Touchscreen Prompts for Preschoolers: Designing Developmentally Appropriate Techniques for Teaching Young Children to Perform Gestures

Alexis Hiniker, Kiley Sobel, Sungsoo (Ray) Hong, Hyewon Suh, India Irish, Daniella Kim, Julie A. Kientz
Human Centered Design and Engineering
University of Washington
{alexisr, ksobel, rayhong, hyewon25, indiai, dkella, jkientz}@uw.edu

ABSTRACT
Though toddlers and preschoolers are regular touchscreen users, relatively little is known about how they learn to perform unfamiliar gestures. In this paper we assess the responses of 34 children, aged 2 to 5, to the most common in-app prompting techniques for eliciting specific gestures. By reviewing 100 touchscreen apps for preschoolers, we determined the types of prompts that children are likely to encounter. We then evaluated their relative effectiveness in teaching children to perform simple gestures. We found that children under 3 were only able to interpret instructions when they came from an adult model, but that children made rapid gains between age 3 and 3-and-a-half, at which point they were able to follow in-app audio instructions and on-screen demonstrations. The common technique of using visual state changes to prompt gestures was ineffective across this age range. Given that prior work in this space has primarily focused on children's fine motor control, our findings point to a need for increased attention to the design of prompts that accommodate children's cognitive development as well.

Categories and Subject Descriptors
H.5.2 User Interfaces: Input devices and strategies, Interaction styles

General Terms
Design, Experimentation, Human Factors.

Keywords
Toddlers, Preschoolers, Gestures, Prompts, Tablets, Touchscreens, CCI, Usability, Instructions.

1. INTRODUCTION
As tablets and mobile devices have proliferated worldwide, so has children's consumption of touch-screen media. Seventy-five percent of American families with a child under 8 own a touchscreen device, and iPad ownership in families with young children saw a five-fold increase (to more than 40% of households) between 2011 and 2013 [29]. This increase in access extends to toddlers and preschoolers as well. Eighty percent of children between the ages of 2 and 4 are mobile device users, and the average American 2-to-4-year-old spends nearly 20 minutes a day using a tablet or smartphone [29].

Though prior work demonstrates that these children are physically capable of performing many of the most common mobile app gestures [3], relatively little is known about how children learn which gestures are contextually relevant or how they learn to perform gestures that are unfamiliar. As developers are increasingly creating educational apps that target this age group, it is important to understand the extent to which preschoolers can meaningfully engage with this content. Only if children understand how to trigger interactive app elements with intention can we expect them to be able to access the designed content or extract intended educational concepts.

In this paper, we examine toddlers' and preschoolers' responses to four different mechanisms for prompting users to perform specific gestures: three unique in-app prompts, and an in-person model of the gesture by an adult. Each of these cues is intended to suggest in advance how a user should interact with a system. To inform this work, we first conducted an initial review of 100 semi-randomly selected apps on the iTunes app store designed for children under the age of 5. We identified the most common techniques across this sample for prompting users to perform specific gestures: visual state changes, audio instructions, and a demonstration by a cartoon hand. Based on this analysis, we designed and implemented materials to comparatively assess the effectiveness of these different styles of prompting in teaching children to perform different gestures.

We conducted an empirical study with 34 children between the ages of 2 and 5 to assess their ability to understand each of these four prompt types. We paired each of these prompting techniques with a set of four gestures, allowing us to examine each technique's effectiveness in a variety of contexts. We examine children's ability to interpret each of these prompt types as a function of age, and we identify the age at which this population develops a stable understanding of each prompting technique. In this investigation, we sought to explore the following research questions:

R1: Do very young children accurately interpret prompts to perform specific tablet gestures?

R2: Which prompts are most likely to result in accurate interpretation?

R3: How does this change with age?

IDC '15, June 21 - 25, 2015, Medford, MA, USA
Copyright is held by the owner/author(s). Publication rights licensed to ACM.
ACM 978-1-4503-3590-4/15/06…$15.00
DOI: http://dx.doi.org/10.1145/2771839.2771851

This research represents the first comparative evaluation of popular prompting techniques for very young children and the first documentation of their relative effectiveness as tools for eliciting common gestures from children. The contribution of this work is to examine 1) how children respond to commonly used prompting mechanisms and 2) at what ages children acquire the ability to understand each of these prompt types. By understanding when prompts will and will not be effective, we hope to support designers in guiding their youngest users toward successful interaction behaviors.

2. RELATED WORK

2.1 Gesture Performance and Children

A large body of prior work has investigated touch-screen gestures and interface design [8, 25, 34]. Though much of this work has considered adult users only, in recent years, a growing body of research has begun to investigate children's and adolescents' use of gestures as well [2, 5, 28]. Previous research suggests that between the ages of 8 and 11, school-aged children approach adult-like maturity in their performance of basic, one-handed gestures such as tap, drag, swipe, and pinch [4, 30]. However, when performing more complex and custom gestures, children and teens' touch interactions are less likely to be interpreted correctly by gesture recognizers [1]. These performance gaps shrink as children age into later school-age years and adolescence; yet even these older children remain a distinct user group with less precise gesture performance than adults [10].

As school-age children mature cognitively, they also demonstrate a more abstract understanding of gestures. In a test of a mixed media prototyping tool, children 8- to 12-years-old were asked to create low-fidelity interfaces for mobile devices using sketches on clear panels. When children were asked to describe their prototypes to a group, they correctly used terms such as "pinch," "tap," "flick," and "scroll" [9]. This suggests that at this age, children have a clear mental model of what these gestures mean and the effect they should each produce on an interface. Similarly, in an assessment with 8- to 11-year-olds, researchers found that organic, child-defined gestures are similar to gestures defined by adult users [30]. This is also consistent with children's mental representations of gestures beginning to resemble adults' in early adolescence.

Across childhood, young users acquire both physical and cognitive skills that bring their gesture-related understanding, intentions, and abilities in line with that of adults. In this investigation, we leveraged this prior work by intentionally examining preschoolers' responses to gesture prompts across an age range of several years. Rather than treating preschoolers as a monolithic design target, we focus on how performance changes with age across our sample.

