Characterizing How Interface Complexity Affects Children's Touchscreen Interactions

Julia Woodward[1], Alex Shaw[1], Annie Luc[1], Brittany Craig[2†], Juthika Das[1], Phillip Hall Jr.[1], Akshay Holla[3‡], Germaine Irwin[4], Danielle Sikich[5†], Quincy Brown[6], Lisa Anthony[1]

[1]Dept of CISE, University of Florida, FL, USA
[2]St. Catherine University, Minneapolis-St. Paul, MN, USA
[3]User Testing, Inc., Mountain View, CA USA
[4]University of Maryland Baltimore County, Baltimore, MD USA
[5]Western Oregon University, Monmouth, OR USA
[6]Bowie State University, Bowie, MD USA

julia.woodward@ufl.edu, alexshaw@ufl.edu, annieluc@ufl.edu, bmcraig@stkate.edu, juthika@ufl.edu, phillipjr@ufl.edu, holla.akshay@gmail.com, germaine.irwin@umbc.edu, dsikich12@mail.wou.edu, qbrown@bowiestate.edu, lanthony@cise.ufl.edu

† Work conducted while these authors were summer interns at the University of Florida.
‡ Work conducted while this author was a student at the University of Florida.

ABSTRACT
Most touchscreen devices are not designed specifically with children in mind, and their interfaces often do not optimize interaction for children. Prior work on children and touchscreen interaction has found important patterns, but has only focused on simplified, isolated interactions, whereas most interfaces are more visually complex. We examine how interface complexity might impact children's touchscreen interactions. We collected touch and gesture data from 30 adults and 30 children (ages 5 to 10) to look for similarities, differences, and effects of interface complexity. Interface complexity affected some touch interactions, primarily related to visual salience, and it did not affect gesture recognition. We also report general differences between children and adults. We provide design recommendations that support the design of touchscreen interfaces specifically tailored towards children of this age.

Author Keywords
Touch interaction; gesture interaction; child computer interaction; gesture recognition; mobile devices.

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION
Because most touchscreen devices are not specifically designed for children, their interfaces often do not optimize interaction for children. Children are an important group to consider when designing for touchscreens because of the widespread use of this technology for education, games, and entertainment targeted at younger users [13,32,42,43]. Previous research shows that existing gesture recognition algorithms designed for adults do not function as well for children [2–4], and children also have more difficulty with touch interactions compared to adults, due to the size and location of touch targets, as well as other application-specific considerations [2,4]. Children are also still developing key motor skills, which affects their touch and gesture interactions when compared to adults [44].

CHI'16, May 07-12, 2016, San Jose, CA, USA
© 2016 ACM. ISBN 978-1-4503-3362-7/16/05…$15.00
DOI: http://dx.doi.org/10.1145/2858036.2858200.

Figure 1. Children participating in our study.

Prior work in three studies with children ages 7 to 17 by Anthony et al. [2–4] inspired our work. However, those studies focused on isolated interactions, such as touching simple targets [2–4], while most real world applications are more complex, with more onscreen interface elements, animations, sounds, feedback, and functionality. Previous work from child psychology literature has found that children's response times are impacted differently by changing visual stimulus than adults'; for example, children (ages 6 to 10) are slower to touch a target in a stimulus-response task when orientation is changed rather than color [15]. What remains an open question is how children's behavior might be affected by other distractors resulting from varying interface complexity. In this study, we focus on how the added complexity of more realistic interfaces might affect children's interactions. In addition, we extend previous work by conducting our study with even younger children, ages 5 to 10. Children of this age range are of particular interest because of the rapid cognitive and motor skills development that occurs during this period [34,44]. We compare touch and gesture interaction data from 30 children and 30 adults collected in two different types of interfaces: (1) a simple, stimulus-response version, and (2) a more complex, realistic version. The added complexity that we introduce into the interface arises from colorful animations, progress indicators and feedback, and a narrative, game-oriented purpose for the interactions.

We analyzed the touchscreen interaction logs and performed gesture recognition experiments on the gesture data. Interface complexity affected some of the touch interactions, primarily related to visual salience, and it did not affect gesture recognition. We also find general differences between children and adults. The findings from our study are two-fold: we first replicate key aspects of Anthony et al.'s studies [2,4] such as miss rate affected by age and target size, but we also go beyond prior work to investigate the effect of interface complexity for target and gesture interactions. The contributions of this work are (1) collection of a touch and gesture corpus of interaction data from children as young as 5 years old; (2) analysis of the effect of interface complexity on touch and gesture interaction in both children and adults; and (3) design implications for touch and gesture interaction for children. The findings from this study enable the design of better touchscreen interfaces specifically tailored toward children.

RELATED WORK

Most previous work focusing on users' interactions with touchscreens has used adults [14,17,18,33,37,47–49]. Some studies have looked at the difference in interaction patterns between children and adults, but have not examined differences across various ages [7,11,12,21,31,35]. Only a few studies have looked at differences between younger and older children [4,28,46]. We focus our review of prior work on four major categories: effect of interface complexity on interaction patterns, touch interactions for children, gesture interactions for children, and gesture recognition.

Interface Complexity

A limitation of prior work on children's touchscreen interactions is that the studies have used very simple apps, designed only to elicit the basic data necessary to study interaction patterns [2–4,7]. These apps can be compared to classic psychology "stimulus-response" tasks [16]. These apps, however, do not resemble real world applications that children might actually use in their daily lives. While these previous studies offer a solid foundation for understanding the ways in which children use touchscreens, we cannot be sure that these results are generalizable to more complex, realistic apps without studying the effect of the increased complexity. We do that in this study by comparing children's touch interactions in a simple, abstract interface, like previous work, and a more complex interface closer to the applications children use in the real world.

Touch Interactions and Children

Previous work on children's touch interactions has focused on helping designers create better interfaces for touchscreens [2,28,36,38]. McKnight and Cassidy [28] studied the interactions of 7- to 10-year-olds with various types of mobile touchscreens, offering guidelines for the design of mobile devices intended for children. Similarly, several other studies have offered design guidelines for user interfaces for apps targeted toward children [2,36,38]. A recent study by Vatavu et al. [46] showed that children improve significantly in their ability to perform tapping interactions between the ages of 3 and 6. While these studies all used relatively simple stimulus-response apps, they have shown that, not only do children's interactions differ from those of adults, they also differ between younger and older children. However, there is still room for more work on comparing interaction patterns across ages at different points of development. We examine children's touch interactions in a study comparing interactions from ages 5 to 10, a time of rich development [34,44].

Gesture Interactions and Children

As with touch interactions, much of the work on children's gesture interactions has examined ways in which designers can improve children's experiences when using gestures in touchscreen interfaces [1,3,20,29,31]. Some recent studies have investigated age-appropriate methods of prompting young children to make gestures within smartphone apps [20,29]. Hiniker et al. [20] found that children responded best to prompts that were specifically designed for them rather than for adults. McKnight and Fitton [29] recommended that applications be designed to compensate for common errors children make, such as delayed timing of interactions and accidental touches during gestures. Anthony et al. [3] demonstrated the importance of providing visual feedback when eliciting gestures from children. In another study, Aziz et al. [1] studied the types of gesture interactions that were appropriate for children ages 2 to 4. Nacher et al. [31] explored multi-touch gesture interaction for children of ages 2 to 3, showing that they can effectively use a variety of common gestures and recommending guidelines to compensate for the challenges very young children face when making gestures. These studies point toward a need for more information about how children's gesture patterns change as they grow and develop. We add to this body of literature by analyzing children's gesture patterns from the ages of 5 to 10, and comparing them to adults. Whereas these previous studies focused on qualitative characteristics of the gestures, we consider how recognition of children's gestures is affected by how children are making gestures.

