Using Conversational Agents to Foster Young Children's Science Learning from Screen Media

Ying Xu
University of California, Irvine
Irvine, CA 92697, USA
ying.xu@uci.edu

IDC '20 Extended Abstracts, June 21–24, 2020, London, United Kingdom

ACM ISBN 978-1-4503-8020-1/20/06.
https://doi.org/10.1145/3397617.3398031

Abstract

My dissertation project leverages an intelligent conversational agent embodied in an on-screen character to add social contingency into children's science video watching experiences. This conversational agent has been developed in an iterative process and embedded in a new PBS KIDS science show, "Elinor Wonders Why." My research aims to understand the design, feasibility, and promise of virtual conversation with media characters in young children's informal science learning through video watching.

Author Keywords

Conversational agents; social learning; screen media; science learning; children

CCS Concepts

•Human-centered computing → Natural language interfaces; Empirical studies in HCI; •Social and professional topics → Children;

Context, Motivation, and Related Work

Science-oriented television programs can be an important source of science learning for young children [4]. However, the educational benefits of these programs have long been limited by children not being able to interact with the content in a meaningful way [11]. One common approach to providing children with interactions during their television/video

Figure 1: Elinor, the main character of *Elinor Wonders Why*. The conversational agent will be embodied in Elinor. ©2020 SHOE Ink, LLC.

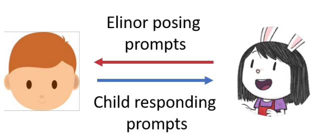

Figure 2: The conceptual design of the CA that allows children to verbally converse with the main character

watching is to embed conversational prompts in the media sources themselves, as exemplified in *Dora the Explorer*. These conversational features typically consist of a character creating faux eye contact, asking the viewer a question, pausing a set amount of time, and responding in a way that does not actually acknowledge the viewer's specific answer. One major limitation of such conversational features is the lack of contingency and responsiveness, which may discourage children from answering the questions [15], misgauge their own learning comprehension [5], or doubt the character as a reliable source of information [6].

With the rapid development of conversational agents (CAs), however, contingent interactions between children and media are now possible. As intelligent systems, CAs have the affordances to understand unconstrained natural language input, allowing for complex dialogue and potentially mimicking human-to-human spoken conversation. These technologies are now prevalent in many homes, and children readily interact with and accept these digital conversational partners in their daily lives [3, 9]. This points to the feasibility of integrating CAs into educational television programmings [14].

My dissertation research uses a CA to enable verbal interactions between children and the main character in a children's science animation series (Figure 1), who can play the role of an engaging and responsive video co-viewer for children (Figure 2). The verbal interactions are designed to promote active viewing, in which children are "minds on," or actively thinking and reasoning rather than passively watching the video [7]. In addition, socially interacting with onscreen characters provides a fun experience, which may enhance children's enjoyment and motivation during video watching [17].

This project is informed by three strands of literature. The first body of literature focuses on the benefits of co-viewing with children during television watching. Research consistently suggests that children learn more from television watching when conversing with parents than when watching alone, with the benefits including improved comprehension, increased language development, and greater enjoyment [1, 2, 10]. However, such interaction with family members is not common [12, 15]. As such, Takeuchi and Steven called for designing digital media to promote "joint-media engagement," facilitating the social aspect of media experiences [16]. The second strand of literature focuses on preschool-aged children's everyday interactions with voice-based CAs. Children appear to interact with CAs in much the same way as they interact with human partners [13]. These studies provide important evidence of the feasibility of building educational applications that revolve around children's verbal interaction with CAs. The third strand of literature focuses on intelligent systems with voice interface, such as robots, for preschool-aged children's learning. For example, Xu and Warschauer developed a CA that narrates picture books to children and engages them in story-related conversation, and found that children enjoyed this learning experience (e.g., [8]). These studies found that intelligent systems can effectively engage children in learning-related conversation.

Research Questions

My dissertation project aims to understand the design, feasibility, and promise of virtual conversation with media characters in young children's informal science learning through video watching. To this end, I have developed conversational videos videos as a supplementary part of Elinor Wonders Why, a new PBS KIDS animated television scheduled to debut in September 2020. The conversational videos allow children to directly speak with Elinor as she solves everyday science mysteries, thus priming children to en-