2.2 Gesture Performance and Preschoolers

Relative to school-age children, very little work has evaluated preschoolers' and toddlers' ability to perform touchscreen gestures and this body of research has focused almost exclusively on children's fine-motor control. Aziz and colleagues demonstrated that children can master some gestures as early as age 2 [4]. In their study, researchers observed children between the ages of 2 and 12 as they played with five different educational tablet apps and performed seven fundamental gestures: tap, flick, slide, drag and drop, rotate, pinch and spread. Two- and three-year olds were found to struggle with drag and drop, rotate, and spread. Four-year-olds were observed to be proficient in all seven gestures. However, as this investigation included only three participants of each age, the extent to which these findings generalize to the broader population of preschoolers is unclear.

A second study of gesture performance in preschoolers demonstrated both that children can perform basic gestures from an early age and that their accuracy increases steadily between the ages of 3 and 6 [33]. All participants attempted tap, double tap, single-touch drag and drop, and multi-touch drag and drop gestures; 3-year-olds were successful 73% of the time, while children older than 5 were successful 89% of the time. Adults remained 30% more accurate on tap tasks and 10% more accurate on straight-line drag and drop tasks than the best performing children. A third study examined 4- to 6-year-olds interactions with a tabletop computer [20]. The researchers found that while younger children required more hints from adults, all children were able to interact with the touchscreen and naturally performed basic gestures such as tapping and dragging. They found no physical barriers to gesture performance for these older preschoolers and kindergarteners.

Thus, while some prior research indicates that very young children can perform gestures, this remains an understudied area where the focus to date has been on fine motor control and the physical performance of gestures. The ways in which children's cognitive development relate to their gesture performance remain unexplored. Our investigation expands this body of literature by examining instructional techniques that predict accurate gesture performance in preschoolers and the extent to which these techniques are comprehensible to preschoolers of varying ages. Though we know preschoolers are at times able to perform productive touchscreen gestures, our work documents the context in which we should expect these children to be successful and the design decisions that facilitate that success.

2.3 Gesture Prompts

The physical ability to perform a gesture is useful only if the user understands when and why he or she should perform it. Thus, many studies have investigated how various design decisions scaffold such knowledge and translate into the ability to elicit specific interactions from users [34]. While past work to define prompting techniques for touchscreen gestures centers on adult users, a few studies have evaluated prompting techniques with children. McKnight and Fitton studied how to teach gestures to 6- and 7-year-old children and evaluated the effectiveness of specific language and terminology choices in written and audio instructions [23]. They report that at this age children are unfamiliar with touch-screen terms such as "select" or "press and hold," but are able to understand terms with real-world applicability, such as "slide" and "swipe" [23]. Based on their analysis, the research team developed a set of design guidelines for creating mobile device interfaces for children age 7 to 10 [22].

However, little prior work has assessed younger children's understanding of gesture prompts. Nacher and colleagues found that 2- to 3-year-old children were more likely to perform drag and scale gestures correctly in response to a prompt from an animated hand symbol than in response to a static hand symbol with arrow icons [24]. In a separate study, Nacher and colleagues found that preschoolers could perform a variety of gestures (tap, double tap, long press, drag, rotation, and scaling) successfully when prompted by an adult. In both studies, child participants completed a gesture-training exercise with a researcher before beginning the assessment.

Our work builds on these early investigations into preschoolers' ability to interpret prompts correctly by evaluating the effectiveness of previously unassessed prompting techniques and by examining the evolution of children's understanding across the preschool years. We also conduct the first assessment of in-app prompts for preschoolers without training participants in advance to recognize the prompts or to perform the associated gestures. This lack of training is intended to increase ecological validity and the likelihood that children's responses to our prompts are reflective of their responses to the unfamiliar prompts they might encounter in the wild.

2.4 Prompting Outside of CCI

Given the limited literature on prompts for preschoolers in digital interfaces, we also looked at evidence-based prompting techniques used with very young children in other contexts. Behavioral therapies for children often involve a highly structured prompt hierarchy that describes successive techniques for guiding a child toward a desired behavior [19]. This hierarchy is designed to grow from least to most intrusive, with more intrusive prompts providing more adult involvement and increasing the likelihood that children will perform the prompted behavior. While the prompt categories can differ by task type, one common hierarchy is listed below [21], ordered from least- to most-intrusive:

- Verbal instructions
- Gestural prompts
- Modeling
- Physical guidance

Commonly, prompt hierarchies recommend first using less-intrusive prompts, such as verbally describing what the child should do (e.g., saying, "Touch the yellow ball") or gesturing toward a target item (e.g., pointing to the yellow ball). When less-intrusive prompts fail, the adult can model the intended behavior (e.g., the adult touches the yellow ball) or physically guide the child to do the intended behavior (e.g., pick up the child's hand and put it on the yellow ball) [7, 14]. Adult models provide more support and increase the likelihood that children successfully perform the desired behavior, relative to prompts that indicate what to do indirectly.

Though these in-person prompting techniques are not available to app developers, we hypothesized that in-app prompts resemble less-intrusive prompts in this hierarchy, such as verbal instructions, and that an adult model would therefore be a more effective teaching tool. As adults frequently facilitate children's use of touchscreen devices [31], we chose to include an in-person model in our assessment of techniques for teaching touchscreen gestures.

3. PRELIMINARY WORK

3.1 App Store Review

As no existing work documents the most common prompts in children's apps, we conducted a preliminary investigation to understand common practices in this space. We searched the iTunes App Store for applications with iTunes' "Kids 5 and Under" designation. No straightforward mechanism exists for viewing all apps in the store and selecting from them randomly, thus we searched for lists of the most popular apps, looked for recommendations from Common Sense Media, reviewed award sites such as the Children's Technology Review and the Parents' Choice Awards, and looked at the apps featured by iTunes. We chose 100 unique titles created by 68 unique app developers. These spanned a variety of categories and included a mix of educational and entertainment content.

We played each app and recorded all of the ways in which the app guides a child toward performing productive gestures. This included sampling multiple levels, exercises, and mini-games; intentionally playing incorrectly; and intentionally waiting for an extended period of time to give the app the opportunity to prompt the user in response to a lack of input. The most common prompts were:

- Audio instructions (50% of apps)
- A visual change to the item to touch (such as sideways motion to indicate that an element can be swiped, or pulsing to indicate that an item can be tapped) (41%)
- Text-based instruction (19%)
- A demonstration by a cartoon hand (14%)

Less common prompts included a path or outline for the user to follow with his finger (9%) or a demo by a character (3%). As we were confident that children 5 and under would largely be unable to read, we chose not to assess whether text instructions would be useful. We used the other common prompts as the foundation of our investigation.