Gesture Recognition and Children

A sizable number of gesture recognition algorithms for touchscreen interfaces have been developed, but they have largely been designed for and tested solely on adults [6,19,22,25,26,39,41,45,50]. A comparison of the efficacy of some of these recognizers [6,45] for children's gestures is presented by Anthony et al. [4], but only for ages 7 to 17, and with ages clustered together in groups of three. Arif and Sylla [7] examined how adults and children make gestures using pen vs. touch input. They found that adults' gestures were better recognized when using pen rather than touch, but that children's gestures were equally poorly recognized for both input modalities. The study reported 80% accuracy for children and just under 90% accuracy for adults, well below acceptance rates established by prior work [24,35]. Kim et al. [23] examined gestures made by young children, but they were identifying children's age and development levels based on their gestures, not the challenges of recognizing children's gestures. More in-depth analysis of children's touchscreen gesture interactions and recognition is needed to provide better support for these ages. To that end, we analyze the performance of a popular multi-stroke gesture recognition algorithm, $P [45], on gestures produced by children aged 5 to 10, as well as adults.

METHOD

In our study, each participant performed two different tasks on mobile touchscreen devices: (1) touching **targets**, and (2) drawing simple **gestures** (e.g., letters, numbers, and shapes). These tasks are based on three related studies from Anthony et al. [2–4]; however, those studies only used simple touch and gesture interfaces, similar to psychology stimulus-response tasks [16]. Our study examines how interactions change when the interface is more realistic. Participants in our study used both a simple version of each task and a more complex version. Overall, each participant performed four different tasks in the same session, which lasted approximately an hour; breaks were offered between the tasks. Task order was counterbalanced across sessions.

During each session there were one or two participants, either children or adults, with two researchers in the room. If there were two children, their ages were similar. (We examined whether being paired or single affected the participants' interactions, and we found no effect. Therefore, we exclude this factor from our analyses.) Each participant was read the informed consent or assent before deciding to participate, and then asked to rank four different incentive prizes. These prizes were small inexpensive toys for children and small office supplies for adults. They were used as motivational items to encourage participants to finish all four tasks, (e.g., [9]) since the tasks might get repetitive for young children. After each task was finished, participants earned the prize for that level, and took home their highest-ranked prize if they finished all four tasks.

Equipment

The applications were run on Samsung Google Nexus S Smartphones with the Android 4.0.4 operating system. The display resolution was 480 x 800 pixels, the pixel density was approximately 234 pixels per inch, and the phones were 4.88 x 2.48 x 0.43 inches, with a 4-inch screen.

Participants

A total of 60 participants (30 adults, 30 children) participated in our study. All children were recruited from the P.K. Yonge Developmental Research School, affiliated with the University of Florida; adults were recruited from the University of Florida. Of the adults, 15 were female and 15 were male; of the children, 16 were female and 14 were male. The adults' ages ranged from 17 to 33 years (M: 23.0 yrs, SD: 3.8 yrs) while the children's ages ranged from 5 to 10 years (M: 7.7 yrs, SD: 1.6 yrs). Of the 60 participants, 47 were right-handed (78%), 7 were left-handed (12%), and 3 were ambidextrous (3 others did not specify).

Participants were asked to rank themselves as either "expert", "average", or "beginner" with touch input devices (e.g., smartphones, tablets). Participants tended to rank themselves highly: 16 (53%) of adults said "expert," 14 (47%) said "average", and 22 (73%) of children said "expert". Only 2 of the children (7%) labeled themselves as "beginners". All of the adults owned a touchscreen smartphone, while children more often had their own tablets (22 out of 30, 73%). Only 7 children (23%) said they owned their own touchscreen smartphone; the majority of the smartphones the children use are owned by a family member (22 out of 30, 73%). Table 1 shows the percentages of adults and children that said they use touchscreen devices daily. Although the interpretation of the self-rankings is subjective, these data reflect the current pervasiveness of touchscreen devices in children's lives.

Target Interactions

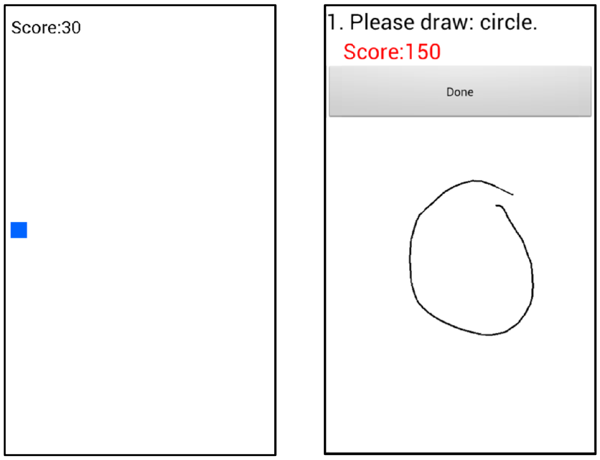
We based our target task application design on the simple application used by Anthony et al. [2,4]. We added game elements to increase interface complexity to study its effect on interaction patterns. In the simple application, the participant was asked to touch blue squares (Figure 2a). In the complex application, the participant was asked to touch fish trapped in ice cubes that melted when tapped (Figure 3). We refer to these applications as the Target Abstract (TA) and the Target Game (TG) applications.

General Design of Target Tasks

In both the TA and TG applications, the participant was asked to touch 104 targets that appeared on the screen. The design and distribution of these targets was as inspired by the previous studies [2,4]. There were four target sizes: very small (0.125 inches), small (0.25 in), medium (0.375 in),

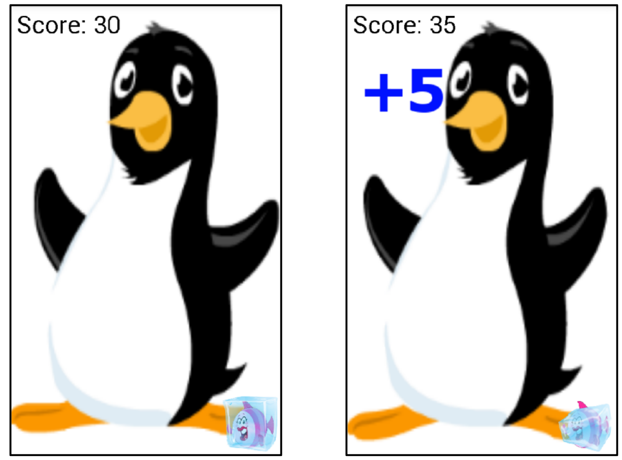
		Mobile Phone	Tablet	MP3 Player	Tablet PC
Adults	**Use it daily**	93%	30%	20%	30%
Children	**Use it daily**	27%	50%	23%	0%

Table 1. Percentage of adults and children who participated in our study and use a touchscreen device daily.



(a) (b)

Figure 2. (a) Example of interface from TA. (b) Example of interface from GA. These applications are similar to those used by Anthony et al. [2–4].



(a) (b)

Figure 3. Example of two screens from TG. (a) The target the participant has to hit (a fish trapped in a block of ice) is in the lower-left-hand corner. (b) After the participant hits the target, the score increases and the ice melts, releasing the fish.

and large (0.5 in). The phone display was divided into a 3 x 5 grid which provided 15 possible locations; only 13 locations were used for the targets, as in the previous studies [2,4]. If the participant did not touch the target within its bounds, the application would not go on to the next target until the participant did so. All of the targets were presented in the same order for all participants and both applications, and no consecutive targets were in the same location or the same size. As in the previous studies [2,4], half the targets had edge padding, meaning the target was slightly inset (10 pixels) from the side of the screen instead of aligned to the edge (Figure 2a). The participant was told that he or she was not being evaluated on how fast the activity was completed. Each participant was instructed to use the hand he or she preferred and was allowed to hold the phone however felt comfortable. (We examined whether the way the participants held the phone had an effect on their interactions; since there was no significant difference, this factor is excluded from our further analyses.)

