# Impact of Voice-based Interaction on Learning Practices and Behavior of Children

Subhasree Sengupta
School of Information Studies
Syracuse University
Syracuse, NY, USA
susengup@syr.edu

Radhika Garg
School of Information Studies
Syracuse University
Syracuse, NY, USA
rgarg01@syr.edu

## ABSTRACT

Smart devices have become an integral part of the everyday lives of children. Today, children can even use voice-based interactions to interact with devices for a wide range of activities. Previous research has shown that voice-driven interfaces have a potential to offer a potent new mechanism for teaching, engaging, and supporting children in daily life. Our paper, therefore, argues that it is critical not only to investigate how children use voice-based interactions to communicate with devices (e.g., smart speakers) but also the nature of relationships that children form with these devices, the influence such use has on children's learning and behavior, and the role that parents or guardians play in deciding the norms of use for children. We also propose to explicitly and intricately investigate complexities in use and its impact relative to entangled identities (conveyed through overlapping attributes of gender, ethnicity, race, class) and larger social systems. To this end, we propose to use Social Learning Theory to understand how children learn through observing and interacting with smart devices, specifically using voice-based commands. Methodologically, we will conduct participatory design sessions and follow-up interviews to get a nuanced understanding of how children mentally contextualize voice-enabled smart devices and how social influence (e.g., parental expectation/norms), social function of identification (e.g., children's emotional connection with technology), and learning goals impact their usage patterns.

## CCS CONCEPTS

• **Human-centered computing** → **User studies**; **Empirical studies in HCI**; **Scenario-based design**; **Participatory design**.

## KEYWORDS

Voice-based interactions; Children's behavior and learning practices; Social learning theory; Parasocial relationships

**ACM Reference Format:**
Subhasree Sengupta and Radhika Garg. 2019. Impact of Voice-based Interaction on Learning Practices and Behavior of Children. In *Joint Proceedings of the ACM IUI 2019 Workshops, Los Angeles, USA, March 20, 2019*. ACM, New York, NY, USA, 3 pages.

*IUI Workshops'19, March 20, 2019, Los Angeles, USA*




## 1 INTRODUCTION

Today, various types of smart devices are deeply integrated in our day to day lives. The use of technology has increased not only for adults but also for children, be it as a source of entertainment or as a learning aid. So much so, that the exposure and use of technology has been considered as a crucial influence on the process of learning and development of children [5]. Bower and Sturman demonstrated that wearable devices offer a range of pedagogical uses (in-situ contextual information, recording, simulation, communication, first-person view, in-situ guidance, feedback, distribution and gamification), afford benefits to educational quality (engagement, efficiency, and presence), and provide logistical advantages (hands-free access and free up space) in a class room setting [2]. More recently, smart devices (e.g., smart phones, tablets, smart speakers) have started to offer conversational assistants (e.g., Amazon Alexa, Siri, and Google Now) that lend flexible means of interacting with the device. Due to the presence of such voice assistants, children no longer need to read or write to be able to interact with the devices [9]. As the amount of background information a child needs to use these devices has reduced, it can have an impact in the information seeking, behavioral (e.g., children might imitate and emulate certain characteristics of these devices), and learning practices pursued by children and the factors that affect these practices. Hence, in this paper we argue that it is critical to investigate how and why children are using these devices (e.g., voice-connected speakers), and the influence voice-based interactions with devices has on children's behavior and learning practices.

We propose to investigate this issue through the lens of Bandura's Social Learning Theory (SLT). It explains 'observational learning' in terms of how people learn through observing othersfi behavior, attitudes, and the outcomes (penalty or reward) one might incur due to such a behavior [1]. However, SLT is a complex and subjective concept with many different facets to it, exploring all of which is beyond the scope of this study. Therefore, to understand 'observational learning' our study centers around the three social factors provided by Over et al [10]. These three factors are: social function of identification (parasocial relationships), the type of role that children associate with voice assistants; learning goals, the type of learning tasks that children use voice assistants for; social norms and customs, particularly focussing on the role that parents/guardians play in regulating children's use of technology. Therefore, we will investigate three fundamental research questions in this paper:

- RQ1: What are the type of parasocial relationships that child form with the voice-enabled smart devices?

- RQ2: What are the type of learning objectives or tasks that children are interested to use voice-enabled smart devices for and how parasocial relationships may or may not impact those?
- RQ3: How do social norms and customs (especially those instilled/followed by parents and guardians) affect the way children use voice-connected smart devices?

To answer these questions we aim to conduct Participatory Design (PD) sessions and set of interviews with design groups consisting of children from different age groups: 7- 9 years, 10-12 years and 13-17 years. This will enable us to investigate the role age and gender of children within the context of our research questions. For example, we will explore if the nature of parasocial relations/role that younger children associate with voice-enabled smart devices differ from those that older children associate with these devices.

