four times across three different days. The first time, she received instructions for making paper airplanes. The other three, she heard the following:

CA: “On the website wonderopolis.org, they say: To make paper from trees, the raw wood must first be turned into "pulp." Wood pulp is a watery “soup" of cellulose wood fibers, lignin, water, and the chemicals used during the pulping process.”

Though this answer partly addresses the question (it was coded as a partial answer), it is an incomplete explanation of how paper is made (i.e., only covering how pulp is made). Moreover, the system repeated the same information multiple times rather than viewing this child's attempt at repetitive question-asking as a way of seeking further or different information.

Other "how" and "why" questions were met with a standard refusal answer from the speaker (i.e., "Sorry, I don't know how to help with that."): "How do you put a computer together?" (age 7, male), "Why do devices need to be charged?" (age 7, female) and "How do we make friends?" (age 8, female).

Data from the debrief interviews provides a more complete picture of the complexities of question-asking behavior and challenges children experienced in receiving answers. Analysis of those interviews alongside logfile data reveals three specific challenges for effectively receiving answers from the conversational agent.

First, children and their parents reported that responses were long or required interpretation. Some of the questions that children asked were met with responses that might contain the answer, but not in a straight forward way. For example, a parent in one interview brought up an incident in which their child (age 6, male) asked, "Who is the second fastest person in the world? The conversational agent responded:

"Here is some information from the web that might possibly help. On the website YouTube.com they say Yohan Blake has been training intensive ahead of the Rio 2016 Olympic where he hopes to compete for the gold medal against Usain Bolt who has the current world record."

After replicating this exchange during the interview, we asked the participant what he thought this meant. He replied, “I have no idea.” What parents meant by “requiring interpretation” could also be described as needing an alternate phrasing, requiring prior knowledge, or requiring inferential reasoning about subcomponents of the question. In this example, the child’s lack of understanding of the response might come from the fact that understanding that Yohan Blake might be the second fastest person in the world requires the child to know (1) that having the “current world record” is the same as being the fastest and (2) that someone who is training in hopes to compete for the gold with the fastest person might be the second fastest.

A second challenge observed across the data is that answers that were meant to be funny to adults were often confusing to children. For example, a mother of a 5-year-old girl explained: “I didn't like that when she asked about unicorns it said that [they are everywhere]; it made it seem that unicorns are real, because I think it was more geared towards adults, so they find it funny. So, if a child is asking a practical question then, you know, it's going to give sort of a misleading answer.”

Looking at the usage logs for further context, the child asked, “Do unicorns exist?” to which the agent responded, “I believe in unicorns. They sound something like this: [sound effects]. There are unicorns everywhere; even in the tech industry.” For devices such as conversational agents that may be shared among multiple household members, tailoring responses based on who is asking the question seems like an essential step for future development of the technology.

A third challenge that emerged from our analysis is that follow up questions required that the user restate the information just mentioned in the previous question, and children often failed to do that. Although conversational agents remember some information over the course of multiple interactions (e.g., user's name), it appears that they do not currently integrate previous exchanges as context for future questions. For example (child age 5, male):

C: "You know how your freezer works?"

CA: "Here's a summary from the website greenlivingideas.com: Essentially, your fridge and freezer work by continually turning a refrigerant from a liquid to a gas, and then condensing that gas back into a liquid. That liquid is then pumped through a capillary which removes any moisture or contaminants before being sent to your fridge/freezer's evaporator coils."

C: "What's the name of the gas?"

CA: "Here's a summary from the website examples.yourdictionary.com: Oxygen is a pure gas because it is made of one type of item. Pure gases may also be compound molecules. Carbon dioxide would be considered a pure gas, but it is also a compound molecule."

Mother: "What's the name of the *gods* in the freezer?" (gas was incorrectly transcribed as gods)

Both the child and the mother appear to expect the agent to understand a reference to a previous exchange, as a person would, but the agent does not yet support this sort of contextual understanding across sequences of children's questions.

My analysis shows that even though the automatic speech recognition (ASR) technology has developed the ability to understand children about 90% of the time, it is still only able to

respond with an adequate answer in 50% of the cases. Almost half (47%) of children's requests of the conversational agent were questions, suggesting that this is an important use case for them -- although it must be said that the first interview, which asked about finding answers to questions, could have primed them to ask more than they would have otherwise. Children do attempt to obtain information from this technology. However, there are hurdles that stand in their way. One is that children, who are still developing their Theory of Mind, do not always know how much context to provide with their questions. Further, children do not always use standard language to communicate what they want to know. Children seem to expect the conversational agent to remember what was just said, as a human would in a conversation. Finally, children often lack background knowledge to be able to extract what they want to know from the answers provided.

In the next chapter, I examine children's conceptions of the smart speaker after using it for two weeks.

Chapter 6: Mental Models of the Conversational Agent

As part of understanding how children interact with conversational agents to obtain information, I was also interested in their mental models of the agent. By mental model here I mean the naturally evolving models that children formulate as they interact with the system (Norman, 2014), especially their ideas about how it finds answers to their questions.

During the second interviews, after having had the smart speaker at home for two weeks, children (N=40) were asked several questions that shed light into how they think of the conversational agent (i.e., the smart speaker technology). First, they were asked how they referred to the device when talking to family or friends. The overwhelming majority said they simply called the speaker "Google" or "the Google," which is not surprising since that is the name the device responds to (through the wake words "Hey, Google" or "OK, Google"). The few children who referred to it differently used descriptive words (e.g. "the Google machine;" "the round gray thing"). I also asked children to explain how they would describe the smart speaker to a friend their age who had never seen one before. The descriptions were largely uniform (e.g., "It's a speaker you can ask questions of."). This may have been influenced by the fact that I encouraged children to ask questions of the device when it was set up.

Secondly, I asked children how they thought the speaker knew the answers to the questions they asked. They were reassured that there was no right or wrong answer, that I was interested in how they thought about it and encouraged to guess if they didn't know.

Most of the younger children (5 and 6-year-olds) described how the speaker works in terms of actions it would perform, as a person would, such as looking up information on a smartphone. One child (age 5, male) explained, "I think she looks things up on her phone."

Another (age 5, female) said, “It looks it up on [its] phone, if it has a phone. […] Just knows stuff. Or it reads books.” Another child (age 6, female) (C) explained to the researcher (R):

C: “Well, I think the Google, it just thinked and thinked and thinked and it got smarter. It is smart, it's learning each day and every day. “

R: “And how does it learn? How do you think it learns?”

C: “By using its phone.”

R: “It has a phone?”

C: “Yeah, just you can't, can't see the hands.”

One child (age 6, male), spoke of Google directly: “It asks another Google, Google tells it, and then I don't know. Maybe it searches online and gives me an answer.”

These responses echo prior work regarding children’s perceptions of other intelligent technologies, which found they tend to expect human-like behavior (Woodward et al., 2018). Children have also been found to attribute mental states such as intelligence and feelings to social robots (Kahn et al., 2012).

Seven and eight-year-olds tended to refer to electronic sources for the information provided by the machine. Several children just said the answers came “from Google”, but when asked to elaborate, had different explanations for what that meant. For example, one participant (age 8, male) said, “I think it just gets answers from the computer at Google. Yeah. I think it's all on Google and Google connects to the thingy that you ask questions on. And then, it gets your answers.” Another (age 8, male) said, “I think there's a big, big, big, big, many, many other Google stations where they keep programming. There's Google-making places and there's Google where they program the Googles to make it work, they test it and give it away. So, I

think they put the answers of every single question they know into every Google." One participant (age 8, female) reflected on how her mental model of Google has evolved:

> C: "When I was like, six and seven, I would think the way that Google was, was that there would be people that studied the universe and they would program it to say those things if you asked this, but once I thought about that a lot, I thought how would they know that we would ask all these different questions, and how would you know the answers to all of them?"
>
> R: "So how do you think it does work?"
>
> C: "Hmm. Maybe they didn't actually have to program it, but they could set it to copy and record so if you asked a question and then something gave it the answer, then it would know if you asked that question, this is the answer. I don't know."