Abstract Target Application (TA)

The "abstract" version of the target applications (TA) was a simple interface consisting of a white canvas on which blue squares would appear (Figure 2a). Before using the application, each participant was instructed to touch each blue square, and that another square would appear after they touched each one. Although the interface was simple stimulus-response, there were gamified elements as inspired by Brewer et al. [9]. After the participant hit each target, the score in the upper left hand corner incremented by five points. After the participant completed all 104 targets, a "game over" message appeared with their final score.

Gamified Target Application (TG)

The "complex" version of the target applications (TG) was a full-fledged game interface showing an animated penguin named Polly, who flipped her flippers and blinked her eyes. A fish frozen in a square block of ice appeared as the target (Figure 3a). Before using the application, the participant was instructed to touch the fish that appeared in order to melt the ice and help Polly catch the fish. After the participant touched the fish, an animation showed the ice melting and the score incremented by five points (Figure 3b). The 104 targets were broken up into six different levels of 17 or 18 targets each, to continue the game metaphor. After each level, a congratulations image appeared with an animation of fireworks. After the participant completed all six levels, another message appeared with a "game over" message and their final score. The design of this interface was inspired by several popular apps for children[1].

Gesture Interactions

We also based our gesture application design on the simple application used by Anthony et al. [2–4]. We added game elements to increase interface complexity to study its effect on interaction patterns. In the simple application, the participant was asked to make a number of gestures (Figure 2b). In the complex application, the participant was asked to make these gestures but in a game-like interface (Figure 3). We refer to these applications as the Gesture Abstract (GA) and the Gesture Game (GG) applications.

General Design of Gesture Tasks

Both the GA and GG applications used the same set of 20 gestures (Figure 4). These gestures have been chosen from prior work [2–4], and represent common letters, numbers, and symbols [8]. Participants provided six samples of each gesture in the set. Our chosen gesture set allows us to collect a representative corpus for recognition experiments, and study the challenges of gestures from younger children.

[1] - Talking Gina http://bit.ly/1jut6VA
- Moonbeeps Fireflies http://apple.co/1OA9ZEJ
- Tap the Frog http://bit.ly/1oM4Lcp
- Pop Balloon Kids http://bit.ly/1h0Rzfg

Figure 4. The gesture set we used in our study, from [2,4].

Abstract Gesture Application (GA)

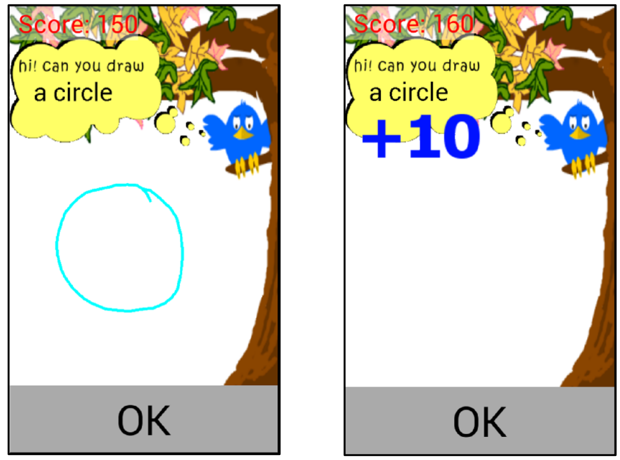
The "abstract" version of the gesture applications (GA) was a simple interface consisting of a white canvas on which the participant could draw gestures (Figure 2b). At the top of the screen, a prompt appeared instructing the participant to draw a specific gesture from the set. The participant drew the gesture and clicked an onscreen "Done" button when finished. To encourage natural interactions, participants were not allowed to erase any gestures they made, as this could have led to them trying multiple times to achieve what they believe is a more aesthetically pleasing gesture result [3]. Although the interface was a simple stimulus-response, there were gamified elements as in Brewer et al. [9]. After the participant drew each gesture, the score in the upper left hand corner incremented by 10 points. After the participant completed all six rounds of 20 gestures, a "game over" message appeared with their final score.

Gamified Gesture Application (GG)

The "complex" version of the gesture applications (GG) was a full-fledged game interface showing a forest space flanked by a tree in which an animated bird, named Billy Bluebird, would sit (Figure 5a). The participant was instructed to help Billy by drawing the gestures that he asked them to produce. Before each of the six rounds, the participant was also asked what color and line width he or she would like to use, to make the gesture drawing more fun and increase engagement. After the participant drew each gesture, an animation showed Billy flapping his wings and the score incremented (Figure 5b). After each round (or level), a congratulations image appeared with an animation of fireworks. After the participant completed all six rounds, another message appeared with a "game over" message and the final score. As with the abstract gesture application, the users could not erase their gestures. The design of this interface was inspired by several popular apps for children[2].

DATA ANALYSIS AND RESULTS

We have analyzed the data from the touch interactions in the target applications, and also conducted gesture recognition experiments on the gesture samples. In both cases, we compared the simple interface to the complex



(a) **(b)**

Figure 5. Example of two screens from GG. (a) Billy prompts the participant for the gesture to draw in a speech bubble. (b) After the participant draws the gesture, the score increases.

interface, and between children of different ages and adults. To compare the data we collected in this study to the findings from the previous studies that inspired our work [2–4], we first replicate similar analysis. We also extend the analysis, examining the effect of interface complexity, which we highlight in the section headers. We consider effect of age as a factor in our analyses as well. Table 2 summarizes the overall effects and findings we will present.

Target Interactions

To analyze the touch interaction data, we first removed eight participants whose data was incomplete due to leaving the study early or logging errors with the applications (13% of our overall dataset, or 2,509 touch events out of 18,892 total touch events). We analyzed 103 targets per person after excluding the first target as practice. Thus, we had a set of 52 participants (29 children, 23 adults) with complete data for the target tasks, consisting of 16,262 touch events. Of the children, there were three 5-year-olds, six 6-year-olds, four 7-year-olds, seven 8-year-olds, four 9-year-olds, and five 10-year-olds. Between the ages of 5 to 10, there is a marked increase in cognitive and motor development [27,34,44], and we wanted to see if there were differences

Factors / Measures	Interface Complexity	Participant Type (Adult / Child)	Age (years)	Target Size	Other Factors
Holdovers	X	X		X	
Misses		X *(overall and gutter misses)*	X *(per-user and edge padding)*	X	*main effect of edge padding*
Pressure	X	X			
Size	X	X			
Location		X			*main effect of region*
Response Time	X	X	X	X	

Table 2. Summary of target task findings. An X indicates a significant effect (p<.05) of the factor on the measure.

[2] - LetterSchool http://apple.co/1TojV6j
- iWriteWords http://bit.ly/1HAq0v5
- ABC Print Big Trace http://apple.co/1Ponuuw

in interactions between younger children and older children, so our analysis considers individual ages (5,6,7,8,9,10 years old, and adults). In some places, our analysis only considers age in general (e.g., *participant type*: child or adult) for simplicity. During the study, the applications logged every touch event and available features of that touch event (e.g., touch location, target size, touch pressure, etc.).