Thus we assessed three in-app prompting techniques: **audio** instructions, **visual** state changes, and a demonstration by a cartoon **hand.** Based on literature suggesting that an adult model is a highly productive mechanism for prompting child behavior [21], we also chose to evaluate these in-app techniques relative to in-person support from an experienced adult user.

3.2 Hypotheses

Based on the literature above we hypothesized that:

H1: Children will be more successful learning from an adult model than from any in-app prompt

H2: Children's success in interpreting prompts will increase with age across the preschool years

Due to the lack of prior work examining the effectiveness of different types of in-app prompts, we made no predictions about their comparative value for this age group.

4. METHODS

We assessed children's ability to perform and understand four different types of gestures in response to prompting. We used four different types of prompts to indicate which gesture to perform. All gestures could be paired with all prompt types, and each child saw all four unique gestures and all four unique prompts (see Table 1). The prompt-gesture combinations varied across participants, as any type of prompt could be paired with any type of gesture. In this section, we describe these prompting techniques, how we used them to teach each gesture, and our method of assessing participants on their ability to successfully interpret each of these types of prompts.

4.1 Participants and Study Site

We conducted this investigation at a private preschool in the city of Seattle for children between the ages of 1 and 5. School administrators sent a solicitation email to the families of all students. Parents of 41 children enrolled their child in the study. Two children under 2 were excluded due to limited language and inability to follow one-step directions. Of the remaining 39 children, 5 declined to participate when asked in person for their assent. A total of 34 children (35% male) between the ages of 2 and 5 (mean = 3;7[1], sd = 0;11) participated in this study. Our sample

[1] We follow traditional linguistic notation where age is reported in yy;mm format (e.g., 2;10 represents 2 years and 10 months) [6]

Table 1: Gestures and prompts seen by each child. Each gesture can be paired with each type of prompt.

Gestures
Double Tap
Horizontal Swipe
Shake the iPad
Vertical Swipe

Prompts
Audio
Hand
Model
Visual

included four sibling pairs of different ages and one pair of identical twins.

We asked parents by email to report their child's past experience with touchscreen technology. Of the 34 children who participated, parents of 25 (74%) responded, as shown in Table 2. All parents who volunteered their child's participation were given a $5 gift certificate to Amazon as a token of appreciation.

Table 2: Participants' Prior Touchscreen Experience

How often does your child use a touchscreen device?	
Never	3%
Less than once a month	6%
Less than once a week	21%
A few times a week	26%
Every day	12%
More than once a day	6%
No response	*26%*

4.2 Materials

We designed and implemented a four-scene task to assess children's ability to correctly interpret prompts to perform specific gestures. We also assessed participants on two additional tasks not analyzed in this paper. All tasks were developed for iOS using the Cocos 2D animation library [13]. The tasks were run on an iPad 2.

In each of the four scenes, exactly one gesture triggers an on-screen event (see Figure 1). An on-screen dog named "Luna" is the focus of each scene, and executing the correct behavior causes Luna to perform a specific action. In one scene, a vertical swipe causes a dog biscuit to hop out of a treat jar and land in Luna's bowl. In another, shaking the iPad causes Luna to bark. In a third scene, a horizontal swipe causes Luna to run across a field of grass and pick up a ball, and in a fourth scene, double-tapping Luna causes her to pant and wag her tail.

We specifically chose gestures that were relatively uncommon in the applications we evaluated in our initial app review to reduce the likelihood of a child already knowing to try such a gesture based on experience. We also chose gestures that would be difficult for a child to perform unintentionally before he or she had a chance to observe the prompt (e.g., we specifically did not choose a simple tap or flick gesture). We were also careful to choose gestures that would still be easy for a young child to perform once they knew what to do (e.g., we eliminated complex gestures such as tapping or dragging with two fingers) [4].

We used four different techniques to prompt the user to perform each gesture (for a total of 16 possible prompt-gesture pairs). All participants saw all four unique scenes and four unique types of prompts. Each gesture could be prompted by any of the following:

1) *Audio Prompt:* An mp3 audio clip tells the child what gesture to perform, such as, "Shake the iPad to see what happens" in a neutral, adult female voice
2) *Hand Prompt:* A cartoon hand demonstrates the gesture on screen
3) *Model Prompt:* A script for the researcher to follow to model the target gesture (the only prompt that did not come from the app), such as saying "Ok, now do this" and performing a horizontal swipe with one finger without quite touching the surface of the iPad
4) *Visual Prompt*: A visual change to the interactive item on screen showing the intended motion

In all scenes, a text label in the top-left corner named the predetermined prompt type for the benefit of the researcher. Tapping the label triggered the prompt. In the case of the model prompt, the label was not interactive and instead named the gesture that the researcher should model (e.g., "Model Vertical Swipe"). The remaining three in-app prompts presented the following content:

Audio Prompts

The following mp3 recordings prompted each gesture:

- **Double Tap**: "Tap the dog two times to see what happens."
- **Horizontal Swipe**: "Move your finger across the screen to see what happens."
- **Shake**: "Shake the iPad to see what happens."
- **Vertical Swipe**: "Move your finger from the bottom of the screen to the top of the screen to see what happens."

Hand Prompts (see Figure 2)

- **Double Tap**: A cartoon hand with extended index finger fades into view over the dog, pauses, animates to a smaller size (to give the appearance of moving closer to the dog and farther from the user) and a pink dot appears momentarily under the tip of the finger when it has reached its smallest size (to indicate making contact with the dog). The hand animates back to its original size, and the shrinking animation is repeated to indicate a second tap. The hand again animates back to its original size and fades out of view.
- **Horizontal Swipe**: A cartoon hand with extended index finger fades into view above the dog, pauses, then animates horizontally across the screen, pauses, and fades out.
- **Shake**: A cartoon image of an iPad with a scene identical to the one currently shown on the iPad fades into view. Two hands are gripping the sides of the iPad. After a pause, the image tilts repeatedly from side to side to give the appearance of the cartoon iPad shaking back and forth.