Another word that was used to discuss the origin of the information given by the speaker was "websites". The Google Assistant usually sources its answers by starting with "according to the website (e.g.) en.wikipedia.org…" However, only one (age 8, female) of the 40 children discussed having heard that. She said she did not like that the conversational agent had sourced an answer about space from stargazer101.com instead of NASA. When I asked a participant (age 7, male), who said the information came from "websites", what that meant, he appeared not to be sure. The following exchange ensued:

> R: "Is that something you would want to know? Where that information came from?"
>
> C: "Yeah. […] Can I ask it right now?"
>
> R: "Yes. Try asking it a question and see if it tells you."
>
> C: "Okay. Hey Google, how do you get the information?"

Conversational Agent: “Sorry, I don’t understand.”

[…]

Parent: “And if you follow up [an answer] and say, “How do you know that?”

C: “Oh, okay. Hey Google. How do lights work?”

CA: “On the website en.wikibooks.org they say, ‘therefore is a result of this. The film then heats up, it starts glowing, converting electrical energy to light energy. This is because of the jewel effect, which means that resistances heat up when electrical current runs through them.’”

C: “Hey Google. How did you know how lights work?”

CA: “Sorry, I don't know how to help with that.”

R: “Did you notice that in the beginning, it did say something about ‘on the website, something, something’?

C: Wikibook.

R: So, what do you think that means? That it said that?

C: Um, I don't know.

The strategy this parent and child pair used here encountered the same problem with regards to context as that discussed in chapter 5. It is also possible that this child is not sure what a website is. Other parents in the study reported that children are not familiar with the idea of browsers and websites. For example, the mother of (another) participant (age 7, male), said that her child relies on the icons of recently visited websites on the search engine’s home page, or auto-complete features for searches and does not understand what a URL or website is. The father of another child (age 7, female), said “The URL goes over their head; they don't know

where answers come from. They haven't learned to use the iPad or phone to look for information." It seems that sourcing the information by giving a URL may be a good strategy for adults but may not be effective with this age group. Perhaps the name of an organization or an author would work better. However, websites like Wikipedia, where many of the answers to children's questions seem to come from, contain anonymous content, which would complicate this approach.

Likert Scale Ratings

To complement interview data and to understand further the children's assessment of using the technology, I asked them to rate the conversational agent along six dimensions using a six-point smiley-o-meter, similar to Druga and colleagues (2017). Table 8 below shows the average and median ratings given by children for each of the six dimensions. In aggregate, children gave the conversational agent a positive rating in all dimensions. In the cases where ratings were mostly positive (friendliness, trustworthiness and safety), the few children who gave negative ratings had specific reasons for doing so.

Regarding friendliness, all but three children rated the device friendly to some degree. Referring to the usage logs, one of the children (age 7, male) who gave it a -3 score on friendliness had made many attempts at social interactions with the conversational agent and complained about its unenthusiastic responses during the second interview. He said, "If it was alive, I would think it was a smart, unfriendly person."

Similarly, 90% of children rated the device as safe and trustworthy. One of the four participants who found it unsafe (male, age 5) was part of a family with three boys, the oldest of which had found a third-party action that made farting noises. This action was difficult to exit,

continuing to make the noises after users had made unrelated requests; it made the device difficult to control. Of the children who found the device unsafe, two (both females, age 8) mentioned having thought that it was being hacked when they started playing with a third-party action (trivia game) that used a male voice, in contrast to the default female voice of the Google Assistant.

Table 8: Children's ratings of the device from a 6-point Likert scale

Is the device...?	**Num. Children Agreed (N=40)**	**Mean Rating (SD)**	**Median**
Friendly	37	2.25 (1.32)	3
Smart	23	1.45 (1.91)	2
Trustworthy	36	1.94 (1.43)	2
Alive	22	0.20 (2.07)	1
Safe	36	2.15 (1.56)	3
Funny	30	0.95 (2.08)	2

After rating the device as smart or not smart, children were asked whether they thought the device was smarter than they are (scored 1), or not as smart as they are (-1). Although this was not offered as an option, some children rated it "as smart as me", which was entered as a score of 0. The average rating was M=0.5, SD=0.74, Median=1. A common explanation for thinking the device was smarter than they were was that it knew more facts. This echoes findings from Druga et al. (2017), who had children rate intelligent devices along similar dimensions after playing with them in a lab setting. Conversely, several of the children who said they themselves were smarter explained that, even though the device could look up information, they were better able to solve problems.

Children were also asked to explain why they thought the device was alive or not alive. Of the children who thought it was alive (55%), most said that it was alive because it could talk

and/or it sounded like a person. The children who said it was not alive said it was a machine, like a robot.

To test whether these concepts were related to age and sex, I calculated simple linear regressions on the scores for the 6 different dimensions, as well as the question about how intelligent the device was thought to be when compared to the child. There was no difference between male and female participants for any of the dimensions.

However, when using age to predict how friendly the participants thought the device was, I found that the friendliness rating decreased significantly with age ($F(1,39)=4.15$, $p=0.0486$, $R^2=0.10$), with about -0.33 rating point per year of age. Additionally, age predicted how smart children thought the device was ($F(1,39)=8.82$, $p=0.0051$, $R^2=0.19$), with about 0.66 rating point gained per year of age. Older children also were found slightly less likely to think the device was alive and more likely to think it was smarter than them, but those differences were not statistically significant.

Younger children in this group tended to explain how the smart speaker finds answers by describing actions a person would perform to find information, such as searching on a smartphone, while older children described electronic information sources, such as Google itself, but had different explanations for what that meant. When rating the smart speaker on personality dimensions, children overwhelmingly found it trustworthy, safe, and friendly -- though how friendly tended to decrease as children got older. Children were more divided when it came to thinking of the device as smart or alive, with a slight majority agreeing. Older children were more likely to think of it as smart.

Chapter 7: Discussion and Conclusion

The present study found that children did use the smart speakers to ask questions about the world around them. Questions accounted for almost half (47%) of all of focal children's interactions with the smart speakers. Children asked questions about a broad range of subjects. Questions about culture, which include topics such as entertainment, history and politics (22% of all questions) and science and technology (19%), which included animals, engineering and the human body, were most frequently represented. These categories contain children's questions about the world around them, suggesting that children think of the device as an information source for topics they were curious about. The two categories of culture and science and technology appeared more frequently than questions about the agent itself (18%) and questions coded as practical (18%), which include the weather.

My analysis found that the technology has evolved to transcribe children's questions correctly about 90% of the time, an improvement over the results of prior research (Kennedy et al., 2017), which found much lower rates. Besides age, a logical predictor of the rate at which the technology transcribes children's speech correctly, the only other factor that predicts the rate of correct transcriptions is the level of education of the participating parent (the average level of education of both parents was not a predictor of the correct transcription of children's questions). It is puzzling that parental education level is found to have a negative relationship with the rate of correct transcriptions. Parental education level is known as a positive predictor of language development (e.g., Hart & Risley, 1995). This may suggest that being transcribed correctly relates to something other than the usual measures of language development, like vocabulary size. A more detailed analysis of children's questions that were not transcribed correctly is

needed to understand the mechanisms that may be generating this relationship. Additionally, the level of education of the participating parent was a predictor of the rate of questions refused by the speaker (i.e., answered with "Sorry, I don't know how to help with that."). This may be because children of highly educated parents could be asking questions that are too detailed or too complex, which could generate refusals.