Holdovers (Effect of Complexity)

Holdovers occur when touches are located in the same vicinity as the previous target instead of the current target [2]. The user has not realized that the touch has already been registered. In our dataset, 283 of the 16,262 touch events were holdovers (1.7%), about the same as the percentage reported by Anthony et al. [4]. Of the holdovers in our dataset, the large majority of them were performed by children (92%), similar to previous studies as well (96% [2], and 81% [4]). These findings indicate that children may not notice as quickly or as accurately as adults when targets have been activated. Also, the majority of holdovers (194 out of 283, 69%) occur on the smallest targets (0.125 in), likely because these were the most difficult and participants tended to try to hit them repeatedly. Note that holdovers did occur on 77 out of the 103 targets, however, so they are not isolated to small targets.

Going beyond previous work, the majority of holdovers in our data occurred in TA (92%), showing a clear *beneficial* effect of interface complexity on this behavior. The effect is most clearly shown among children: there were a total of 238 holdovers in TA for children but only 21 in TG. This pattern may be due to the increased visual salience of the ice melting when the touch was registered, making it easier for children to notice and move on.

Target Misses (No Effect of Complexity)

The original studies by Anthony et al. [2,4] emphasized the miss rates exhibited by children and adults, so we also report the pattern of misses across users in our study.

Overall Misses. We first calculated overall miss rate. We excluded holdovers here, so we only counted misses when aiming for the current target. Children missed on the first try 23% of the time for TA and 24% for TG. Adults only missed the target 17% of the time for TA and 15% for TG. These miss rates match Anthony et al.: children 23%, adults 17% [2]. The higher miss rates for children are likely due to less touchscreen experience and ongoing motor skills development. Because the overall miss rates are similar for TA and TG, we can conclude that interface complexity does not affect users successfully hitting onscreen targets.

Per-User Misses. We calculated the per-user proportion of misses over all targets. Like previous studies [2,4], we excluded holdovers and only looked at first attempts. We also examined the effect of target size, which Anthony et al. [2,4] found significant. A repeated measures ANOVA on the *per-user miss rate* with within-subjects factors of *interface complexity* (TA or TG) and *target-size* (very small, small, medium, large) and a between-subjects factor

Figure 6. Average proportion of misses overall by target size. Errors bars indicate the 95% confidence interval.

of *participant type* (adult or child) found a significant main effect of *participant type* ($F_{1,50}$=44.76, p<.0001), in which children miss more often than adults, as above. Also, tests of within-subjects effects with a Greenhouse-Geisser (G-G) correction showed a significant main effect of *target size* ($F_{2.15,107.64}$=897.7, p<.0001). Smaller targets were significantly harder to hit for both children and adults (Figure 6). The same repeated measures ANOVA showed no significant effect of *complexity* ($F_{1,50}$=0.00, *n.s.*), validating that complexity did not affect misses in general. However, it is worth noting that tests of within-subjects effects (G-G applied) showed a marginal three-way interaction between *target size*, *participant type* and *interface complexity* ($F_{2.46,122.86}$=2.33, p<.10). Children missed smaller targets more, but larger targets less, when complexity was introduced; adults were fairly consistent between the two levels of complexity. The visual animations may be the reason: while distracting to children for the smaller targets, they had the opposite effect for the larger targets by making them more noticeable.

Edge Padding. *Edge padding* refers to the target being inset from the side of the screen [2]. We examined the per-user miss rate on targets with and without edge padding, only first attempts, while again excluding holdovers. A repeated measures ANOVA on the *miss rate* with within-subjects factors of *complexity* (TA or TG) and *edge padding* (yes or no) and a between-subjects factor of *age* (grouped into individual years) found a significant main effect of *edge padding* ($F_{1,45}$=262.78, p<.0001). The miss rate was nearly double for targets with edge padding (children: 31%, adults: 22%) versus those without (children: 17%, adults: 11%), replicating prior work [2]. There was also a significant main effect of *age* ($F_{6,45}$=7.17, p<.0001). The younger the participants were, the more misses they had. Notably, there was no interaction between *age* and *edge padding*, contradicting Anthony et al. [4], but we divided the analysis into individual ages, whereas they examined coarse age groups. Edge-padded targets pose more difficulty in part because there are more edges possible to over- or undershoot. On targets without edge padding, the participant only has three sides of the target where a miss could occur since it is aligned to the screen edge. Going beyond prior work, we found no significant effect of *complexity* ($F_{1,45}$=0.56, *n.s.*), meaning the miss rate on edge-padded targets was not influenced by interface complexity. Like Anthony et al. [2,4], we also computed the percent of misses that were within the "gutter", that is, the 10-pixel space between the target and the edge of the screen on edge-padded targets. Nearly all of the misses on edge-padded targets occurred within the gutter (98% across all users), consistent with Anthony et al. (99%) [2,4]. The proportion of misses in the gutter varied by less than 1% for children between TA and TG, and less than 2% for adults, indicating that complexity did not affect this behavior. Children had a slightly higher average rate of misses within the gutter (99%) than adults (96%).

Touch Pressure and Size (Effect of Complexity)

Touch pressure, or the amount of pressure registered by the touchscreen when the participant touches it, and *touch size*, or the size of the contact area registered, were logged by the apps for every touch that the participant performed. A repeated measures ANOVA on *touch pressure* with a within-subjects factor of *complexity* (TA or TG) and a between-subjects factor of *participant type* (child or adult) found a significant main effect of *complexity* ($F_{1,50}$=19.97, p<.0001), and a marginal interaction between *complexity* and *participant type* ($F_{1,50}$=3.497, p<.10). Both children and adults had a lower average touch pressure when using the more complex interface (children: 0.474, adults: 0.481) than the simpler interface (children: 0.493, adults: 0.526). Children had lower average touch pressures than adults, but adults exhibited a larger difference between interfaces.

A repeated measures ANOVA on *touch size* with the same factors found a significant effect of *complexity* ($F_{1,50}$=16.87, p<.0001), and a significant interaction between *complexity* and *participant type* ($F_{1,50}$=4.196, p<.05). Children and adults had larger touch sizes in TA (children: 0.164, adults: 0.176) than TG (children: 0.159, adults: 0.163), though the effect was small. Children also have smaller touch sizes than adults overall, likely due to their smaller finger size.

Both children and adults had smaller touch pressure and a smaller touch size in the complex interface (TG). Participants may have been trying to be more precise when hitting the targets in TG due to the immersion of the game experience. Also, when applying more pressure to the screen, the resulting touch size registered is usually bigger.

Location (No Effect of Complexity)

We also explored the data beyond the measures reported by Anthony et al. [2,4]. One new area we examined was characterizing how children's touches were distributed around targets, and whether interface complexity affected this behavior. We analyzed whether targets in some locations were more difficult than others for children or adults. We calculated the miss rate by target location on the right and left sides of the screen and the top and bottom of the screen. Two separate repeated measures ANOVAs were run on the *miss rate* for the within-subjects factor of either *vertical region* (top, center, bottom) or *horizontal region* (right, middle, left), with an additional within-subjects factor of *interface complexity* (TA or TG) and a between-subjects factor of *participant type* (adult or child).

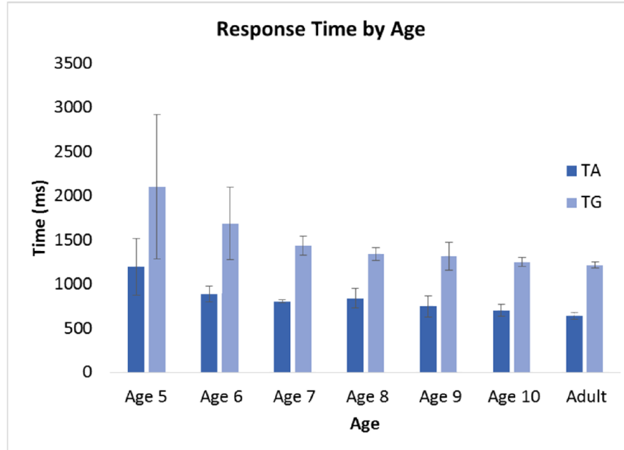


Figure 7. The average response time to hit the first target. Errors bars indicate the 95% confidence interval.

