Designing Smarter Touch-Based Interfaces for Educational Contexts

Lisa Anthony†, Quincy Brown‡, Berthel Tate‡, Jaye Nias‡, Robin Brewer†, Germaine Irwin†

UMBC†, Bowie State University‡

lanthony@umbc.edu, qbrown@bowiestate.edu

Abstract

In next-generation classrooms and educational environments, interactive technologies such as surface computing, natural gesture interfaces, and mobile devices will enable new means of motivating and engaging students in active learning. Our foundational studies provide a corpus of over 10,000 touch interactions and nearly 7,000 gestures collected from nearly 70 adults and children ages 7 years old and up, that can help us understand the characteristics of children's interactions in these modalities and how they differ from adults. Based on these data, we identify key design and implementation challenges of supporting children's touch and gesture interactions, and we suggest ways to address them. For example, we find children have more trouble successfully acquiring onscreen targets and having their gestures recognized than do adults, especially the youngest age group (7 to 10 years old). The contributions of this work provide a foundation that enables touch-based interactive educational apps that increase student success.

Keywords: Touch interaction, gesture interaction, child-computer interaction, mobile devices, gesture recognition, $N, $P, educational technology.

Introduction

Over the past decade, interactive technologies such as surface computing, natural gesture interfaces, and mobile devices have made the transition from niche markets to mainstream products that are available to the end-user. With this transition comes new opportunities for the design of interactive experiences that provide seamless, natural interactions using modalities users can apply across contexts. This set of new technologies will enable new means of motivating and engaging students of all ages in active learning experiences. As students are able to use natural interaction modalities such as touch, gesture, voice, and motion in new applications that support them, they will no longer need to consciously adapt their interactions to each new digital device or tool. Rather, they will be able to use existing tools and abilities they have learned in other apps, allowing them to focus on the learning tasks or goals at hand.

We present our work on the fundamental building blocks of this next generation of education technology, specifically dealing with touch and gesture interactions enabled by surface computing devices such as smartphones. We believe that this set of modalities will be common in the near-term next wave of educational applications, because of the growing adoption of mobile touchscreen devices such as smartphones and tablets. These devices enable *personalized learning* [11], since the cost for each student to have his or her own device is no longer prohibitive [32]. The mobile form factor also enables *situated learning* [29], facilitating use of the devices anywhere, from the classroom, to the science laboratory, to field trips, and in the home [10,42].

Educational technology serves children of all ages. However, younger children, in elementary through middle school (about ages 7 to 13), present a special challenge for the use of mobile touchscreen devices for their educational activities. Typically, these devices and their underlying interactions are not designed specifically for children. Rather, the devices are developed for the "typical" adult user, as mass-market, consumer-oriented products. The current design of the most widespread devices has not taken into account the specific needs of young users. Furthermore, literature on developmental psychology [22,34,44] reveals that children are inherently different and not just "smaller adults." Their motor skills are still developing, often at a rapid pace [22]. Their performance and sensory processing speeds are markedly different than that of adults [44]. Specifically, children have smaller fingers, weaker arms, less fine motor control, and (typically) less experience with technology than adults. These factors may contribute to key differences in how children can, expect to, and do use touch and gesture interaction and will impact their success with mobile devices. So far, this impact has not been examined thoroughly.

To investigate these differences, we present two foundational studies, conducted on touchscreen smartphones, with nearly 70 adults and children ages 7 years old and up. We describe a corpus of over 10,000 touch interactions and nearly 7,000 gestures that can help us understand the characteristics of children's interactions in these modalities and how they differ from adults. Based on these data, we identify key design and implementation challenges of supporting children's touch and gesture interactions, and suggest ways to address them. For example, we find that children miss onscreen touch targets far more often than adults, especially for smaller targets, and children exhibit distinctive behaviors such as unintentional touches in the location of the previous target, a phenomenon we term *holdovers*. We also find that children's gestures are less accurately recognized than adults' gestures by current gesture recognition approaches. Our results indicate that these challenges are experienced most heavily by the youngest age group in our studies, children ages 7 to 10 years old. These findings build on our prior work in this space [2,8], increasing the robustness of these conclusions across a wider range of users.