Even though the conversational agent transcribed 90% of children's questions correctly, it was still only able to address about half of their inquiries. One reason for this may be the lack of commercial focus on use cases that are important to children, such as asking information- and explanation-seeking questions. Although there is at least one smart speaker in the market that targets children specifically (the Amazon Echo Dot Kids Edition), it appears to prioritize branded content featuring media characters such as *Elmo* and *SpongeBob SquarePants* – the side of the box it comes in is filled with child-focused media company logos. Further, children may be significantly underrepresented in the training data used in these systems, widening a gap between children's use practices and behaviors and those of adults.

Children appeared to lose interest in the speaker quickly, with use declining steeply after the first few days. Interestingly, it was not asking questions specifically that declined, but overall usage. This suggests that within the first week of use, children might have exhausted experimentation, having perhaps made up their minds about what the smart speaker can do for them.

Nevertheless, the fact that children used the device to ask curiosity questions suggests that there is potential for this technology to play a role in children's self-directed learning. Self-directed learning, which is the ability to decide what to learn about, has been found to generate

better information retention both in adults (Markant et al., 2014) and in younger children (Partridge et al., 2015). Just as others have shown that children have difficulty reformulating questions to ask of a conversational agent (Yarosh et al., 2018), this study shows that children have challenges extracting information from a conversational agent's complex response. Inferential reasoning and theory of mind (e.g., Wellman & Liu, 2004) seem essential to the ability to reformulate questions and understand complex responses, making information seeking through conversational agents challenging for younger children who may not have yet developed these capacities.

The children in the families I studied did not have a smart speaker in their homes prior to participating in this research. Although most of them mentioned having interacted with Siri or the Google Assistant through tablets or parents' smartphones, this was the first time they interacted with a conversational agent without the benefit of a screen in which to view content. This being a relatively new technology to them, it is not surprising that more than half of the children think of the conversational agent as being alive and explain its behavior in human terms. Sherry Turkle (1984) found that children in the 1980s thought of computer games that offered variation in how they performed (i.e., imperfectly vs. perfectly) as being alive. Conversational agents offer much more variation than a handheld game of tic-tac-toe, and they have evolved to sound remarkably natural. This also raises user expectations about their abilities (Nass & Brave, 2005): the more like a person a machine sounds, the more like a person it will be expected to behave, and the fewer concessions users will be willing to make. The children in this study expected the agent to carry out a conversation as a person would, a finding similar to prior work on how children understand intelligent devices (e.g., Woodward et al., 2018). Prior work on

children's perceptions of computers (e.g., Mikropoulos, Misailidi, & Bonoti, 2003) had been done with older devices made up of a monitor, CPU, keyboard and mouse, and had found that children understood that computers were not alive beginning at around age 5. A "computer" that understands one's speech, replies in kind, and works with artificial intelligence appears to create a new ontological category, similar to what Khan and colleagues have found with older children and robots (Kahn et al., 2012).

In television's early days, it was not uncommon for viewers, especially children, to believe that there were people acting out a scene inside the box and that the people on the screen could hear them (e.g., Nikken & Peters, 1988). It might be harder to find children who still subscribe to those ideas in today's screen-saturated world, where children are able to flip a tablet upside down and see that the furniture in the scene does not fall to the ceiling. Likewise, children in a not-too-distant future might have fewer misconceptions about the nature of conversational agents.

Interestingly, younger children in the sample were less likely to find the speaker smart. This may be due to an evolving awareness of one's own knowledge (or lack thereof). In the early days of data collection, I piloted a task (removed after the first several participants), which consisted in asking parents and children to act out how they would find information to answer a couple of pre-scripted questions. I had prepared a list of questions to draw from, varying in complexity depending on the child's age, that were mostly science-related (e.g., "How do airplanes stay up in the air?", "Why do zebras have stripes?"). The task did not work because the younger children, aged 5 and 6 years, would immediately answer the question themselves with

their best guesses and consider the matter closed, losing interest in what the parent might find[4]. This illustrates what the data suggest: that younger children seem to rely more on what they know (or think they know) and are therefore less impressed with the conversational agent's ability to retrieve facts.

Children also were found to think of the speaker as less friendly as they got older -- although they still overwhelmingly rated it as friendly to some degree. This may be due to changing expectations of what it means to be friendly. Research has found that as children (ages 3 to 7) grow older, they tend to place a stronger focus on affective characteristics of friendship (Furman & Bierman, 1983), and affection is not a strong suit of the conversational agent.

Implications for Design

Several challenges appeared in the analysis of children asking questions of the Google Assistant that can help inform subsequent design work. Some of the key design insights based on this empirical investigation are detailed below:

Tailor answers to unique users. Some difficulties impacting children could be ameliorated if the system knew that a young child was speaking rather than an older sibling or adult. For example, responses that assume a humorous question, such as the existence of unicorns, could be answered differently to children than to adults. As I realized in the process of analyzing the data, the identification of children's voices is not yet reliable. Therefore, a stop-gap solution while voice ID improves could be, as one parent suggested, allowing users to self-identify as they initiate an inquiry. This would allow the service to offer simple answers to

[4] The task was also unnecessary because parents' own descriptions of how they find answers to children's questions were richer than the artificial situation the task created.

younger children, but be able to offer more options to older ones after giving an answer, such as "Would you like to hear more?" or "Would you like to hear an answer from a different source?"

Simplify and decompose answers to children's questions. The quality of the responses offered, as well as their length and complexity, seem to be the result of a combination of the information available online and the algorithm used to select what information to include when providing responses. For example, in the exchange described earlier between a participant and the conversational agent about squirrels liking pumpkins, the answer included a sequence that seemed to be part of a photo caption ("A squirrel stands on pumpkins carved into Halloween jack-o'-lanterns Oct. 25, 2004, in Washington."), which is not needed to answer the question, "Do squirrels like pumpkins?" Length and complexity are added in the response. The answer that was heard at the end would have been better on its own: "Here's the thing, squirrels don't even particularly like pumpkins."

Other responses required children to perform higher level reasoning and draw on other worldly knowledge (e.g., understanding that training to race against the fastest person might mean someone is second fastest). Continued development in the ability to identify information that answers a question succinctly and directly will benefit all users, not just children. Given the challenges of achieving this in real-time conversational AI systems, a crowd-powered approach may be a promising alternative (Huang, Chang and Bigham, 2018).

Understand context through prior questions and responses. Children often fail to provide the necessary context for initial questions, as Yarosh and colleagues (2018) also demonstrated. Beyond this, children (and adults) expect the agent to remember what was just said and be able to use that information as context for asking a follow-up question. That is, they

see interaction as building across a sequence of questions and responses, and they expect these prior interactions to contribute to conversational grounding (i.e., a mutual understanding between two parties regarding what they are talking about) (Brennan, 1998; Clark & Brennan, 1991). Such grounded, contextual understandings that take a sequence of interactions (e.g., multiple questions and responses) into account are not yet built into commercial conversational agents, as this is a difficult challenge for conversational AI systems in general (IBM, 2018). An interim solution could be to either ask the child to clarify, as prior work suggests (Yarosh et al., 2018), or to confirm. For example, in the case described above where a participant wanted further information about the gas that makes the refrigerator work, when subsequently asked "what's the name of the gas?" the agent could confirm if the user is indeed still talking about the same topic.

Adapt responses based on repetitive questioning. In my data, there were a few cases when children asked the same question over and over. This may be because they did not understand the answer or perhaps wanted more or different information on the same topic. If the conversational agent is able to identify and tailor responses to specific users, it could take repeat question-asking behavior into account for each individual as part of their own history. With this information, the system could then limit how many times it will repeat the same answer, after which it could either offer another option or suggest that the user ask the question in a different way. The open-ended fact-finding question-asking behavior observed in these data is distinct from other uses of conversational agents, such as asking practical questions (e.g., "What's the weather today?") that are likely to be repeated daily. Thus, having the system take into account the *kind* of question asked in addition to frequency seems like an important next step in supporting children's question-asking behavior.