For the *horizontal regions*, there was a significant main effect of the *region* ($F_{1.76,87.9}$=4.38, p<.05, G-G applied) and also of *participant type* ($F_{1,50}$=28.74, p<.0001). There was no effect of *complexity* ($F_{1,50}$=0.19, *n.s.*). Children missed more than adults; however, both children and adults missed significantly more when the targets were on the right side of the screen (M=20.9%, SD=7.2%) than the left and center sides (M=19.8%, SD=9.3%). Most of our participants were right-handed (like the general population), which may contribute to the location of the misses.

In the second ANOVA with the *vertical regions*, we found a significant main effect of the *region* ($F_{2,100}$=10.61, p<.0001) and a significant main effect of *participant type* ($F_{1,50}$=23.59, p<.0001). There were more misses on the top of the screen (M=20.6%, SD=9.6%) than on the bottom and center (M=19.5%, SD=8.0%). There was also a marginal interaction between *complexity* and *participant type* ($F_{1,50}$=3.26, p<.10), and a significant interaction between *vertical region* and *participant type* ($F_{2,100}$=7.73, p<.01). The interaction between complexity and participant type might be due to the animation of the penguin's eyes near the top of the screen, as extra visual stimulus that distracted children more than adults. The extra stimulus would have also made targets harder to locate and touch precisely.

Response Time (Effect of Complexity)

Another measure we considered beyond previous studies [2,4] is *response time*, that is, the amount of time it took the participant to generate the first touch event for a target after it appeared onscreen. We ran a repeated measures ANOVA on *response time* with within-subjects factors of *complexity* and *target-size* and a between-subjects factor of *participant type*. There was a significant main effect of *complexity*

($F_{1,50}$=410.64, p<.0001), a significant main effect of *target size* ($F_{2.22,110.9}$=100.62, p<.0001, G-G applied), and a significant interaction between *complexity* and *target size* ($F_{2.49,124.5}$=11.94, p<.0001, G-G applied). We expect that the effect of complexity is due to the use of animation slowing down the experience. Also, smaller targets had a longer response time for both children and adults. The significant interaction between complexity and target size may be caused by an increase in visual stimulus which made smaller targets more difficult to locate. There was also a significant main effect of *participant type* ($F_{1,50}$=15.75, p<.0001). Children (M=1175 ms, SD=283.9) had slower response times than adults (M=931 ms, SD=72.9).

We ran a separate repeated measures ANOVA similar to the above but with the *participant type* broken down into individual *ages* (e.g., 5,6,7,8,9,10 years old, and adults). There was a significant main effect of *complexity* ($F_{1,45}$=376.7, p<.0001), and a significant main effect of *age* ($F_{6,45}$=11.21, p<.0001). There was also a significant interaction of *complexity* and *age* ($F_{6,45}$=2.53, p<.05). The younger the child, the slower their response time (Figure 7). Young children may be more distracted by visual stimulus than older children and adults. In the graph, it is evident that the animation causes a longer response time in TG than TA, but this effect does not cloud the clear effect of age.

Summary of Touch Interactions

Our analysis of the touch interactions shows that interface complexity affects some interaction patterns for children and adults but not others. Most of the effects of interface complexity were on measures that are affected by visual salience of interface elements like response time, occurrence of holdovers, and touch size and pressure.

Gesture Interactions

We analyzed our participants' gestures in GA and GG by running recognition experiments to understand how age and interface complexity affect recognition rates. The gesture recognition experiments were performed using our own Java-based implementation of the $P recognizer [45], a multi-stroke gesture recognizer, which we selected due to its relatively high popularity, high accuracy, and ease of implementation. Both user-dependent and user-independent experiments were run, explained next. As described earlier, each gesture application consisted of six rounds of 20 gestures each. The first round was always treated as practice, leaving a total of five samples of each gesture per participant in each application.

User-Dependent Results (No Effect of Complexity)

The user-dependent experiment tested recognition rates when the recognizer was trained on the same participant whose gestures would be tested. This test helps designers understand the accuracy levels they can expect when applications are trained on each user, rather than on pre-loaded templates. We used the testing procedure established by Wobbrock et al. [50] and used in later studies [5,6,45]. In our study, the number of training samples was systematically increased from T = 1 to 4 (one must be chosen for testing, leaving a maximum of four for the training set). We had a total of 4 x 10^5 recognition tests (100 trials x 51 participants x 4 values of T x 20 gestures) in the user-dependent case. Because nine participants had some missing data, their maximum number of training examples was reduced. Two had four samples of each gesture, three had three samples, one had two samples; three others had only one sample so they were excluded from the tests.

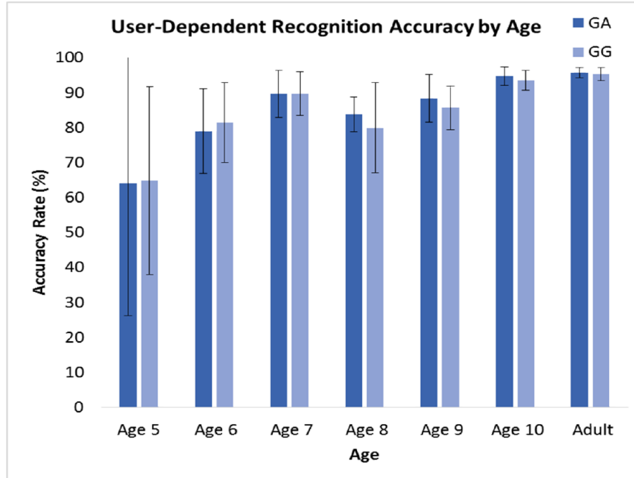


Figure 8. Average user-dependent recognition accuracy by age. Errors bars indicate the 95% confidence interval.

Recognition Accuracy. A repeated measures ANOVA on *per-user recognition accuracy* with a within-subjects factor of *interface complexity* (GA or GG) and a between-subjects factor of *age* (individual ages) found no significant main effect of *complexity* ($F_{1,50}$=0.31, *n.s.*). This result indicates that the increased complexity of the game-like interface of the GG application did not significantly alter the way users made the gestures. We also found a significant main effect of *age* ($F_{6,50}$=9.08, p<.0001). Recognition accuracy was lower the younger the participant was (Figure 8). This finding is consistent with previous studies of adults' and children's gesture recognition rates [2–4], but we examine recognition rates for children at a finer granularity of ages. To illustrate this point further, we ran a bivariate correlation between age and recognition accuracy on just the children, as done in prior work [2,4]. A significant correlation was confirmed (r=.52, p<.0061, N=27). The general rule is: the higher the user's age, the better the recognition accuracy.

Studies of children's drawing abilities [8,27], which may affect the production of gestures by children, can explain these results. Younger children are at an earlier stage of development and as such do not have the motor control needed to make gestures consistently. Children also may not have as much practice making gestures on touchscreens as adults do. Even though we trained the recognizer on children's gestures for these tests, the recognizer was much more accurate when trained and tested on adults. This finding points to a need for recognizers to be designed that can compensate for the inconsistency in children's gestures while they are still developing. Current recognition accuracy rates for children are far from those reported to be acceptable: 91% to 97% according to prior work on children and adults, respectively [24,35]. Prior studies reported 81% recognition for children ages 7 to 17 [2], 84% for children ages 7 to 10 [4], and 77-81% for children ages 10 to 13 [3], but we show that these are not representative for all ages of children. The gestures in our study made by 10-year-olds were recognized with 94% accuracy; however, gestures by 5-year-olds were only at 64% accuracy. The reason that we obtained higher accuracy for older children than previous studies is possibly because those studies grouped younger and older children together. Thus, application developers must carefully consider their target age groups and how this will affect gesture recognition.