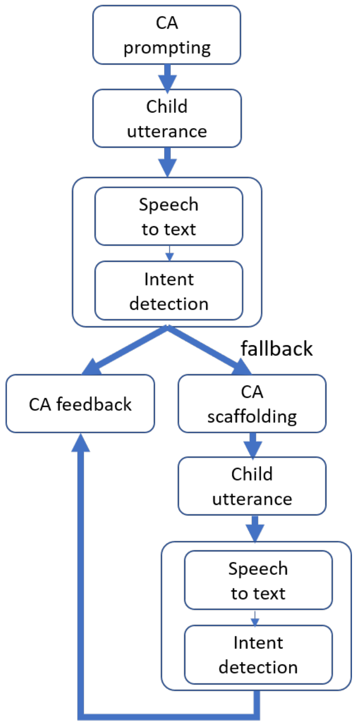

Figure 3: The CA workflow. The scaffolding mechanism will be triggered if a child's response to an initial prompt is classified as "fallback" by the CA.

gage in observation, prediction, pattern identification, and problem solving through scaffolded conversation. The agent also offers contingent feedback that varies based on children's responses. The project is guided by five foci:

1. Can CAs feasibly help young children engage with and learn from science video watching?
2. If so, how should the conversational experiences be best designed to enhance children's learning and engagement?
3. What are the effects of using CAs to scaffold children's video watching?
4. Do these effects vary by children's age, gender, English language proficiency, and prior CA experiences?
5. How do parents and children perceive these interactive videos?

These questions will be answered by a three-phase study, including the development, field testing, and an efficacy study involving randomized control trial (RCT). Children aged 4 to 6 years will be recruited to participate in the field testing and pilot RCT from two communities with diverse cultural and socioeconomic backgrounds in Southern California. A wide range of data sources will be collected.

Initial Work

CA Development

I have developed a CA for one 11-minute episode of *Elinor Wonders Why* where Elinor and her friends explore how bees make honey. Three conversation prompts were integrated at the beginning of the episode as casual talk to evoke children's curiosity of the content (e.g., *"Do you like honey?" "How does it feel when you touch honey?"*). Another six prompts were embedded throughout the episode, with a goal of involving children in the scientific inquiry process with the main character (e.g., *"How do bees make the runny juice into goopy honey?" "How can we get honey off Ari's <one character in the show> face"*)

Using Google's Dialogflow engine, the agent underlying Elinor learns to understand children's responses both from the pre-trained language models already built into this engine as well as training phrases that we provide, which are sample phrases of what children may say to respond to a particular conversational prompt. For each conversational prompt, we predefine four types of intent categories that we want Elinor to classify the utterances into, including (1) the anticipated answers, (2) requests to repeat the question, (3) indications of an inability to answer the question, and (4) a fallback intent that handles the edge cases where the child's utterance does not match any of the other intents or the child does not respond to the prompt at all. Given the open-ended nature of the prompts and that children may respond to particular questions in several distinct ways, there is usually more than one intent related to the anticipated answers. After the agent classifies the child's responses into one of the intent categories, differentiated feedback is given based on the classification. Elinor's feedback is designed not only to tell children whether their responses are correct or incorrect, but to encourage children to reflect on their responses, think more deeply, and maintain their curiosity and interest in Elinor's further explorations.

Moreover, the fallback intent is combined with a scaffolding mechanism. Children are provided with two follow-up opportunities to answer one set of questions, with the follow-up prompt rephrasing the original prompt into a multiple-choice format. In this way, the system restricts interaction within its capabilities. The scaffolding mechanisms combined with edge case handling ensures the system would still be able to drive the conversation along the context even if the CA does not understand the exact utterance of a participant. Figure 3 displays the workflow of the CA.

This language training model is being optimized constantly during the course of field testing and pilot RCT as we collect more data on children's responses. Specifically, we have modified the intents (e.g., add more intents to encompass other common response categories) and included more training phrases to increase the accuracy of intent classification.

Figure 4: Child participants using conversational videos during the field testing. The field testing was carried out in childcare centers. Children interacted with the conversational video independently.

Findings from Field Testing
I have field tested the usability of the conversational video with 8 children, in order to understand how children actually interacted with and learn from the CA. During the field-testing, I utilized the built-in camera of the laptop to capture children's visual attention and also assessed children's learning of scientific concepts after the video watching. I found that the conversational moments with Elinor effectively enhanced children's attention to the episode: children were frequently observed to redirect their visual attention back to the screen when Elinor asked them questions. Children also learned the scientific content taught in the episode: they correctly answered the majority of the post-viewing assessment questions, which was a significant improvement from what these children had known about this topic before the video viewing.

Planned Next Steps
Insights gained from the field study have been used to iteratively improve the embedded conversational function in the mobile application. I will carry out an efficacy study to further evaluate the effects of the improved CA by conducting a four-way experiment with 40 children in each group (160 children in total): (1) watching *Elinor Wonders Why* with a CA (embodied in Elinor), (2) watching *Elinor Wonders Why* with a human partner, (3) watching *Elinor Wonders Why* with pseudo interactions where Elinor asks questions but does not listen to or understand children's answers, and (4) watching *Elinor Wonders Why* with no dialogue. We will collect baseline data on demographics, language level, and interest in science learning; engagement data on children's visual attention, facial expression, verbal utterances, and self-reported perceptions; and outcome data on story comprehension and learning of scientific concepts. This will allow me to examine what benefits, if any, children receive from engaging with a CA while watching a science video, and how these benefits compare to conversing with a human partner. This evaluation will provide a comprehensive evaluation of the promise of incorporating CAs into screen media to foster young children's science learning and engagement.