Source curation and presentation. Another point worth noting is that, even though the Google Assistant mentions the source of many of its responses, reading out the URL before the answer to the question, only one of the children remembered this when asked how they thought the device worked. This is perhaps not surprising in a media world dominated by apps on touchscreens, where few 5-to-8-year-olds have a need to understand URLs. However, if children are able to seek information independently through a conversational agent before they are able to understand where the information comes from, they might benefit from curation managed by parents and educators. This way, the adults could be given access to source options to be drawn from when the children in their care are interacting with the technology. Alternatively, redesigning the presentation of source information could be a way to help children learn about where information comes from and source credibility at a young age.

Limitations and Future Work

One limitation of the present analysis is that the sample included predominantly middle- and upper-class families with highly educated parents. Therefore, I do not know how children of other socio-economic backgrounds and parental education levels might use this technology and what challenges they might encounter. For example, research has shown that children of lower SES react differently than wealthier, more educated counterparts when they receive dismissive or incomplete answers from adults (Kurkul & Corriveau, 2018; Tizard & Hughes, 1984), being less likely to insist or to come up with their own explanations.

A limitation of conducting research with a live technology is that I could not control Google's updates to its software, or the way it logged user's interactions with the smart speaker. Luckily, the only noticeable change – including audio files in the usage logs only for users who

had been voice-matched – happened at the very end of the data collection and only affected one participating family.

Finally, conversational agents are now available in many languages and countries. Prior work has shown differences between cultures in children's question-asking behavior (Gauvain et al., 2013; Henderson & Garcia, 1973). Even though this sample includes several children who spoke a language other than English in the home, they were still all English-speaking children growing up immersed in the culture of the United States. Future work should include children in other cultures and how their question-asking habits might impact their experiences trying to obtain answers from a conversational agent.

As conversational agents continue to expand their reach into homes, cars and additional electronic devices, children might grow more knowledgeable about how they work. The present analysis uncovers initial question-asking behaviors and practices in situ, however, a longitudinal study of how children incorporate this technology into their repertoire of information sources over a longer period could extend our understanding and inform future child-focused versions of conversational agents.

Another aspect of children's interactions with the conversational agent that could be better studied longitudinally is whether, and how, interacting with this technology impacts children's learning of language structures. There is a possibility that children might change how they express themselves in order to be better understood by the agent. We do not know whether this type of adaptation might transfer to other forms of expression such as interactions with people, or writing.

Part of the reason behind the complexity of the answers that children hear from the conversational agent may be related to the fact that the answers come from texts found on the web that were not written to be read aloud. An analysis of the answers received and future experiments testing the comprehension of multiple versions of the same information (focusing on the style one might use for an audio program instead of what is written to be read silently), may be an important step toward improving accessibility of responses for all users, but especially for children.

The data collected for this project includes many examples of children's interactions with the smart speakers that were not coded as questions, most interestingly their attempts at social interaction, including declarations of love and friendship ("Hey Google, I love you!"), greetings ("Hey Google, I'm home!"), and even attempts to show the conversational agent one's achievements in music instruction ("Hey Google, listen to this!" followed by recorder playing). A detailed analysis of these interactions could further our understanding of how children think of this technology and attempt to relate to it.

Conclusion

The work undertaken in this dissertation revealed that children do use a smart speaker to ask fact-finding questions when one is available to them. However, even though the technology is able to understand their words and transcribe them correctly the majority of the time, children only hear answers that directly address their questions in 50% of cases. Even when children do hear direct answers, they face challenges interpreting the information. The fact that children use conversational agents to find information suggests that it may be worth continuing to investigate ways to make this technology work better for them.

This dissertation helps to advance our understanding of how early elementary-aged children use conversational agents to find information. As opportunities for voice interaction continue to expand rapidly, the importance of studying how children may learn from conversations with artificial intelligent-based systems will continue to grow as well.

References

Abramovich, G. (2018). Study Finds Consumers Are Embracing Voice Services. Retrieved September 2018 from https://www.cmo.com/features/articles/2018/9/7/adobe-2018-consumer-voice-survey.html#gs.6DCN6JNj

Beneteau, E., Richards, O., Zhang, M., Kientz, J. A., Yip, J. C., & Hiniker, A. (2019). Communication Breakdowns Between Families and Alexa. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems-CHI* (Vol. 19, pp. 1-14).

Bilal, D. (2002). Children's use of the Yahooligans! Web search engine. III. Cognitive and physical behaviors on fully self-generated search tasks. *Journal of the American Society for information science and technology* 53, 13: 1170-1183.

Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analysis and code development.* Sage.

Braun, V. & Clarke, V. (2006). Using Thematic Analysis in Psychology. In *Qualitative Research in Psychology*, 3, 2: 77-101.

Brave, S., & Nass, C. (2003). Emotion in human-computer interaction. The human-computer interaction handbook: fundamentals, evolving technologies and emerging applications, 81-96.

Brennan, S. E. (1998). The grounding problem in conversations with and through computers. In *Social and cognitive approaches to interpersonal communication* .201-225.

Callanan, M. A. & Oakes, L. M. (1992). Preschoolers' questions and parents' explanations: Causal thinking in everyday activity. *Cognitive Development*, 7, 2: 213-233.

Chall, J. S., Roswell, F. G., Curtis, M. E., & Strucker, J. (2003). *Quick adult reading inventory.* Continental Press, Inc.

Cheng, Y., Yen, K., Chen, Y., Chen, S., & Hiniker, A. (2018). Why Doesn't It Work? Voice-Driven Interfaces and Young Children's Communication Repair Strategies. *Proceedings of the 17th ACM Conference on Interaction Design and Children*, pp. 337-348.

Chiasson, S., & Gutwin, C. (2005, April). Testing the media equation with children. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 829-838). ACM.

Chouinard, M. M., Harris, P. L., & Maratsos, M.P. (2007). Children's questions: A mechanism for cognitive development. *Monographs of the Society for Research in Child Development*: i-129.

Clark, H.H. & Brennan, S.E. (1991). Grounding in communication. *Perspectives on socially shared cognition* 13, 1991: 127-149.

Common Sense Media & SurveyMonkey. (2019). What's That You Say? Smart Speakers and Voice Assistants Toplines. Retrieved from https://www.commonsensemedia.org/sites/default/files/uploads/pdfs/2019_cs-sm_smartspeakers-toplines_final-release.pdf March, 2019

Corriveau, K., & Harris, P. L. (2009). Choosing your informant: Weighing familiarity and recent accuracy. *Developmental Science*, 12(3), 426-437.

Danovitch, J. H. & Alzahabi, R. (2013). Children show selective trust in technological informants. *Journal of Cognition and Development* 14, 3: 499-513.

Diethelm, I., Wilken, H., & Zumbrägel, S. (2012). An investigation of secondary school students' conceptions on how the internet works. In *Proceedings of the 12th Koli Calling International Conference on Computing Education Research* (pp. 67-73). ACM.

Druga, S., Williams, R., Breazeal, C., & Resnick, M. (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*, pp. 595-600.

Druin, A., Foss, E., Hatley, L., Golub, E., Guha, M. L., Fails, J., & Hutchinson, H. (2009). How children search the internet with keyword interfaces. In *Proceedings of the 8th International conference on interaction design and children*, pp. 89-96.

Druin, A., Foss, E., Hutchinson, H., Golub, E., & Hatley, L. (2010). Children's roles using keyword search interfaces at home. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'10)*, pp. 413-422.

Foss, E. & Druin, A. (2014). Children's Internet Search: Using Roles to Understand Children's Search Behavior. *Synthesis Lectures on information concepts, retrieval, and services*, 6, 2: 1-106.

Furman, W., & Bierman, K. L. (1983). Developmental changes in young children's conceptions of friendship. Child Development, 54(3), 549.