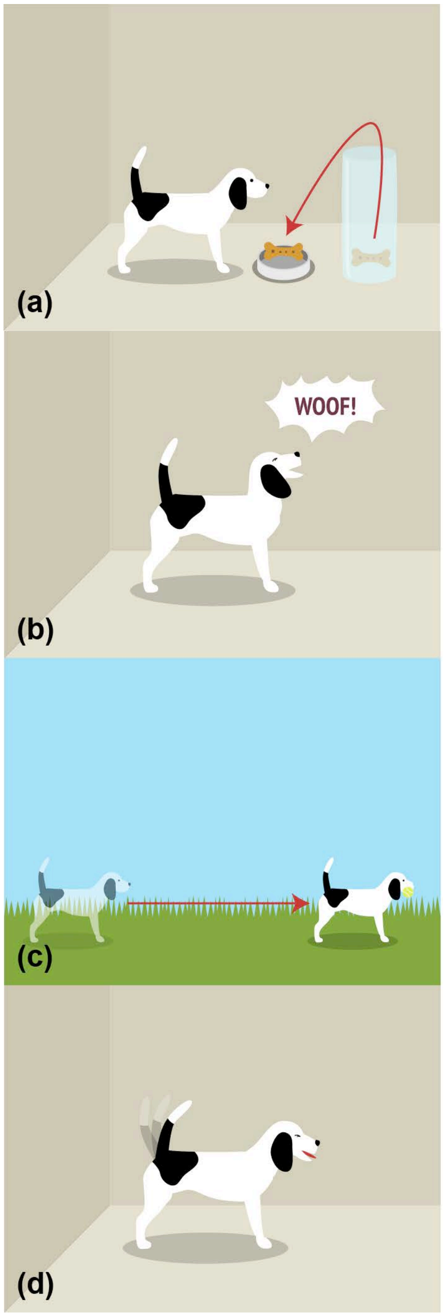

Figure 1: Scenes. All children saw all four scenes, each exactly once, each paired with a different type of prompt.

(a) Treat hops out of treat jar in response to vertical swipe, (b) Luna barks in response to shaking the iPad, (c) Luna runs across the screen and picks up her ball in response to horizontal swipe, (d) Luna hangs out her tongue and wags her tail when double-tapped.

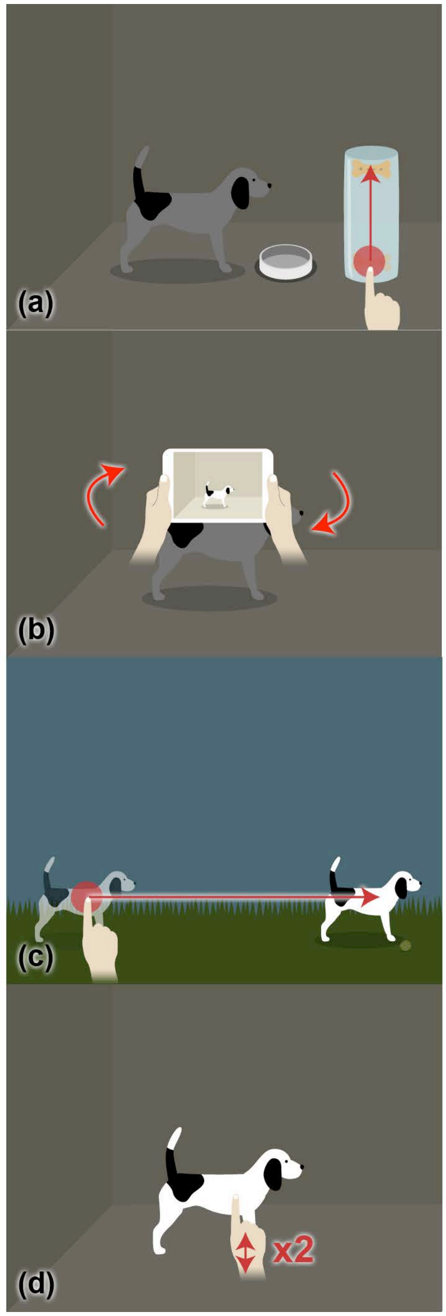

Figure 2: Hand prompts, (a) vertical swipe, (b) shake, (c) horizontal swipe, and (d) double tap. Red arrows are annotations to show interaction and were not visible to users.

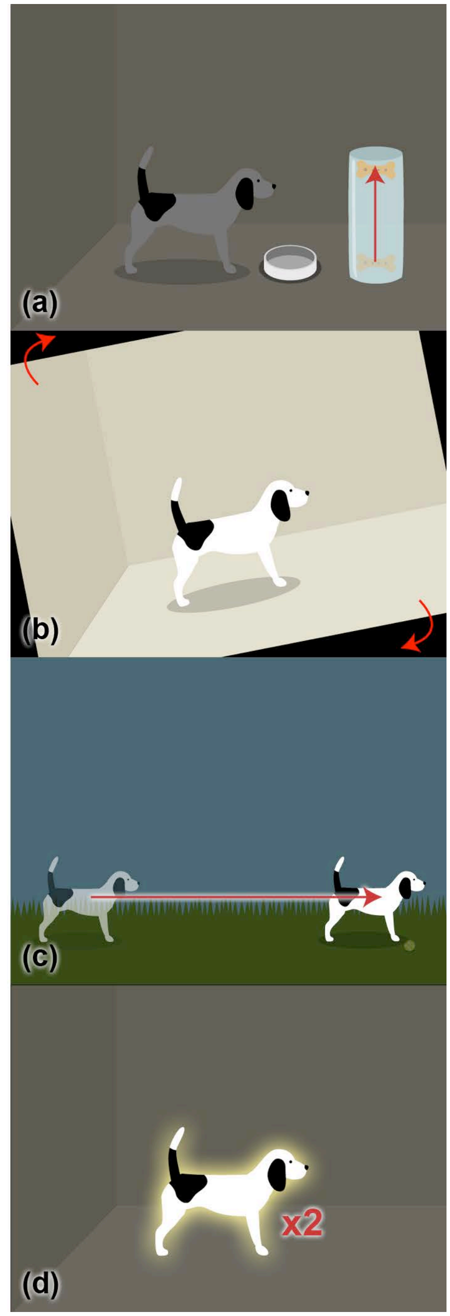

Figure 3: Visual prompts, (a) vertical swipe, (b) shake, (c) horizontal swipe, and (d) double tap. Red arrows are annotations to show interaction and were not visible to users

- **Vertical Swipe**: A cartoon hand with extended index finger fades into view over the dog biscuit, pauses, moves to the top of the treat jar, pauses, and fades out.