What I Will Bring to and Benefit from DC
First, my dissertation research examines the potential for conversational agents to function as children's social learning partners in a screen-based science activity, which may expand other scholars' awareness of this possible application for CAs as well as of this pioneering technology to support children's science learning. Second, at the time of the conference, I will have preliminary results regarding the design, feasibility, and effectiveness of incorporating a conversational agent into children's screen media, which may be useful for scholars studying in relevant areas. Third, my research could benefit from insightful feedback from other IDC attendees. Given that I will have 11 months left in my doctoral program following the conference, the feedback I receive from other researchers could be fundamental in helping me refine my research.

REFERENCES

[1] Brigid Barron, Caitlin Kennedy Martin, Lori Takeuchi, and Rachel Fithian. 2009. Parents as learning partners in the development of technological fluency. (2009).

[2] Sabrina L Connell, Alexis R Lauricella, and Ellen Wartella. 2015. Parental co-use of media technology with their young children in the USA. *Journal of Children and Media* 9, 1 (2015), 5–21.

[3] Stefania Druga, Randi Williams, Cynthia Breazeal, and Mitchel Resnick. 2017. Hey Google is it OK if I eat you?: Initial explorations in child-agent interaction. In *Proceedings of the 2017 Conference on Interaction Design and Children*. ACM, 595–600.

[4] Shalom M Fisch. 2014. *Children's learning from educational television: Sesame Street and beyond*. Routledge.

[5] James Paul Gee. 2005. Learning by design: Good video games as learning machines. *E-learning and Digital Media* 2, 1 (2005), 5–16.

[6] Paul L Harris and Melissa A Koenig. 2006. Trust in testimony: How children learn about science and religion. *Child development* 77, 3 (2006), 505–524.

[7] Kathy Hirsh-Pasek, Jennifer M Zosh, Roberta Michnick Golinkoff, James H Gray, Michael B Robb, and Jordy Kaufman. 2015. Putting education in "educational" apps: Lessons from the science of learning. *Psychological Science in the Public Interest* 16, 1 (2015), 3–34.

[8] Jacqueline Kory and Cynthia Breazeal. 2014. Storytelling with robots: Learning companions for preschool children's language development. In *The 23rd IEEE international symposium on robot and human interactive communication*. IEEE, 643–648.

[9] Silvia B Lovato, Anne Marie Piper, and Ellen A Wartella. 2019. Hey Google, Do Unicorns Exist? Conversational Agents as a Path to Answers to Children's Questions. In *Proceedings of the 18th ACM International Conference on Interaction Design and Children*. 301–313.

[10] Marie-Louise Mares and Zhongdang Pan. 2013. Effects of Sesame Street: A meta-analysis of children's learning in 15 countries. *Journal of Applied Developmental Psychology* 34, 3 (2013), 140–151.

[11] Lydia Plowman, Christine Stephen, and Joanna McPake. 2010. Supporting young children's learning with technology at home and in preschool. *Research Papers in Education* 25, 1 (2010), 93–113.

[12] Victoria J Rideout and Elizabeth Hamel. 2006. *The media family: Electronic media in the lives of infants, toddlers, preschoolers and their parents*. Henry J. Kaiser Family Foundation.

[13] Alex Sciuto, Arnita Saini, Jodi Forlizzi, and Jason I Hong. 2018. Hey Alexa, What's Up?: A mixed-methods studies of in-home conversational agent usage. In *Proceedings of the 2018 Designing Interactive Systems Conference*. ACM, 857–868.

[14] Subhasree Sengupta and Radhika Garg. 2019. Impact of Voice-based Interaction on Learning Practices and Behavior of Children.. In *IUI Workshops*.

[15] Gabrielle A Strouse, Katherine O'Doherty, and Georgene L Troseth. 2013. Effective coviewing: Preschoolers' learning from video after a dialogic questioning intervention. *Developmental psychology* 49, 12 (2013), 2368.

[16] Lori Takeuchi, Reed Stevens, and others. 2011. The new coviewing: Designing for learning through joint media engagement. In *New York, NY: The Joan Ganz Cooney Center at Sesame Workshop*.

[17] Ying Xu and Mark Warschauer. 2019. Young Children's Reading and Learning with Conversational Agents. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, CS10.