Gauvain, M., Munroe, R. L., & Heidi Beebe. (2013). Children's questions in cross-cultural perspective: A four-culture study. *Journal of Cross-Cultural Psychology* 44, 7: 1148-1165.

Harris, P. (2012). *Trusting What You're Told: How Children Learn From Others*. Harvard University Press.

Henderson, R. W. & Garcia, A. B. (1973). The effects of parent training program on the question-asking behavior of Mexican-American children. *American Educational Research Journal* 10, 3: 193-201.

Huang, T. K., Chang, J. C., & Bigham, J. P. (2018). Evorus: A Crowd-powered Conversational Assistant Built to Automate Itself Over Time. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, p. 295. ACM.

IBM. (2018). How conversation (with context) will usher in the AI future. Retrieved December 2018 from https://www.ibm.com/watson/advantage-reports/future-of-artificial-intelligence/ai-conversation.html

Kabali, H. K., Irigoyen, M. M., Nunez-Davis, R., Budacki, J. G., Mohanty, S. H., Leister, K. P., & Bonner, R. L. (2015). Exposure and use of mobile media devices by young children. Pediatrics, 136(6), 1044-1050.

Kahn Jr, P. H., Gary, H. E., & Shen, S. (2013). Children's Social Relationships With Current and Near□Future Robots. *Child Development Perspectives*, 7(1), 32-37.

Kahn Jr, P. H., Kanda, T., Ishiguro, H., Freier, N. G., Severson, R. L., Gill, B. T., … Shen, S. (2012). "Robovie, you'll have to go into the closet now": Children's social and moral relationships with a humanoid robot. *Developmental psychology* 48, 2: 303.

Kammerer, Y., & Bohnacker, M. (2012, June). Children's web search with Google: the effectiveness of natural language queries. In *Proceedings of the 11th international Conference on Interaction Design and Children* (pp. 184-187). ACM.

Kennedy, J., Lemaignan, S., Montassier, C., Lavalade, P., Irfan, B., Papadopoulos, F. … Belpaeme, T. (2017). Child speech recognition in human-robot interaction: evaluations and recommendations. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 82-90.

Kuhlthau, C. C. (1991). Inside the search process: Information seeking from the user's perspective. *Journal of the American Society for Information Science*, 42(5), 361-371.

Kurkul, K. E. & Corriveau, K. H. (2018). Question, explanation, follow-up: A mechanism for learning from others. *Child Development* 89, 1: 280-294.

Lagattuta, K. H., Kramer, H. J., Kennedy, K., Hjortsvang, K., Goldfarb, D., & Tashjian, S. (2015). Beyond Sally's missing marble: Further development in children's understanding of mind and emotion in middle childhood. In *Advances in Child Development and Behavior* (Vol. 48, pp. 185-217). JAI.

Lovato, S., & Piper, A. M. (2015). “Siri, is this you?”: Understanding young children's interactions with voice input systems. *Proceedings of the 14th International Conference on Interaction Design and Children*, pp. 335-338.

MacWhinney, B. (2000). The CHILDES Project: Tools for analyzing talk. Third Edition. Mahwah, NJ: Lawrence Erlbaum Associates.

Maier, A., Hönig, F., Steidl, S., Nöth, E., Horndasch, S., Sauerhöfer, E. … Moll, G. (2011). An automatic version of a reading disorder test. *ACM Transactions on Speech and Language Processing (TSLP)* 7, 4: 17

Marchionini, G. (1989). Information-seeking strategies of novices using a full-text electronic encyclopedia. *Journal of the American Society for Information Science* 40, 1: 54-66.

Markant, D., DuBrow, S., Davachi, L. & Gureckis, T. M. (2014). Deconstructing the effect of self-directed study on episodic memory. *Memory & Cognition* 42, 8: 1211-1224.

McGregor, J. (2017). Amazon Echo vs. Google Home Black Friday: Which is Best? Retrieved January 2018 from https://www.forbes.com/sites/jaymcgregor/2017/11/15/black-friday-amazon-alexa-vs-google-home/#415774512cb7

McKenney, S. & Voogt, J. (2010). Technology and young children: How 4–7-year-olds perceive their own use of computers. *Computers in Human Behavior* 26, 4: 656-664.

Mikropoulos, T. A., Misailidi, P. & Bonoti, F. (2003). Attributing human properties to computer artifacts: Developmental changes in children's understanding of the animate-inanimate distinction. *Psychology*. 10, 1: 53-64.

Mills, C. M., Legare, C. H., Grant, M. G. & Landrum, A. R. (2011). Determining who to question, what to ask, and how much information to ask for: The development of inquiry in young children. *Journal of Experimental Child Psychology.* 110: 539-560.

Nass, C. I. & Brave, S. (2005). *Wired for speech: How voice activates and advances the human-computer relationship.* MIT Press.

Nielsen. (2018). (Smart) Speaking My Language: Despite Their Vast Capabilities, Smart Speakers Are All About The Music. Retrieved December 2018 from https://www.nielsen.com/us/en/insights/news/2018/smart-speaking-my-language-despite-their-vast-capabilities-smart-speakers-all-about-the-music.html

Nikken, P. & Peeters, A. L. (1988). Children's perceptions of television reality. *Journal of Broadcasting & Electronic Media* 32, 4: 441-452.

Norman, D. A. (2014). Some observations on mental models. In *Mental Models* (pp. 15-22). Psychology Press.

Papastergiou, M. (2005). Students' mental models of the Internet and their didactical exploitation in informatics education. In *Education and Information Technologies*, 10(4), 341-360.

Papert, S. (1993). The children's machine: Rethinking school in the age of the computer. BasicBooks, 10 East 53rd St., New York, NY 10022-5299.

Partridge, E., McGovern, M. G., Yung, A., & Kidd, C. (2015). Young children's self-directed information gathering on touchscreens. In *Proceedings of the 37th Annual Conference of the Cognitive Science Society, Austin, TX. Cognitive Science Society*.

Piaget, J. (1929). *The child's conception of the world* (J. Tomlinson & A. Tomlinson, Trans.). London: Kegan Paul, Trench, Trubner & Co.

Piotrowski, J. T., Litman, J. A., & Valkenburg, P. (2014). Measuring epistemic curiosity in young children. *Infant and Child Development*, 23(5), 542-553.

Purington, A., Taft, J. G., Sannon, S., Bazarova, N. N. & Hardman Taylor, S. (2017). Alexa is my new BFF: Social Roles, User Satisfaction, and Personification of the Amazon Echo.

Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '17, 2853–2859.

Pyae, A., & Joelsson. T. N. (2018). Investigating the usability and user experiences of voice user interface: a case of Google home smart speaker. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*, pp. 127-131

RBC Capital Markets. (2019). Hey Alexa, What's Next for Voice Tech? Five Insights into a Voice-First World. Retrieved from https://www.rbccm.com/en/insights/story.page?dcr=templatedata/article/insights/data/2019/04/hey_alexa_whats_next_for_voice_tech_five_insights in April, 2019.

Reeves, B., & Nass, C. I. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. Cambridge university press.

Rideout, V. (2017). The Common Sense census: Media use by kids age zero to eight. Robb, M., (ed.) *Common Sense Media*

Rothbart, M. K., Ahadi, S. A., Hershey, K. L., & Fisher, P. (2001). Investigations of temperament at three to seven years: The Children's Behavior Questionnaire. Child development, 72(5), 1394-1408.

Rücker, M. T. & Pinkwart, N. (2016). Review and discussion of children's conceptions of computers. *Journal of Science Education and Technology* 25, 2: 274-283.

Sciuto, A., Saini, A., Forlizzi, J. & Hong, J. I.. (2018). Hey Alexa, What's Up?: A Mixed-Methods Studies of In-Home Conversational Agent Usage. In *Proceedings of the 2018 on Designing Interactive Systems Conference 2018*, pp. 857-868.