Number of Training Examples. We also examined the effect of increasing the number of samples used to train the recognizer. We ran a separate repeated measures ANOVA on *recognition accuracy* with within-subjects factors of *interface complexity* (GA or GG) and *number of training examples* (one to four), and a between-subjects factor of *age* (individual ages). Tests of within-subjects effects with a Greenhouse-Geisser correction found a significant effect of the *number of training examples* on recognition accuracy ($F_{1.11,50.8}$=186.1, p<.0001). As expected, and shown by prior work [3,45,50], recognition accuracy improves as the number of training examples is increased. The test also confirmed a significant impact of *age* on recognition accuracy ($F_{6,46}$=7.73, p<.0001). Furthermore, the test found a significant interaction between the *number of training examples* and *age* ($F_{6.63,50.8}$=5.38, p<.0001, G-G applied). The accuracy of recognition tests with only one training example is much worse for younger children, and it catches up more slowly as the number of training examples is increased. This finding implies that the number of training examples used to train a recognizer will depend on the age of the target audience. Developers should expect accuracy rates to "level off" relatively quickly for adults (three to four training examples), while access to more training examples for young children will enable better accuracy.

User-Independent Results (No Effect of Complexity)

We also tested children's gestures in a user-independent experiment. In this test, the recognizer is trained on gestures made by participants other than the one who made the gesture which is being tested. While this results in lower accuracy than the user-dependent test, it gives designers an idea of accuracy rates they can expect for "out of the box" applications with pre-trained recognizers. Because user-independent tests require consistent numbers across users, we first removed the gestures from the nine participants who did not completely finish the gesture tasks (13% of the gestures). We had a total of 51 participants (27 adults and 24 children) with complete data for the gesture tasks.

We did not have enough training examples per person to run the user-independent tests by individual ages. Therefore, we ran the tests by age group: 5- to 7-year-olds (10 participants), 8- to 10-year-olds (14 participants), and adults (30 participants). We used the procedure explained by Vatavu et al. [45] with 10 trials, 9 participants (P), and 5 training examples (T). In total, there were 2.7 x 10^5 recognition tests (3 age groups x 9 values of P x 5 values of T x 10 trials per combination of P and T x 10 trials per testing participant and test set x 20 gestures).

The average user-independent recognition rate we found was 81% for adults, 76% for ages 8 to 10, and 65% for ages 5 to 7. All of these values are well below reported acceptance rates [24,35]. As in the user-dependent case, a repeated measures ANOVA on *recognition error rate* per number of participants in the training pool with a within-subjects factor of *interface complexity* and a between-subjects factor of *age* found no significant effect of *complexity* ($F_{1,24}$=1.81, *n.s.*). The test showed a significant effect of *age* ($F_{2,24}$=9.88, p<.01), as expected.

Summary of Gesture Interactions

Our gesture recognition experiments show that there is a significant effect of age on recognition accuracy in both user-dependent and user-independent scenarios, supporting previous work [2–4]. There was no significant effect of interface complexity on recognition rates, so visual aspects of the interface do not affect how users make gestures.

DISCUSSION AND DESIGN IMPLICATIONS

This study's findings supplement previous work that has examined children's touchscreen interactions in isolation [2–4,7]. Our study has expanded the focus from simple stimulus-response apps to more realistic applications. Our work shows how the findings from simpler applications can *generalize* to real-world applications. Our findings point to several *design recommendations* for touchscreen interfaces for children of this age group. Regarding the effect of interface complexity, significant to mention is the fact that children did not miss targets more often when using a more complex application, nor did they miss by a greater margin, nor did they make their gestures any differently (e.g., same recognition accuracy). Therefore, our findings support design recommendations from Anthony et al. [2–4] pertaining to ignoring holdovers, using reasonable target sizes, allowing out-of-bounds touches, aligning targets to screen edges, training age-specific recognizers, and designing gesture sets for children. We provide new recommendations that follow from our data.

Provide salient visual feedback of accepted input to prevent holdovers. In our study, participants, especially children, experienced many fewer holdovers in the more complex application. We speculate that this finding is due to the increased visual saliency of the feedback received during gameplay for correct touches that hit the fish targets. Therefore, in applications designed for children, ensuring that activation of interface widgets via onscreen feedback is prominent and immediately noticeable (e.g., animation) will prevent the occurrence of holdovers that may distract children from their main tasks. An example is a touchscreen keyboard on a smartphone: children may benefit from larger and longer feedback when pressing an onscreen key.

Avoid small targets at the screen edges, especially in visually complex interfaces. We found that children tended to miss small targets on the right-hand side or top edge of the screen more often than other locations or other target sizes. This effect was even more pronounced in the visually complex application, possibly due to decreased visual saliency of targets in those locations in relation to onscreen graphics. We emphasize that in both conditions, the smaller targets had a much lower success rate overall. The right-hand side problem is at least in part likely due to the demographics of our participants, who were predominantly right-handed, and unintended touches would be more likely to occur on that side. (Left-handedness appears in approximately 10% of people [30].) We recommend designers avoid laying out small targets in places that will be more difficult for children to notice and touch or more susceptible to unintended touches.

Consider the trade-off between visual saliency and response time when designing games or applications for speedy input. Children, especially ages 5 and 6, exhibited a slower response time in the more visually complex application than in the simpler application. Many games rely on speed of response as part of the challenge of gameplay or increase the required speed to increase the challenge of the game, e.g., Pop Balloon Kids[1]. For young children, this strategy must balance between onscreen visual stimulus and desired reaction speed. An animation-centric game will be more distracting to the children, slowing their reaction times. Play-testing with the target age group will help designers balance their games and account for this effect. Future work is needed to understand how other types of interface changes affect response time.

Train gesture recognizers for younger children with more examples. In our gesture recognition experiments, we noted that recognition accuracy was worse the younger the child, for both user-dependent and user-independent tests. In addition, increasing the number of training examples per gesture continued to improve accuracy for the youngest children, while for older children and adults accuracy “leveled off” after only a few samples in both testing cases. The recognizer we used in this work, $P [45], is notable for requiring few training examples, like other members of the $-family [5,6,45,50]. However, the inconsistency in gesture styles exhibited by young children requires more examples to compensate. Designers of applications for very young children will need to collect more training data upfront to achieve reasonable levels of accuracy. Our results do show that designers can collect examples from many different children rather than requiring many examples from one child. Future studies could answer how many samples and children are needed to obtain a desired accuracy. Further work is also needed on algorithms to recognize young children’s gestures more accurately in general.

LIMITATIONS AND FUTURE WORK

The study we present goes beyond prior work by studying younger children and effects of interface complexity on interaction. Still, the study has several limitations. First, the children we recruited were ages 5 to 10, an age range experiencing rapid cognitive and motor skills development [34,44]. However, findings for this age group may not transfer to younger children. Second, gesture recognition accuracy in our study did not “level off” for the youngest children. Having more training examples might provide better guidance on how much data is needed to support young children; however, issues of attention during empirical studies make this difficult [9,10]. Future work could collect data in more natural activities, e.g., integrated into games children already play at home or at school. Third, even the complex interfaces used in this study only show one target onscreen at a time and prompt users for clearly segmented gestures. Recognition of intended targets and gestures is more challenging in real tasks on touchscreen devices, and should be considered in future work. Finally, this study only used small screen devices and may not generalize to larger screen devices. For example, tabletops afford different types of interaction than small screens, including collaboration [18,21,40], which should be considered in future work. This study also points toward new investigations into recognition of young children’s gesture input, including tailored or even personalized gesture recognizers, adaptive gesture recognition interfaces, and design of natural gestural user interfaces for children.