Visual Prompts (see Figure 3)

- **Double Tap**: The dog "pulses" twice in rapid succession via a yellow glow animation.
- **Horizontal Swipe**: The dog animates horizontally across the screen towards a ball on the other side; just before reaching it, the dog snaps back to her original position.
- **Shake**: The entire scene (including the background) rotates quickly by 30 degrees back and forth in each direction.
- **Vertical Swipe**: The dog biscuit "hops" to the top of the treat jar by animating vertically, pauses, and falls back down to the bottom.

4.3 Procedures

Prompts were assigned to gestures using a Latin Square design, ensuring that each participant saw each gesture (double-tap, horizontal swipe, shake, vertical swipe) and each prompt (audio, hand, model, visual) exactly once, with prompt order and gesture order counter-balanced across participants.

All data were collected at school during the school day over a one-week period in December 2014. Data were collected during periods of free play in order to avoid disrupting structured parts of the school day (such as meals or naps). A researcher asked one child at a time if he or she would like to "be a helper" by "playing some games on a little computer." If the child responded affirmatively, the researcher escorted him or her to a nearby office (away from the distractions of the classroom). The researcher showed the child the iPad and asked if he or she had used one before. The researcher then supported the child in playing a warm-up app where the child could draw shapes on the iPad and pop them by tapping them. Once the child was successfully performing touch interactions with the warm-up app, the researcher moved on to the experimental procedures. All participants were able to easily navigate the warm-up exercise and were creating and "popping" shapes within a few minutes of exposure.

After the warm-up exercise, the researcher introduced the child to "Luna." She explained that Luna can do tricks, and that the child can figure out how to make her do these tricks. The researcher sequentially presented each scene in the task, with scene-order predetermined and automatically populated based on participant ID. In each scene, the researcher silently pressed the label to initiate the prompt; in the case of the model prompt, the researcher instead said "Do this" while modeling the correct gesture by hovering just over the iPad screen without touching it. The researcher then gave the child the opportunity to experiment with the screen, asking "What should we do?" The child then had the freedom to perform gesture(s) or other touch interactions.

Participants were not trained in any way or told what types of gestures to perform. At no point did the researcher suggest what to do (other than triggering the appropriate prompt). Participants only knew what gesture to perform by interpreting the prompt, performing trial-and-error, or possibly drawing on prior experience with other applications.

If the child asked the researcher for assistance, said that he or she was unsure what to do, or stopped experimenting, the researcher asked the child if he or she would like to try anything else. After the second time the child stopped experimenting with the iPad, the researcher asked the child if he or she was ready to move on to the next scene. If, at any point, the child performed a gesture successfully the researcher praised the child and asked first "How did you make that happen?" If the child described what he or she did (even if it was not an accurate description), the researcher then asked, "How did you know how to do that?" Children typically completed all four scenes in less than five minutes.

After the child completed all tasks, the researcher thanked the child for his or her help and escorted him or her back to class. A second researcher was present for the duration of the session and took notes. All sessions were audio- and video-recorded.

4.4 Data Analysis

For each scene that each child saw, we coded:

- Whether the child understood what gesture to perform
- Whether the child successfully performed the gesture in a way that was correctly interpreted by the application

To determine whether a child understood what to do, we examined whether his or her action appeared deliberate, whether he or she paused after performing the gesture to observe its effect, any comments the child made spontaneously, and the child's responses to our follow-up questions about the action they performed.

Coding was performed via video analysis of session recordings. Two researchers each coded half of the data, spot-checking each other's codes for agreement. A third researcher formally assessed interrater reliability by independently coding a randomly selected 20% of all data. Cohen's κ was .926. Disagreements were discussed until consensus was reached.

5. RESULTS

5.1 Children's Understanding by Gesture

Children understood and performed all four gestures (double tap, horizontal swipe, shake, or vertical swipe) equally well. Repeated-measures ANOVAs revealed no significant differences in understanding or in performance based on gesture type. There were also no significant differences based on prior experience with touch screen technology.

5.2 Children's Understanding by Prompt

However, the mechanism by which a gesture was prompted (audio, hand, model, or visual) significantly predicted whether or not a child understood what to do (see Figure 4). A repeated-measures ANOVA revealed a highly significant effect of prompt type (F(3,

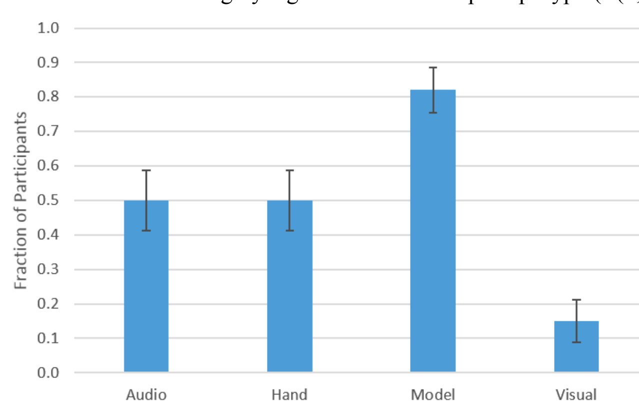

Figure 4: Fraction of participants who understood what gesture to perform as a function of prompt type

30) = 16.32, p < .001, η^2 = .338). Pairwise comparisons between prompt types (see Table 3) revealed that children were significantly more likely to understand what to do when the gesture was modeled by the researcher than when it was prompted by any in-app mechanism. Children were significantly less likely to understand what to do when prompted by a visual change to the item to touch. A Bonferroni correction was applied to all pairwise comparisons.

Table 3: Pairwise comparisons of understanding by prompt

Prompt Type		Mean Difference	Std. Error	p
Audio	Hand	.000	.087	1.000
	Model	.333*	.104	.018
	Visual	.364*	.105	.009
Hand	Model	.333*	.112	.034
	Visual	.364*	.096	.004
Model	Visual	.697*	.092	< .001

**Significant difference between two prompts*

5.3 Children's Understanding by Age

The prompt-dependent differences in children's understanding persisted after controlling for age. A repeated-measures ANCOVA of understanding by prompt type with age as a covariate produced the same pattern of results (F(3, 30) = 7.14, p < .001, η^2 = .187): the model prompt remained the most effective teaching tool, while the visual prompt remained the least effective. The same test also revealed a significant interaction effect between prompt type and age (F(3, 30) = 6.53, p < .001, η^2 = .174). We examined this interaction effect further with point-biserial Pearson product correlations of age and prompt type. These revealed that our participants became significantly more likely to understand the audio and hand prompts as they got older, while their understanding of the model and visual prompts did not change with age (see Table 4).