Shulevitz, J. (2014). Siri, You're Messing Up a Generation of Children. *The New Republic.* Retrieved from https://newrepublic.com/article/117242/siris-psychological-effects-children in March, 2019.

Solomon, P. (1993). Children's Information Retrieval Behavior: A Case Analysis of an OPAC. *Journal of the American Society for information Science* 44 (5): 245-264.

Teale, W. H. & Sulzby, E. (1986). Emergent Literacy: Writing and Reading. *Writing Research: Multidisciplinary Inquiries into the Nature of Writing Series.* Ablex Publishing Corporation.

Tizard, B. & Hughes, M. (1984). *Young Children Learning*. John Wiley and Sons.

Turkle, S. (1984). *The Second Self: Computers and the Human Spirit.* MIT Press.

Van Duuren, M., Dossett, B., & Robinson, D. (1998). Gauging Children's Understanding of Artificially Intelligent Objects: A Presentation of "Counterfactuals." *International Journal of Behavioral Development* 22(4), 871-889.

Vygotsky, L. S. & Cole, M. (1978). *Mind in society: The development of higher mental process*. Harvard University Press.

Ward, W., Cole, R., Bolaños, D., Buchenroth-Martin, C., Svirsky, E., Van Vuuren, S., … Becker, L. (2011). My science tutor: A conversational multimedia virtual tutor for

elementary school science. *ACM Transactions on Speech and Language Processing (TSLP)*7, 4 :18.

Weizenbaum, J. (1966). ELIZA--A computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36-45.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta□analysis of theory□of□mind development: The truth about false belief. *Child Development*, 72(3), 655-684.

Wellman, H. M. & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development* 75, 2: 523-541.

Woodward, J., McFadden, Z., Shiver, N., Ben-Hayon, A., Yip, J. C, & Anthony, L. (2018). Using Co-Design to Examine How Children Conceptualize Intelligent Interfaces. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, p. 575.

Yarosh, S., Thompson, S., Watson, K., Chase, A., Senthilkumar, A., Yuan, Y., & Brush, A. J. (2018). Children asking questions: speech interface reformulations and personification preferences. In *Proceedings of the 17th ACM Conference on Interaction Design and Children*, pp. 300-312

Appendix A: Parent Questionnaire

Children's Questions and Voice Interfaces

Parent Survey

Part 1: Child Demographics

1. What is your child's first name or nickname? _____

2. What is your child's sex?
   ☐ Male ☐ Female

3. What is your child's birth date? ____._____._____

4. What is your child's race? Please mark all that apply.
   ☐ African American
   ☐ American Indian and Alaska Native
   ☐ Asian
   ☐ Native Hawaiian or Other Pacific Islander
   ☐ White
   ☐ Some Other Race (please specify)

5. Is your child Hispanic or Latino/a?
   ☐ Yes ☐ No

6. What grade is your child in?
   ☐ Not in school
   ☐ Preschool/Pre-K
   ☐ Kindergarten
   ☐ 1st grade
   ☐ 2nd grade
   ☐ 3rd grade

7. What kind of school, if any, does your child attend?
   ☐ Public school
   ☐ Private school
   ☐ Preschool or day care center
   ☐ Home-based day care
   ☐ Homeschool
   ☐ My child does not attend school

Part 2: Technology in the Home (adapted from Common Sense Media 2017 survey)

8. Which of the following technologies do you have in your household? Please check all that apply.
   ☐ Television
   ☐ An antenna (to watch over-the-air TV)
   ☐ Cable or satellite TV

- ☐ DVR (a device to record shows from TV to watch later)
- ☐ A device that allows you to watch online video on your TV (such as Roku, Apple TV, Chromecast, etc.)
- ☐ Video subscription service (such as Netflix, Amazon Prime, Hulu, etc.)
- ☐ DVD player
- ☐ Computer (desktop or laptop)
- ☐ High-speed Internet access
- ☐ Video game console (such as X-box, PlayStation, Wii)
- ☐ Virtual Reality headset (such as Oculus Rift, Google Cardboard, PlayStation VR)
- ☐ A handheld video game (such as Game Boy, PSP, Nintendo DS)
- ☐ Tablet (such as iPad, Samsung Galaxy Tab, Microsoft Surface or Kindle Fire)
- ☐ iPod Touch or similar
- ☐ E-reader (such as a Nook or Kindle)
- ☐ A smart speaker (such as the Amazon Echo or Google Home, including the Dot or Mini versions)

9. What type of cell phone, if any, do you (the parent or guardian) have?

- ☐ A smartphone
- ☐ A cell phone just for texting and talking
- ☐ I don't have a cell phone

10. Which of the following items does your child have? Please check all that apply.

- ☐ [His/her] own cell phone
- ☐ [His/her] own tablet (iPad, Kindle Fire or similar)
- ☐ [His/her] own iPod Touch or similar device
- ☐ [His/her] own educational game device (such as a Leapster/LeapPad, LeapFrog Epic, Playtime Pad, or V-Tech device (V-Smile, Mobigo, or Innotab)
- ☐ [His/her] own handheld video game player like a Game Boy, PSP, or Nintendo DS
- ☐ [His/her] own internet-connected toys, such as Hello Barbie, Fisher-Price Smart Toy Bear, or CogniToys Dino

11. If you check the first box in question 10 (your child has his/her own cell phone), is it a smart phone?

- ☐ Yes
- ☐ No

12. How often is a smartphone, tablet or laptop in your child's bedroom overnight?

- ☐ Every night
- ☐ Most nights
- ☐ Some nights
- ☐ Hardly ever
- ☐ Never

Part 3: Parent and Child Technology Use

13. How often does your child use a smartphone, tablet or computer to…?

	Several times a day	Once a day	Several times a week	Once a week	Less than once a week	Has never done
Watch TV or movies						
Watch videos						
Play games						
Listen to music						
Read books						
Take photos or videos						
Look at photos or videos						
Draw or write						
Look something up/search						
Talk to Siri, Google Assistant or similar						

14. If your child uses a smartphone, tablet or computer to look up information, please tell us as much detail as you remember about the last time it happened (i.e., the context, what were they looking for, what they found, how much help they needed to search and interpret results).

15. If your child uses a smartphone, tablet or computer to speak to a virtual assistant such as Siri, the Google Assistant or Microsoft's Cortana, please tells us about the last time it happened with as much detail as you remember.

16. If your child watches videos online (on any device), how often does she or he watch the following types of videos?

	Often	Sometimes	Hardly ever	Never	I don't know.
How-to videos (e.g., how to draw, cook, dance, make crafts, make things with Legos or Play-Doh)					
"Unboxing" videos (i.e., video if someone opening a new toy), or product demonstrations (showing off toys, electronics, cosmetics, etc.)					
Music videos					
Video gaming/gameplay videos					
Celebrity/behind-the-scene videos					
Challenges/stunts/tricks videos					
Animal videos					
Videos featuring family and friends, posted by people or organizations your family knows (e.g., school events)					
Learning videos (alphabet, numbers, math, homework help, etc.)					
Television shows					
Other (please explain below)					

Other kinds of videos watched:

17. Thinking of the kinds of videos your child watches the most, what is **the main** way he or she found those videos?
    - ☐ Through friends
    - ☐ Through siblings or cousins
    - ☐ You or another adult selected them
    - ☐ Your child found them him/herself
    - ☐ I don't know
    - ☐ My child doesn't watch videos online
    - ☐ Other (please specify below)

18. Has your child ever found something online that you wish they had not? If yes, please tell us what happened.

19. How much help does your child need when using technology, most of the time? Please leave blank if your child does not use a device.