CONCLUSION

We presented an empirical study of 30 children ages 5 to 10 and 30 adults who interacted with touchscreen apps we built to log all their touch and gesture interactions. Two of the apps used an abstract stimulus-response interface to elicit touchscreen interactions, and two of the apps used game-like interfaces to immerse the participants in the task. We analyzed the interaction logs for similarities and differences between children and adults, and between the two types of apps. We also conducted offline gesture recognition experiments to estimate the accuracy on young children’s gestures and how much data is needed for reasonable accuracy. Interface complexity affected some touch interactions related to visual salience, and did not affect gesture recognition. We also find general differences between children and adults. Based on our findings, we presented design recommendations for touchscreen interfaces for children of this age. Our work contributes to practical design of interfaces for young children, and answers research questions about how age, interface complexity, and task elements affect children’s interactions.

ACKNOWLEDGMENTS

This work is partially supported by National Science Foundation Grant Awards #IIS-1218395 / IIS-1433228 and IIS-1218664. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect these agencies’ views. The authors thank Thaddeus Brown, Monique Ogburn, Chiamaka Okorohoa, Femi Williams, and Qingchuan “Bruce” Zhao for assistance with this work.

REFERENCES

1. Nor Azah Abdul Aziz, Firat Batmaz, Roger Stone, and Paul Wai Hing Chung. 2013. Selection of touch gestures for children's applications. *Science and Information Conference (SAI), 2013*, 721–726.

2. Lisa Anthony, Quincy Brown, Jaye Nias, Berthel Tate, and Shreya Mohan. 2012. Interaction and Recognition Challenges in Interpreting Children's Touch and Gesture Input on Mobile Devices. *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, ACM Press, 225–234. http://doi.org/10.1145/2396636.2396671

3. Lisa Anthony, Quincy Brown, Jaye Nias, and Berthel Tate. 2013. Examining the need for visual feedback during gesture interaction on mobile touchscreen devices for kids. *Proceedings of the International Conference on Interaction Design and Children*, ACM Press, 157–164. http://doi.org/10.1145/2485760.2485775

4. Lisa Anthony, Quincy Brown, Berthel Tate, Jaye Nias, Robin Brewer, and Germaine Irwin. 2014. Designing smarter touch-based interfaces for educational contexts. *Personal and Ubiquitous Computing* 18, 6, 1471–1483. http://doi.org/10.1007/s00779-013-0749-9

5. Lisa Anthony and Jacob O. Wobbrock. 2010. A lightweight multistroke recognizer for user interface prototypes. *Proceedings of Graphics Interface*, Canadian Information Processing Society, 245–252.

6. Lisa Anthony and Jacob O. Wobbrock. 2012. $N-protractor: a fast and accurate multistroke recognizer. *Proceedings of Graphics Interface*, 117–120.

7. Ahmed Sabbir Arif and Cristina Sylla. 2013. A comparative evaluation of touch and pen gestures for adult and child users. *Proceedings of the International Conference on Interaction Design and Children*, ACM Press, 392–395. http://doi.org/10.1145/2485760.2485804

8. Keith E. Beery, Norman A. Buktenica, and Natasha A. Beery. 2004. *The Beery-Buktenica Developmental Test of Visual-Motor Integration, 5th Edition*. Modern Curriculum Press, New Jersey.

9. Robin Brewer, Lisa Anthony, Quincy Brown, Germaine Irwin, Jaye Nias, and Berthel Tate. 2013. Using gamification to motivate children to complete empirical studies in lab environments. *Proceedings of the International Conference on Interaction Design and Children*, ACM Press, 388–391. http://doi.org/10.1145/2485760.2485816

10. Quincy Brown, Lisa Anthony, Robin Brewer, Germaine Irwin, Jaye Nias, and Berthel Tate. 2013. Challenges of Replicating Empirical Studies with Children in HCI. *Proceedings of the CHI2013 Workshop on the Replication of HCI Research*, CEUR Workshop Proceedings, 53–57.

11. Quincy Brown, Lisa Anthony, Jaye Nias, Berthel Tate, Robin Brewer, and Germaine Irwin. 2013. Towards Designing Adaptive Touch-Based Interfaces. *Proceedings of the ACM SIGCHI 2013 Third Mobile Accessibility Workshop*, 4pp.

12. Quincy Brown and Lisa Anthony. 2012. Toward comparing the touchscreen interaction patterns of kids and adults. *Proceedings of the SIGCHI Workshop on Educational Software, Interfaces and Technology*, 4pp.

13. Cynthia Chiong and Carly Shuler. 2010. *Learning: Is there an app for that? Investigations of young children's usage and learning with mobile devices and apps*. The Joan Ganz Cooney Center at Sesame Workshop, New York, NY. Retrieved August 23, 2012 from http://dmlcentral.net/resources/4496

14. Andy Cockburn, David Ahlström, and Carl Gutwin. 2012. Understanding performance in touch selections: Tap, drag and radial pointing drag with finger, stylus and mouse. *International Journal of Human-Computer Studies* 70, 3, 218–233. http://doi.org/10.1016/j.ijhcs.2011.11.002

15. Nick Donnelly, Kyle Cave, Rebecca Greenway, Julie A Hadwin, Jim Stevenson, and Edmund Sonuga-Barke. 2007. Visual search in children and adults: top-down and bottom-up mechanisms. *Quarterly Journal of Experimental Psychology* 60, 1, 120–36. http://doi.org/10.1080/17470210600625362

16. Henry Gleitman, James Gross, and Daniel Rossberg. 2010. *Psychology*. W. W. Norton & Company.

17. Tracy Anne Hammond and Brandon Paulson. 2011. Recognizing sketched multistroke primitives. *ACM Transactions on Interactive Intelligent Systems* 1, 1, 1–34. http://doi.org/10.1145/2030365.2030369

18. Amanda Harris, Jochen Rick, Victoria Bonnett, Nicola Yuill, Rowanne Fleck, Paul Marshall, and Yvonne Rogers. 2009. Around the table: Are multiple-touch surfaces better than single-touch for children's collaborative interactions? *Proceedings of the International Conference on Computer Supported Collaborative Learning*, International Society of the Learning Sciences, 335–344. http://doi.org/10.3115/1600053.1600104

19. James Herold and Thomas F. Stahovich. 2012. The 1 ¢ Recognizer: a fast, accurate, and easy-to-implement handwritten gesture recognition technique. *Proceedings of the International Symposium on Sketch-*

Based Interfaces and Modeling, Eurographics Association, 39–46. http://doi.org/10.2312/SBM/SBM12/039-046

20. Alexis Hiniker, Kiley Sobel, Sungsoo Ray Hong, Hyewon Suh, India Irish, Daniella Kim, and Julie A. Kientz. 2015. Touchscreen prompts for preschoolers. *Proceedings of the International Conference on Interaction Design and Children*, ACM Press, 109–118. http://doi.org/10.1145/2771839.2771851

21. Uta Hinrichs and Sheelagh Carpendale. 2011. Gestures in the wild: studying multi-touch gesture sequences on interactive tabletop exhibits. *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, ACM Press, 3023–3032. http://doi.org/10.1145/1978942.1979391

22. L.B. Kara and Thomas F. Stahovich. 2005. An image-based, trainable symbol recognizer for hand-drawn sketches. *Computers & Graphics* 29, 4, 501–517. http://doi.org/10.1016/j.cag.2005.05.004