Table 4: Correlation between Age and Prompt Understanding

		Audio	Hand	Model	Visual
Age	r	.708**	.644**	.193	.127
	p	.000	.000	.273	.480

These results were consistent with group differences in age based on understanding. Independent-samples t-tests of children who did and did not understand what to do in response to each type of prompt revealed that the children who understood the hand prompt were significantly older (mean = 4;2, sd = 0;7) than the children who did not understand the hand prompt (mean = 3;0, sd = 0;10, t = -4.77, p < .001). Similarly, the children who understood the audio prompt were significantly older (mean = 4;3, sd = 0;6) than the children who did not understand the audio prompt (mean = 2;11, sd = 0;9, t = -5.68, p < .001). There were no significant differences in the ages of those who did (mean = 3;8, sd = 0;11) and did not (mean = 3;2, sd = 1;1) understand the model prompt, as children were relatively successful in following model prompts regardless of age. There were no significant differences in the ages of those who did (mean = 3;11, sd = 0;11) and did not (mean = 3;7, sd = 0;11) understand the visual prompt, as all children were relatively unsuccessful in following visual prompts.

Plotting the fraction of participants who understood each prompt type as a function of age further corroborated these analyses. Figure 5 shows these effects visually. Even the youngest participants were able to reliably interpret model prompts and this capacity remained high across this age range. However, participants less than 2-and-half years old were almost entirely unable to follow any of the in-app prompts, and it was only at age 3 that participants began to interpret these accurately with any regularity (though still with only 30% success). Children acquired the ability to interpret the hand and audio prompts gradually across this age range. At age 3, less than a third of participants could use these prompts effectively, but by 3-and-a-half roughly two-thirds of participants were interpreting them successfully. Across our sample, the visual prompts were ineffective for children of all ages. Prior touchscreen experiences was not correlated with age, nor was it correlated with understanding of any of the four prompt types.

5.4 Executing on Understanding

We compared participants' understanding of the gestures to their ability to execute these gestures and found that children were highly effective in translating their understanding into action. A paired samples t-test revealed no significant differences between the frequency with which children understood what to do and the frequency with which they did it successfully. This was true for each gesture type, for each prompt type, and for the collection of scenes as a whole. Across all scenes for all children, participants who understood what gesture to attempt successfully performed the gesture in a way that was understood by the system 87% of the time. A repeated-measures ANCOVA comparing the number of gestures a participant understood to the number of gestures he or she successfully performed with age as a covariate revealed no significant interaction between age and performance measure.

6. DISCUSSION

Our results reveal that even very young children can be successful in performing common gestures. They also reveal that, for these common and relatively simple motions, understanding which gesture will be productive – and not the physical demands of the gesture itself – is a hurdle to children's success. While much of the limited prior work in this space has assessed children's fine motor control and gesture performance, our findings point to a need for increased focus on designing prompts to guide children toward contextually useful gestures. Once children understand which gestures will be productive in a given scene, they are highly successful in executing them.

Our results confirmed both of our hypotheses, specifically that in-

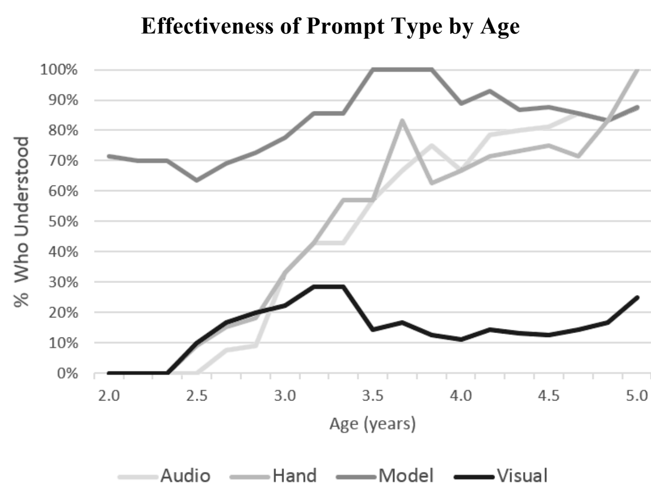

Figure 5: Percentage of participants who understood each type of prompt at each age. For smoothing, percentages represent a sliding window of six months on either side of the target age.

person modeling is at least as effective as any in-app prompting mechanism and that children's success in working with prompts increases with age across the preschool years. We further found that children's ability to respond appropriately to a prompt is a function of both the prompt type and the child's age. Here we discuss our participants' ability to work with each of the four prompt types we assessed.

6.1 Model Prompts

Our results show that in-person modeling is a more productive teaching technique for this population than any of the most common in-app prompt types. However, the gains from modeling are greatest for younger children and, by the end of the preschool years, in-app techniques begin to rival its effectiveness. These results suggest that while adult guidance is a valuable mechanism for supporting children in engaging with tablet content across this age range, by the age of 5 other prompting techniques become accessible and an adult demonstration may not be essential.

6.2 Audio Prompts

Surprisingly, none of the common in-app prompting mechanisms we assessed were effective instructional techniques for children under 3. We were surprised to find that audio prompts were ineffective for this age group, despite the fact that these participants could follow simple verbal directions and converse easily with the research team. As our audio prompts were short and intentionally spoken slowly and clearly, we expected them to be accessible to young children.

However, adults and older children consistently, automatically, and unconsciously employ *child-directed speech* when speaking to children under 4 [11]. Child-directed speech is a way of speaking that is tailored with precision to a very young child's limited verbal ability and involves: expanding the pitch contour, adding extended pauses between words and statements, reframing and repeating target words and phrases, and eliminating disfluencies. Given the sophistication and degree of tailoring in adults' statements to young children, it is possible that pre-recorded audio content may lack an element of in-person verbal communication that is essential to reaching very young children.

However, audio prompts quickly became highly effective teaching tools as children grew older. Older children were as likely to follow gesture instructions from an in-app voice-recording as they were an in-person model. Our review of commercially available apps for preschoolers indicates that this is the most common prompting technique for content targeting this user group and is used by nearly half of all apps to provide instructions about where and how to touch the screen. Our results indicate that this is largely appropriate and that app developers can expect children age 4 and up to use these verbal instructions effectively. However, many of these application authors might be surprised to learn that their prompts are ineffective for 3-year-olds. Further work is needed to understand if and under what conditions in-app audio instructions successfully prompt these younger preschoolers to perform the contextually appropriate gestures.