    He or she can use a…

	tablet	smartphone	computer	game console	other (please specify)
Independently (someone else might turn on or unlock the device)					
Mostly independently (with some help finding an activity or getting unstuck)					
With light assistance (someone else helps but the child does most of the actions)					
With heavy assistance (the child does less than half of the actions)					

20. How independently does your child decide what to do on each device, most of the time? Please leave blank if your child does not use a device.

	tablet	smartphone	computer	game console	other (please specify)
Independently The child is allowed to choose freely.					
Mostly independently There are some guidelines and off-limits features.					
Under some supervision There are some pre-approved features; anything else must be discussed.					
Under strict supervision					

The child asks for permission to use each feature.					

21. Thinking of tablets and smartphones, where does your child **typically** use them? Check all that apply

	Tablet	Smartphone
At home		
in the living room		
in the kitchen/dining area		
in the child's own bedroom		
other home location (please specify)		
In the car		
In restaurants		
In waiting rooms (such as at the doctor's)		
At the grocery store		
At the library		
Other (please specify)		

22. What are your child's favorite activities in each device?

Computer:

Tablet:

Smartphone:

23. Does your child use apps or features that are NOT specifically designed for children? Please check all that apply:

- ☐ YouTube (not YouTube Kids)
- ☐ Netflix, Amazon Prime, Hulu or equivalent
- ☐ Email or text messaging
- ☐ Camera (taking photos or video)
- ☐ Photo or video editing features

- ☐ Word processing (MS Word, Google Docs, etc.)
- ☐ App store, Google Play or equivalent
- ☐ Maps
- ☐ Social media (Instagram, Facebook, etc.)
- ☐ Other (please describe below)

Other features NOT specifically designed for children that your child uses:

24. Do you ever use any type of software or a device to monitor your child's access to the internet, such as NetNanny or Circle with Disney?
    - ☐ Yes
    - ☐ No
    - ☐ I don't know

25. Please tell us about how you use technology. When you are **NOT at work**, how often do you…

	Send and receive text messages (SMS or apps such as WhatsApp or Facebook messenger)	Send and receive emails	Make and receive phone calls (including FaceTime and Skype)	Look up information	Use a social networking app or website (Facebook, Instagram, Twitter, etc.)
Several times per hour or more often					
Once per hour					
Several times a day					
Once a day					
Several times a week					

Once a week or less often					
Never					

26. Do you use spoken language to interact with a device, such as talking to Siri or Google Assistant? If yes, tell us about the most frequent cases, including where and when. For example, "I ask Siri to text my spouse that I am running late when I am driving."

27. How often is your child with you when you use technology?

☐ Always
☐ Most of the time
☐ About half the time
☐ Rarely
☐ Never

Part 4: Child Personality

For questions 28 to 53, please use the following scale:

Circle #	If the statement is:
1	extremely untrue of your child
2	quite untrue of your child
3	slightly untrue of your child
4	neither true or false of your child
5	slightly true of your child
6	quite true of your child
7	extremely true of your child

If you cannot answer one of the items because you have never seen the child in that situation, for example, if the statement is about the child's reaction to your singing and you have never sung to your child, then circle NA (not applicable).

My child:

28. Sometimes prefers to watch rather than join other children playing.
1 2 3 4 5 6 7 NA

29. Usually rushes into an activity without thinking about it.
1 2 3 4 5 6 7 NA

30. Is comfortable in situations where s/he will be meeting others.
1 2 3 4 5 6 7 NA

31. Seems to be at ease with almost any person.
    1 2 3 4 5 6 7 NA

32. Sometimes interrupts others when they are speaking.
    1 2 3 4 5 6 7 NA

33. Decides what s/he wants very quickly and goes after it.
    1 2 3 4 5 6 7 NA

34. Gets embarrassed when strangers pay a lot of attention to her/him.
    1 2 3 4 5 6 7 NA

35. Acts very friendly and outgoing with new children.

    1 2 3 4 5 6 7 NA

Circle #	If the statement is:
1	extremely untrue of your child
2	quite untrue of your child
3	slightly untrue of your child
4	neither true or false of your child
5	slightly true of your child
6	quite true of your child
7	extremely true of your child

36. Often rushes into new situations.

    1 2 3 4 5 6 7 NA

37. Joins others quickly and comfortably, even when they are strangers.

    1 2 3 4 5 6 7 NA

38. Takes a long time in approaching new situations.

    1 2 3 4 5 6 7 NA

39. Usually stops and thinks things over before deciding to do something.

    1 2 3 4 5 6 7 NA

40. Is sometimes shy even around people s/he has known a long time.

1 2 3 4 5 6 7 NA

41. Is slow and unhurried in deciding what to do next.

1 2 3 4 5 6 7 NA

42. Sometimes seems nervous when talking to adults s/he has just met.

1 2 3 4 5 6 7 NA

43. Tends to say the first thing that comes to mind, without stopping to think about it.

1 2 3 4 5 6 7 NA

44. Acts shy around new people.

1 2 3 4 5 6 7 NA

45. When eager to go outside, sometimes rushes out without putting on the right clothes.

1 2 3 4 5 6 7 NA

Circle #	If the statement is:
1	extremely untrue of your child
2	quite untrue of your child
3	slightly untrue of your child
4	neither true or false of your child
5	slightly true of your child
6	quite true of your child
7	extremely true of your child

46. Is comfortable asking other children to play.

1 2 3 4 5 6 7 NA

47. Approaches slowly places where s/he might hurt her/himself.

1 2 3 4 5 6 7 NA

48. Talks easily to new people.

1 2 3 4 5 6 7 NA

49. When s/he sees a toy or game s/he wants, is eager to have it right then.

1 2 3 4 5 6 7 NA

50. Is among the last children to try out a new activity.

1 2 3 4 5 6 7 NA

51. Sometimes turns away shyly from new acquaintances.

1 2 3 4 5 6 7 NA

52. Is "slow to warm up" to others.

1 2 3 4 5 6 7 NA

53. Seems completely at ease with almost any group.

1 2 3 4 5 6 7 NA

54. How often does your child present the characteristics or behaviors below?

	Always or almost always	Often	Sometimes	Almost Never or Never
My child has fun learning about new topics or subjects.				
My child is attracted to new things in his/her environment.				
My child enjoys talking about topics that are new to him/her.				
My child shows visible enjoyment when discovering something new.				
When my child is learning about something new, she/he asks many questions about it.				
When presented with a tough problem, my child focuses all his/her attention on how to solve it.				
My child devotes considerable effort trying to figure out things that are confusing or unclear.				

My child is bothered when she/he doesn't understand something, and tries hard to make sense of it.				
My child will work for a long time to solve a problem because he/she wants to know the answer.				
My child carefully examines things by turning them around and looking at them from all sides.				

Part 5: Parent Demographics (If you are the only parent, please skip the "other parent" questions.)

55. What is your age? _____
56. What is your gender? _____
57. What is the age of your child's other parent? _____
58. What is the gender of your child's other parent? _____

59. Marital status:

☐ Single/Never been married
☐ In a domestic partnership
☐ Married
☐ Separated
☐ Divorced
☐ Widowed

60. What is your education level? Please mark your highest degree received.

☐ Less than high school
☐ High school diploma or equivalent (like GED)
☐ Some college, or Associate's degree
☐ Bachelor's degree
☐ Master's degree
☐ Doctoral degree (JD, MD, PhD, etc.)

61. What is the education level of your child's other parent?

☐ Less than high school
☐ High school diploma or equivalent (like GED)
☐ Some college, or Associate's degree
☐ Bachelor's degree
☐ Master's degree
☐ Doctoral degree (JD, MD, PhD, etc)

62. What is your occupation? _____

63. What is the occupation of your child's other parent? _____

64. Please tell us about other children living in your household 50% of the time or more (check here if no other children: [ ]).

First Name/Nickname	Male/Female	Age

65. Does your family speak a language other than English at home?

☐ Yes ☐ No

If yes, what language(s)? _____

If yes, what percent of the time? Please make an x along the line for each language other than English. If more than one, please label each x.