## 2 BACKGROUND AND RELATED WORK

In this section we first discuss the framework by Over et al [10] that acts as a foundation of our proposed study. Then, we discuss prior studies that have looked into what type of activities children are using voice-enabled devices for, parental role in use of voice-enabled devices by children, and type and influence of parasocial relationships that children develop with such devices.

Oven et al. highlighted three crucial factors that impact the selectivity in 'observational learning' (i.e the fact that people selectively choose to imitate or emulate certain behavior they observe) [10]. These factors were used to operationalize SLT to be able to understand children's use, how they relate with these devices, and how it impacts their learning practices.

**Parasocial relationships:** The first factor that Over et al. [10] stated is 'social function of identification' that explains that children emulate and establish an emotional connection with those they feel they resemble or want to be like. For example, they may find the voice of a device relatable or they may find a character in a game that they play on a device relatable and start to personate the characteristics of those. More importantly, children sometimes assert roles (such as friend, mentor, pet) to these devices thereby personifying them and forming a relationship with these devices. These sort of connections have also been termed as parasocial relationships (i.e., one-sided, emotionally driven relationships that children develop with media characters) [3, 7]. Druga et al. demonstrated that the ability to have voice-based interactions, elements of social realism, and human-like characteristics makes these smart devices more relatable and easy to use for children [6]. Brunick et al. [3] highlighted how parasocial relationships can be useful for developing educational tools for children by embedding in intelligence agents ability to generate parasocial interactions, such as conversational timing and response personalization. Gray et al [7] emphasized that factors such as social realism and personification should be considered when designing an intelligent agent for children. Therefore, we propose to explore the type of relationships children form particularly with voice-enabled smart devices and how that impacts the type of learning goals that children use these devices for.

**Learning goals:** The second factor that Over et al. [10] put forth is 'learning goals', which includes self-established or self-motivated learning goals of children. For example, if a child does not know how to do something and wants to learn about it, he/she may ask the device for help and use the information gained from these devices to perform that task. Preliminary work by Lovato et al. highlighted that children in general either 'explore' voice assistants such as Siri and Google Now or use them to 'seek new information' [9]. In 'exploration' children use the voice assistants for fun and are even able to develop a bond with the voice assistants. In 'information seeking' children use the voice assistants to find facts and develop knowledge base. Both these forms of use have impact on children's development. However, the main source of data for their study were Youtube videos of children's activity, which might not be representative of children's actual usage patterns. We aim to add to this work, by investigating actual logs of voice commands to understand the categories of use. Further, usage patterns can greatly differ by age and gender, thus while annotating the voice commands we will also explore how usage differs based on age and gender. In order to understand this better we will conduct a design session to investigate the kind of devices children would like to use for varying learning goals that they might have the role of parasocial relationships in this process.

**Role of parents:** The final influencing factor that Over et al. [10] put forth is 'social influence', which comprises of social customs, expectations, and norms that might affect the way children use or communicate with a device. Particularly we will focus on the role parents or guardians play a key in role in introducing a smart device to children and establishing norms of use and the extent of use, and children's use of devices might also be influenced by the way their parents or other older members of the family use these devices. Cheng et al. presented four roles parents play in helping children communicate with voice-controlled devices [4]. Parents may also help establish boundaries for device usage by children as illustrated by related work (e.g., [8]). They may control the amount of time, the type of content and the nature of interactions children may have with these devices thereby influencing the learning practices of children. Hiniker et al. demonstrated that parents are vital in scaffolding children's use of a novel/relatively newer technology [8]. Therefore, we propose to expand on prior work by including parents in PD sessions to identify how they support or regulate children's use of voice-connected devices for learning, and how this differs by age and gender of children.

## 3 PROPOSED METHOD

Our study will be conducted in following two stages:

### 3.1 Historical Log Analysis

In this stage we will deploy a survey on Amazon Mechanical Turk (AMT) to collect children's voice history logs comprising of their interactions with smart speakers. The survey will also consist of questions on on family structure, number of children in families, types of smart speakers. The primary goal of the survey is to annotate voice history logs to get an understanding of the type of activities that children use smart speakers for and get an estimate of the percentage of those used for learning. Further, we will also analyze how the usage patterns, learning tasks differ by age and gender of children.

### 3.2 Participatory Design Sessions

In second stage we will employ Cooperative Inquiry [11] for conducting PD sessions that will focus on co-designing devices/technology with children that they might like to use for learning. Each of these sessions will be divided in two parts, the first part called circle time will be used to help the participant better contextualize the task they are about to do in the session and the second part will be the actual design prompt based on which the participants will perform a design activity.