23. Hong-hoe Kim, Paul Taele, Stephanie Valentine, Erin McTigue, and Tracy Hammond. 2013. KimCHI: a sketch-based developmental skill classifier to enhance pen-driven educational interfaces for children. *Proceedings of the International Symposium on Sketch-Based Interfaces and Modeling*, ACM Press, 33–42. http://doi.org/10.1145/2487381.2487389

24. Mary J. LaLomia. 1994. User acceptance of handwritten recognition accuracy. *Extended Abstracts of the ACM SIGCHI Conference on Human Factors in Computing Systems*, ACM, 107–108. http://doi.org/10.1145/259963.260086

25. WeeSan Lee, Levent Burak Kara, and Thomas F. Stahovich. 2007. An efficient graph-based recognizer for hand-drawn symbols. *Computers & Graphics* 31, 4, 554–567. http://doi.org/10.1016/j.cag.2007.04.007

26. Yang Li. 2010. Protractor: a fast and accurate gesture recognizer. *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, ACM, 2169–2172. http://doi.org/10.1145/1753326.1753654

27. Viktor Lowenfeld and W. Lambert Brittain. 1982. *Creative and Mental Growth*. MacMillan, New York, NY, USA.

28. Lorna McKnight and Brendan Cassidy. 2010. Children's Interaction with Mobile Touch-Screen Devices. *International Journal of Mobile Human Computer Interaction* 2, 2, 18 pp. http://doi.org/10.4018/jmhci.2010040101

29. Lorna Mcknight and Daniel Fitton. 2010. Touch-screen Technology for Children : Giving the Right Instructions and Getting the Right Responses. *Proceedings of the International Conference on Interaction Design and Children*, ACM Press, 238–241. http://doi.org/10.1145/1810543.1810580

30. Ian Christopher McManus. 2009. The history and geography of human handedness. In *Language lateralization and psychosis* (1st ed.), Iris E. C. Sommer and René S. Kahn (eds.). Cambridge University Press, 37–58.

31. Vicente Nacher, Javier Jaen, Elena Navarro, Alejandro Catala, and Pascual González. 2015. Multi-touch gestures for pre-kindergarten children. *International Journal of Human-Computer Studies* 73, 37–51. http://doi.org/10.1016/j.ijhcs.2014.08.004

32. Cathleen Norris and Elliot Soloway. 2004. Envisioning the handheld-centric classroom. *Journal of Educational Computing Research* 30, 4, 281–294. http://doi.org/10.2190/MBPJ-L35D-C4K6-AQB8

33. Pekka Parhi, Amy K. Karlson, and Benjamin B. Bederson. 2006. Target size study for one-handed thumb use on small touchscreen devices. *Proceedings of the International Conference on Human Computer Interaction with Mobile Devices and Services*, ACM Press, 203–210. http://doi.org/10.1145/1152215.1152260

34. Jean Piaget. 1983. Piaget's Theory. In *Handbook of Child Psychology*, P Mussen (ed.). Wiley & Sons, New York, NY, USA.

35. Janet C. Read, Stuart MacFarlane, and Chris Casey. 2003. "Good enough for what?": acceptance of handwriting recognition errors by child users. *Proceedings of the International Conference on Interaction Design and Children*, ACM, 155–155. http://doi.org/10.1145/953536.953565

36. Glenda L. Revelle and Emily Reardon. 2009. Designing and testing mobile interfaces for children. *Proceedings of the International Conference on Interaction Design and Children*, ACM Press, 329–332. http://doi.org/10.1145/1551788.1551876

37. Shyam Reyal, Shumin Zhai, and Per Ola Kristensson. 2015. Performance and User Experience of Touchscreen and Gesture Keyboards in a Lab Setting and in the Wild. *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, ACM Press, 679–688. http://doi.org/10.1145/2702123.2702597

38. Geoff Romeo, Suzy Edwards, S. McNamara, Ian Walker, and Christopher Ziguras. 2003. Touching the

screen: Issues related to the use of touchscreen technology in early childhood education. *British Journal of Educational Technology* 34, 3, 329–339. http://doi.org/10.1111/1467-8535.00330

39. Dean Rubine. 1991. Specifying gestures by example. *SIGGRAPH Computer Graphics* 25, 4, 329–337. http://doi.org/10.1145/127719.122753

40. Kathy Ryall, Meredith Ringel Morris, Katherine Everitt, Clifton Forlines, and Chia Shen. 2006. Experiences with and observations of direct-touch tabletops. *Proceedings of the IEEE International Workshop on Horizontal Interactive Human-Computer Systems*, IEEE, 8pp. http://doi.org/10.1109/TABLETOP.2006.12

41. Tevfik Metin Sezgin and Randall Davis. 2005. HMM-based efficient sketch recognition. *Proceedings of the International Conference on Intelligent User Interfaces*, ACM Press, 281–283. http://doi.org/10.1145/1040830.1040899

42. Mike Sharples, Inmaculada Amedillo Sanchez, Marcelo Milrad, and Giasemi Vavoula. 2009. Mobile learning: small devices, big issues. In *Technology Enhanced Learning: Principles and Products*, N. Balacheff, S. Ludvigsen, T. de Jong and S. Barnes (eds.). Springer, Heidelberg, 233–249.

43. Carly Shuler. 2009. *Pockets of Potential: Using Mobile Technologies to Promote Children's Learning*. Joan Ganz Cooney Center at Sesame Workshop, New York, NY. Retrieved August 23, 2012 from http://www.joanganzcooneycenter.org/publication/industry-brief-pockets-of-potential-using-mobile-technologies-to-promote-childrens-learning/

44. Jerry R. Thomas. 1980. Acquisition of motor skills: information processing differences between children and adults. *Research quarterly for exercise and sport* 51, 1, 158–173. http://doi.org/10.1080/02701367.1980.10609281

45. Radu-Daniel Vatavu, Lisa Anthony, and Jacob O. Wobbrock. 2012. Gestures as point clouds: a $P recognizer for user interface prototypes. *Proceedings of the ACM International Conference on Multimodal Interaction*, ACM Press, 273–280. http://doi.org/10.1145/2388676.2388732

46. Radu-Daniel Vatavu, Gabriel Cramariuc, and Doina Maria Schipor. 2015. Touch interaction for children aged 3 to 6 years: Experimental findings and relationship to motor skills. *International Journal of Human-Computer Studies* 74, 54–76. http://doi.org/10.1016/j.ijhcs.2014.10.007

47. Radu-Daniel Vatavu, Daniel Vogel, Géry Casiez, and Laurent Grisoni. 2011. Estimating the perceived difficulty of pen gestures. *Proceedings of the IFIP TC13 Conference on Human-Computer Interaction*, Springer, 89–106. http://doi.org/10.1007/978-3-642-23771-3_9

48. Daniel Vogel and Ravin Balakrishnan. 2010. Direct Pen Interaction with a Conventional Graphical User Interface. *Human–Computer Interaction* 25, 4, 324–388. http://doi.org/10.1080/07370024.2010.499839

49. Jacob O. Wobbrock, Htet Htet Aung, Brandon Rothrock, and Brad A. Myers. 2005. Maximizing the guessability of symbolic input. *Extended Abstracts of the ACM SIGCHI Conference on Human Factors in Computing Systems*, ACM, 1869–1872. http://doi.org/10.1145/1056808.1057043

50. Jacob O. Wobbrock, Andrew D. Wilson, and Yang Li. 2007. Gestures without libraries, toolkits or training: a $1 recognizer for user interface prototypes. *Proceedings of the ACM Symposium on User Interface Software and Technology*, ACM, 159–168. http://doi.org/10.1145/1294211.1294238