0%----------50%----------100%

66. What is your annual family income before taxes?

☐ Less than $25,000
☐ $25,000 to $49,000
☐ $50,000 to $74,999
☐ $75,000 to $99,999
☐ $100,000 to $124,999
☐ $125,000 to $149,999
☐ $150,000 to $174,999
☐ $175,000 to $199,999
☐ more than $200,00

67. What is your race?

☐ African American
☐ American Indian and Alaska Native
☐ Asian
☐ Native Hawaiian or Other Pacific Islander
☐ White
☐ Some Other Race (please specify)

68. Are you Hispanic or Latino?

☐ Yes ☐ No

69. What is your child's other parent's race?

☐ African American
☐ American Indian and Alaska Native
☐ Asian
☐ Native Hawaiian or Other Pacific Islander
☐ White
☐ Some Other Race (please specify)

70. Is he or she Hispanic or Latino?

☐ Yes
☐ No

Appendix B: Interview Guide

Rapport Building

- What are some things that you like to do in your free time?
    - When did you start doing that?
    - How did you learn about/how to do it?

- I hear that many children are curious. Do you think you are a curious person?
- What are you curious about? Are you curious about [topic related to free time activity]?
- When you have a question, how do you usually find the answer?
- Can you remember the last time you had a question? What was it about?
- How did you find the answer? (if they answer that they asked an adult, ask how they think the adult found the answer).

Devices

- Do you ever use a tablet, computer or smartphone?
    - Which one do you use the most?

- What is it for? What do you use it for?
- How about in school? Do you use any tablet or computer at school? What for?
- How does new stuff appear in the device?
- Can you get videos or new games anywhere you go? Like, can you use it in the car? What do you need for it to work?
- Have you heard of the Internet? If a kid your age who had never heard about it asked you what it was, what would you say?

Finding Information Online

- Do you ever try to find answers to your questions through these things?
- How do you do it?
- Can you tell me about the last time you used this device to find information about something you wanted to know? What happened?
- Where do you think the answer came from?
- When you search on Google, what appears on the screen? Who do you think wrote that? If videos, who put that video there?
- How do you decide which link/video to click on?
- Does your device have a voice in it that you can talk to, like Siri or Alexa? Do you ever do that?
- What is that like? What do you say to it? Can you tell me about the last time you talked to it?
- Do you ever try to get answers by talking to [Siri/Alexa]?

- How does it know the answer?

Questions for the parent:

- We hear that children can ask a lot of questions! Is that true of your child? Can you tell us a little bit about her/his questions from your perspective? Do you remember some recent ones?
- What if you don't immediately know the answer? What would you say is the most common way in which you help answer a question?
- Have you ever learned something new or interesting because of a question your child asked?
- Does [child] give you suggestions for how to find the answer? Does he or she ask you to look it up, or "google it"? Do you remember at what age they started doing that?
- Have you noticed any change in the type of questions they ask over the years? If yes, how have the questions changed?

Appendix C: Descriptive Subject Codes

Wider Subject Area	Descriptive Codes	Definition	Examples
agent	agent	questions about the GA	How old are you? What kind of ice cream do you like? What's your favorite color?
agent	device	questions about device's functionality	Are you on? How much time left on my timer?
agent	other agents	questions about Siri/Alexa/Cortana	do you know Siri?
cultural	art	questions about visual arts and art museums	where's the monalisa painting?
cultural	books	questions about who wrote what, books in general	who's the most famous children's author? How many books are there in the world?
cultural	business	questions about companies	what is the biggest company in the world?
cultural	civics	questions about national symbols or how the government works	how wide is the American flag
cultural	fantasy	questions about dragons, fairies, unicorns, etc	are unicorns real?
cultural	food	info about foods/cuisines	what different kinds of donuts are there?
cultural	geography	political and administrative facts about countries and states	what's the capital of jamaica?
cultural	history	questions about historic events	When was the medieval times? What is the World War?
cultural	human behavior	mostly reasons behind behaviors	why do people collect pokemon cards?
cultural	landmarks	facts about famous places, buildings, monuments, etc.	how tall is the sears tower?
cultural	media and technology	questions about the behind-the-scenes of movies, TV programs, apps and software in general.	how are movies made?
cultural	music	questions about musical instruments, notation, composers	What does a bassoon sound like?
cultural	names	questions about first names or family names	is Kennedy a boy name?
cultural	opinion	questions with subjective answers, not fact-based	what is the best store? What are some fun games?

cultural	philosophy	philosophical questions, not God-related.	what is ying yang?
cultural	pop culture	questions about bands, movies, TV shows, cartoons and their characters	do you know My Little Pony?
cultural	public figures	questions about famous people, politicians, movie stars, etc.	how tall is Donald Trump? who is Donald Trump? where does Sophie Kinsella live?
cultural	religion	questions about God, the Bible, the Torah, other sacred texts	when did God die? who are Adam and Eve?
cultural	social norms	age limits, other social customs	Can a 13 year old write a will? How old do you have to be to get your ears pierced?
cultural	sports	questions about teams, rankings, historical sports data	who won the master's tournament? who is batting for the Cubs? what club dos Cristiano Ronaldo play for?
joke	joke	joke questions	Why did the chicken cross the road? What does the fox say?
language	translation	requests for translations	can you say Google is stupid in Indian?
language	words	word meanings, origins, spellings	what is the name pizza? what do you call the thing that birds live in? how do you spell google? what does ethiopia mean?
math	math	calculations with simple numbers (not time, distance, ingredients)	how much is 1 + 1?
personal info	personal info	anything the user teaches google (name, etc)	what is my name? what is my dog's name?
personal info	research team	questions about the research project or people	do you know the name of the person who brought you here?
personal info	user's context	question about the user's surroundings	what the heck is that? where is my mom?
practical	advice	advice	what movie should I watch? how do I get my kid to eat vegetables? What should I bring to the beach today?
practical	calendar	questions about holidays and dates	what day is today? how many days until Christmas? what's the shortest month? when is Columbus Day?
practical	distances/length	distance between two points	how far is New York from Chicago?
practical	food	recipes, nutrition info, kitchen measurements, substitutions	how do I make cookies? how many calories in a potato chip? how many teaspoons in a sugar cube?
practical	how-to	requests for instructions or directions	how do you make a bird-feeder? how do I play candyland? How do you get to Savannah?

practical	local	local information	Where's the nearest McDonalds?
practical	money	questions about the cost of things or anything related to money	how much does tape cost?
practical	time	questions related to time and durations	what time is it? how long does it take to…
practical	weather	forecast, local or not	what's the weather?
science	animals	facts about animals in general, including questions about pets (dogs/cats/goldfish/etc)	how many bones do giraffes have? can dogs eat chocolate?
science	chemistry	questions about the composition or process through which different compounds are made.	How is acid made?
science	engineering	questions about how things are built	how much concrete does it take to fill an entire house?
science	everyday objects	questions about the workings and materials of every day things.	how much ink does a fountain pen hold? how are pillows made? Does marker dry out?
science	geography	animal and human populations, rivers, mountains…	how many people are there on earth?
science	human body/health	questions about the human body, bodily functions, vaccines, medicine, etc.	how many teeth do adult humans have? Why do we poop? How tall is the average 5th grader?
science	nature	questions about the natural world other than plants and animals	why is the sky blue?
science	plants	vegetable kingdom	what's inside a flower?
science	science	catch all category for questions about scientific classifications, species, experiments…	are humans today considered homo sapiens sapiens?
science	space	questions about space, planets, stars…	what is the orange spot in Venus?
science	transportation	questions about airplanes, buses, garbage trucks…	how do helicopters carry people?
unclear	unclear	when we humans cannot understand what the child wanted to know (but we know it is a question)	how can english count? how can an oodle eat a soodle?
varies	future	prediction of the future	Will dad buy Chinese food for dinner?