**Design Session 1 (DS1)**

Prior research has shown that children perceive interactive media characters as enjoyable companions and develop different parasocial relationship with them. Therefore, the design session will begin with the circle time (15 minutes), where we will ask participants to share with us their favorite cartoon or media character. The aim of circle time is to ask "question of the day" to get adults and children started. After that we will ask them to design an interactive technology/device to identify kind of roles children would like the such devices to take, specifically as they use them for various learning tasks (e.g explore unknown facts, improve their language skills, or to help them to hone their deductive reasoning). We will utilize *Bags-of-stuff*, *Big Paper* and *Layered Elaboration* PD techniques [11]. Through such activities we will investigate two things: 1) The type of parasocial roles that children see these smart devices to take, and 2) the connection between the parasocial role and the type of learning task.

**Design Session 2 (DS2)**

In this design session we will elicit information regarding how children think about using different speech agents (e.g, Amazon Alexa, Google Now) for various learning tasks, using *Stickies* PD [11] technique. The circle time prompt for this session will ask participants how they (would) use a device/technology to explore a fact they wish to learn about. The design prompt ask participants to note what they liked, disliked or would like to improve based on the current devices, using the design technique of Stickies.

**Design Session 3 (DS3)**

In the final session we will use *Stickies* [11], *Layered Elaboration* [11] as design methods to have children and parent revisit design ideas that children built in DS1 and build them with the help of parents/guardians. One of the member of the research team will present the ideas generated during DS1 through storyboards. The circle time prompt will have participants think about how parents influence children's use of technology. For the design prompt, the parents and children will be then asked to explain their likes, dislikes, and further design ideas. In such a way parents and children would work together to make the storyboards better based on each other's ideas and opinions. For example, parents might want to include the possibility of regulating children's use (e.g., permitted duration, tone of the device) in the designs proposed by children.

## 4 CONCLUSION

Our position paper proposes to investigate how voice-based interactions with smart devices are affecting or can affect learning practices of children. Particularly, we use the three factors by Oven et al. to operationalize the use of SLT as a tool to answer our research questions on how children mentally contextualize voice-enabled smart devices and how social influence (e.g., parental expectation/norms), social function of identification (e.g., children's emotional connection with technology), and learning goals impact their usage patterns. To this end, we propose to employ both historical log analysis and participatory design sessions.

## REFERENCES


[1] Albert Bandura. 1977. *Social learning theory*. Prentice Hall.

[2] Matt Bower and Daniel Sturman. 2015. What are the educational affordances of wearable technologies? *Computers & Education* 88 (2015), 343–353.

[3] Kaitlin L Brunick, Marisa M Putnam, Lauren E McGarry, Melissa N Richards, and Sandra L Calvert. 2016. Children's future parasocial relationships with media characters: the age of intelligent characters. *Journal of Children and Media* 10, 2 (2016), 181–190.

[4] Yi Cheng, Kate Yen, Yeqi Chen, Sijin Chen, and Alexis Hiniker. 2018. Why Doesn't It Work? Voice-Driven Interfaces and Young Children's Communication Repair Strategies. In *Proceedings of the 17th ACM Conference on Interaction Design and Children*. ACM, 337–348.

[5] Cynthia Chiong and Carly Shuler. 2010. Learning: Is there an app for that. In *Investigations of young children's usage and learning with mobile devices and apps. New York: The Joan Ganz Cooney Center at Sesame Workshop*. 13–20.

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

[7] James H Gray, Emily Reardon, and Jennifer A Kotler. 2017. Designing for Parasocial Relationships and Learning: Linear Video, Interactive Media, and Artificial Intelligence. In *Proceedings of the 2017 Conference on Interaction Design and Children*. ACM, 227–237.

[8] Alexis Hiniker, Bongshin Lee, Kiley Sobel, and Eun Kyoung Choe. 2017. Plan & play: supporting intentional media use in early childhood. In *Proceedings of the 2017 Conference on Interaction Design and Children*. ACM, 85–95.

[9] Silvia Lovato and Anne Marie Piper. 2015. Siri, is this you?: Understanding young children's interactions with voice input systems. In *Proceedings of the 14th International Conference on Interaction Design and Children*. ACM, 335–338.

[10] Harriet Over and Malinda Carpenter. 2012. Putting the social into social learning: explaining both selectivity and fidelity in children's copying behavior. *Journal of Comparative Psychology* 126, 2 (2012), 182.

[11] Greg Walsh, Elizabeth Foss, Jason Yip, and Allison Druin. 2013. FACIT PD: a framework for analysis and creation of intergenerational techniques for participatory design. In *proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2893–2902.