6.3 Hand Prompts

Hand demos were also inaccessible to children under 3. As children in this age range are just beginning to acquire the capacity for symbolic representation [15], it is possible that they are unable to mentally represent a cartoon hand as a symbol referencing their own hand, and that they therefore fail to map on-screen actions to actions of their own. For example, in an unrelated study, researchers placed stickers on children's bodies and asked them to identify the corresponding spot on a doll, a symbolic representation of the child's own body [16]. Because very young children struggle to form simultaneous mental representations of a symbolic object as both an object in its own right and a stand-in for its referent [15], study participants were unable to map between the doll's body and their own. One explanation for our younger participants' failure to imitate the actions of a cartoon hand could be an inability to interpret this as a representation. If a child does not recognize that an on-screen hand is a stand-in for the user's hand, he or she would have no reason to interpret the demo as an instruction. Further work is needed to understand whether children's challenges in using these prompts stem from an immature capacity for symbolic thinking.

Despite the difficulties of very young children, the hand prompts, like audio prompts, became steadily more accessible as children grew older, with dramatic acquisition between age 3 (when less than a third of children could interpret these prompts) and 3-and-a-half (when more than two-thirds of children could interpret these prompts). These results suggest that designers targeting older preschoolers can expect their users to interpret hand and audio prompts productively without adult facilitation. They also suggest that designers targeting toddlers and younger preschoolers cannot expect their users to understand traditional prompts.

6.4 Visual Prompts

Finally, we were surprised to find that visual state changes to interactive items were uninterpretable to our participants and that this remained true even for the oldest children. Prior work demonstrates that *within-stimulus prompting* (changes to the stimulus to which a child should attend, such as increasing the size of the number '5') can be more effective than *extra-stimulus prompts* (changes which are external to the stimulus to which a child should attend, such as underlining the number '5') in early-education settings [32]. Thus we expected these visual prompts to be more effective than an external stimulus, such as a cartoon hand, and were surprised to find that this was not the case.

It remains likely that these visual prompts are effective in drawing children's attention to interactive items, as these visual-state changes leverage users' orienting response and would be hard to miss [17]. And it is also possible that this pulsing, glowing, and movement make these items seem more enticing or touchable. However, we found that they were unable to serve as teaching tool for demonstrating with precision how users should touch these on-screen items. Further work remains to determine at what age basic visual-state changes can successfully communicate swipe-ability or the need to double-tap, and whether within-stimulus visual changes are able to teach gestures to users of any age.

We found these prompts to be the most difficult to design and, relative to audio content or hand demos, the visual prompts we selected were drawn from a larger set of possibilities. It is possible that with evidence-based guidelines for creating visual prompts, designers could develop solutions that work well for preschoolers. However, as our prompts were very similar to the prompts we saw in our app review, today's visual prompts are likely to be ineffective for teaching young children to swipe, double-tap, or perform other simple gestures. Visual prompts were common in the apps we reviewed, occurring in more than 40% of these titles. Our results suggest that these widespread techniques are likely unproductive and could benefit from redesign. Future work to define at what age, if at all, these prompts become interpretable and the features of visual prompts that predict success would be valuable.

6.5 Future Work

There is a clear opportunity for the research community to innovate novel prompting mechanisms for very young children. While joint media engagement with an adult is recommended for preschoolers' using tablets [31], these devices are also used by children independently, and this trend is likely to continue as parents view young children's screen time as a useful tool for keeping children occupied [18]. New techniques for guiding young children toward meaningful patterns of interaction would increase the chances that children can access educational content and engage with an app productively when playing alone. Our results point to avenues for investigating how best to refine prompt designs for each of these prompt types.

As we found a lag between children's acquisition of conversational language and their ability to understand audio prompts, there is a clear opportunity to examine the source of this gap. Future work might examine whether tailored audio prompts which incorporate elements of child-directed speech and adjust to support children's age and verbal ability increase children's success and lower the age at which audio prompts become useful. A systematic evaluation of the differences between adults' in-person verbal instructions and pre-recorded in-app audio instructions could also illuminate reasons why children under 3 struggle to use app audio prompts. Further, comparing identical audio content delivered by a computer and by a human to children at a fixed age could help isolate whether young children struggle with the audio content itself or the digital delivery mechanism.

As we found modeling to be highly effective with preschoolers of all ages, it would be useful to identify the components of adult modeling that facilitate learning and extrapolate opportunities, if any, for translating these components into in-app supports. In this study, participants were required to interpret prompts they had never seen before and to interpret them without assistance. It is possible that with adult facilitation to interpret these prompts initially, children might be more likely to interpret them correctly in subsequent instances. Future work remains to determine whether the benefits of an adult model are sustained and whether an initial teaching session enables very young children to respond correctly and independently to subsequent audio or hand prompts.

Further work is also needed to understand the benefits and drawbacks of combining prompts, as these may be more than the sum of their parts. Using a visual prompt to call attention to a target item while describing a gesture with an audio prompt may be more effective than using either prompt in isolation. By leveraging the power of visual prompts to harness the user's attention and the power of audio prompts to add precision, it is possible that the designer would create a more effective teaching tool. Alternatively, it is also possible that simultaneous multi-modal prompts would be distracting or require the user to synthesize information in a way that is challenging for young children.

Finally, our laboratory-style study is limited and may not reflect children's responses to prompts in more natural contexts. In daily life, children may be more experimental, may attempt more gestures without certainty, and may be more likely to discover correct gestures in absence of useful prompts. With repeated exposure over an extended period of time, not possible in this context, children may be able to interpret prompts with greater accuracy. Or other cues or pure experimentation may result in children discovering productive gestures without interpreting prompts. Further work is needed to understand how our findings translate to more natural usage scenarios. Further work is also needed to understand whether our findings generalize to more complex gestures.

7. CONCLUSION

Though it is well-established that digital media can confer learning benefits, children's access to these gains is gated by their ability to interact meaningfully with the content. Though the popular television series *Mister Rogers' Neighborhood* increases young children's empathy and supports emotional learning [12], these messages are lost on children under 2 who are unable to distinguish between normal and incomprehensibly scrambled videos [26]. Understanding children's ability to interpret digital content and the way in which this changes with age is essential to creating media experiences that are beneficial [27].

In this study, we found that the way in which children are prompted to interact with tablet applications directly impacts their understanding and their ability to engage with content. Our results show that children under 3 are unable to interpret these prompts and should be guided by an adult model. They also show that children make rapid gains in understanding between the ages of 3 and 3-and-a-half, at which point they are moderately successful in interpreting on-screen demos and audio instructions. Common visual-state-change prompts appear to be ineffective instructional tools across this age range. These findings point to the critical importance of designing developmentally appropriate prompts when creating interfaces for very young children and define several means of doing so.

8. ACKNOWLEDGMENTS

We would like to thank Chris Peloquin who assisted with study design and data collection. We would also like to thank the Seattle preschool where we conducted this research and all of our participants and their families. The Institutional Review Board at the University of Washington approved this research.

9. REFERENCES

[1] Anthony, L. et al. (2012). Interaction and recognition challenges in interpreting children's touch and gesture input on mobile devices. *ITS'12* (225).

[2] Arif, A.S. and Sylla, C. (2013). A comparative evaluation of touch and pen gestures for adult and child users. *IDC'13* (392–395).

[3] Aziz, N. A. A. et al. (2013). Selection of touch gestures for children's applications. *SAI'13* (721–726).

[4] Aziz, N.A.A. (2013). Children's interaction with tablet applications: Gestures and interface design. *Children*. *2*(3), 447-450.

[5] Baloian, N. et al. (2013). Tablet gestures as a motivating factor for learning. *ChileCHI'13* (98–103).

[6] Baron, N.S. (1993). *Growing up with language: How children learn to talk*. Da Capo Press.

[7] Berkowitz, S. (1990). A comparison of two methods of prompting in training discrimination of communication book pictures by autistic students. *J Autism Dev Disord*. *20*(2), 255–262.

[8] Bragdon, A. et al. (2011). Experimental analysis of touch-screen gesture designs in mobile environments. *CHI'11* (403–412).

[9] Brown, Q. et al. (2010). Clear Panels: a technique to design mobile application interactivity. *DIS'10* (360–363).

[10] Brown, Q. and Anthony, L. (2012). Toward comparing the touchscreen interaction patterns of kids and adults. *SIGCHI EIST Workshop 2012*.

[11] Clark, E. V (2009). *First language acquisition*. Cambridge University Press.

[12] Coates, B. et al. (1976). The influence of "Sesame Street" and "Mister Rogers' Neighborhood" on children's social behavior in preschool. *Child Dev*. *47*, 138–144.

[13] Cocos2d-x: World's #1 Open-Source Game Development Platform: *http://www.cocos2d-x.org/*. Accessed: 2015-04-02.

[14] Csapo, M. (1981). Comparison of Two Prompting Procedures to Increase Response Fluency among Severely Handicapped Learners. *Journal of the Association for the Severely Handicapped*. *6*(1), 39–47.

[15] DeLoache, J.S. (2004). Becoming symbol-minded. *Trends in Cogn Sci*. *8*(2), 66-70.

[16] DeLoache, J.S. and Marzolf, D.P. (1995). The use of dolls to interview young children: issues of symbolic representation. *J Exp Chi Psychol*. *60*(1), 155–173.

[17] Diao, F. (2004). Orienting Response and Memory for Web Advertisements: Exploring Effects of Pop-Up Window and Animation. *Comm Res*. *31*(5), 537-567.

[18] Evans, C.A. et al. (2011). Only Two Hours?: A Qualitative Study of the Challenges Parents Perceive in Restricting Child Television Time. *J Fam Issues*. *32*(9), 1223-1244.

[19] Hilsen, L. (2013). *A Step-by-step ABA Curriculum for Young Learners with Autism Spectrum Disorders (age 3-10)*. Jessica Kingsley Publishers.

[20] Kammer, D. et al. (2014). Investigating interaction with tabletops in kindergarten environments. *IDC'14* (57–66).

[21] MacDuff, G.S. et al. (2001). Prompts and prompt-fading strategies for people with autism. *Making a difference: Behavioral intervention for autism*. 37–50.

[22] Mcknight, L. and Cassidy, B. (2010). Children's Interaction with Mobile Touch- Screen Devices : Experiences and Guidelines for Design. *Int Journal of Mobile HCI*. *2*(2), 1–18.

[23] McKnight, L. and Fitton, D. (2010). Touch-screen Technology for Children : Giving the Right Instructions and Getting the Right Responses. *IDC'10* (238–241).

[24] Nacher, V. et al. 2014. Exploring Visual Cues for Intuitive Communicability of Touch Gestures to Pre-kindergarten Children. *ITS'14* (159–162).

[25] Oh, U. and Findlater, L. (2013). The Challenges and Potential of End-User Gesture Customization. *CHI'13* (1129–1138).

[26] Pempek, T.A. et al. (2010). Video comprehensibility and attention in very young children. *Dev Psychol*. *46*(5), 1283–1293.

[27] Radesky, J.S. et al. (2015). Mobile and Interactive Media Use by Young Children: The Good, the Bad, and the Unknown. *Pediatrics*. *135*(1), 1–3.

[28] Rahman, M.S.A. et al. (2013). A Study on the Naturalness of Gesture-Based Interaction for Children. *Advances in Visual Informatics*. Springer. 718–728.

[29] Rideout, V. (2013). Zero to eight: Children's media use in America 2013. *Common Sense Media*.

[30] Rust, K. et al. (2014). Understanding childdefined gestures and children's mental models for touchscreen tabletop interaction. *IDC'14* (201–204).

[31] Takeuchi, L. et al. (2011). The new coviewing: Designing for learning through joint media engagement. *New York, NY: The Joan Ganz Cooney Center at Sesame Workshop*.

[32] VanDerHeyden, A.M. et al. (2002). Comparison of within-stimulus and extra-stimulus prompts to increase targeted play behaviors in an inclusive early intervention program. *The Behavior Analyst Today*. *3*(2), 188.

[33] Vatavu, R. D. et al. (2015). Touch interaction for children aged 3 to 6 years: Experimental findings and relationship to motor skills. *Int J Hum-Comput St*. *74*, 54–76.

[34] Wigdor, D. and Wixon, D. (2011). *Brave NUI World: Designing Natural User Interfaces for Touch and Gesture*. Elsevier.