The contributions of this work will assist designers and developers who are creating interactive educational apps for surface computing platforms, especially small-screen devices such as smartphones, but also extending to larger devices such as tablets. This work can help apps increase student success by allowing users to focus on learning goals instead of on making themselves understood by the device or app. In our vision of the future of educational apps, rather than the user adapting to the device, the device will adapt to the user and to the context, improving learning.

Related Work

Touch and Surface Gesture Interactions

Previous research has examined the usability of interaction on touch- and surface-gesture-based platforms, for both finger and pen, including Tablet PCs and tablet computers [46,47], tabletop and surface displays [6,16,38,40], and mobile devices [4,18,20,30,33]. Most of this work has focused on adult users without explicit consideration of younger users. In cases where children have been included, findings have been general conclusions rather than direct comparisons between adults and children [40]. Likewise, other research has included children only [16,38], further hampering the comparison of children and adults.

As the use of touch-based platforms by young users [43] has increased, so has the need for further work on direct comparisons of children and adults as users to understand how best to accommodate children. Prior work on stylus-based interaction on Tablet PCs has examined handwriting input for children to compare their performance with adults' performance [36]. Recent work has begun to explicitly explore differences between children and adults on touchscreen devices [2,8,19], but more work is needed. The Sesame Workshop has developed a set of "best practice" interaction design recommendations for touchscreen apps for pre-school children (ages 4 and under) based on more than 50 touchscreen studies with this age group [41], but they do not compare these recommendations with those for older children or adults. We extend this work to school-age children, and we prioritize cross-generational comparisons throughout our studies.

Pointing Interactions

Pointing interactions on the desktop have been well-studied, including pointing for children. Prior studies have explored how well Fitts' law [14] applies to children by comparing adults and children acquiring targets using mice or other pointing devices [12,21,25]. Other work has found that, as children grow older, their movements to complete aiming tasks such as pointing become more stable, and that, as children repeat these tasks, their behavior begins to remain consistent by age twelve [39].

A few studies have investigated pointing on touchscreens specifically, exploring the relationship between factors such as target size, target location, and finger size to understand the effects of "fat fingers" and occlusion while using small screens [20,35]. New interaction paradigms have been proposed to improve touchscreen interaction in unstable environments, such as applications of *goal crossing* [1] to touchscreens [27]. However, again such evaluations of pointing-based interactions on touchscreens have been conducted only on adults, hindering the comparison between adults and children.

Results from a few studies that *have* compared children and adults have revealed that children experience difficulty with drag-and-drop interactions on both desktops and mobile devices [23,24]. For example, children have difficulty maintaining contact with the touchscreen as required to complete dragging gestures [9]. These findings begin to point to design recommendations for touchscreen interaction for children, but more work is needed.

Cognitive and Physical Factors

Latency, defined as the time between a user's action and a system's response, has been studied in mobile devices with respect to touch and haptic feedback [26]. Adults' latency perception has been found to be between 25 to 40 milliseconds (ms) for visual-auditory asynchrony and haptic-auditory asynchrony [26]. Other research has revealed a latency effect of more than 80 ms from when a finger touches a touchscreen, the time required for the vibration motor to start, and the expected application response [17]. These latencies can cause unexpected interactions and poor performance, if the device does not respond as expected. Since more children are now using mobile devices, we must continue to explore the effects of touch latency on adult and child users.

In addition to latency perception, manual dexterity is another factor that can contribute to a user's ability to successfully acquire touch targets or use surface gesture interactions. Previous research has revealed that people's gross motor skills (e.g., large body movements) and fine motor skills (e.g., small body movements) refine as they progress from infancy to adulthood [22,39,44]. This knowledge has been used by educational psychologists in the development of instruments that are used to assess an individual's ability to draw shapes. For example, it is developmentally appropriate for a child aged twelve years to draw a three-dimensional cube, though only appropriate for a child aged four years to draw a one-dimensional square [5]. Furthermore, the impacts of these stages of development have been evident in previous research that found that little hands unexpectedly activate interactors on mobile touchscreen devices [31], and that children have difficulty maintaining contact with touchscreens to effectively complete drag-and-drop interactions [9]. Thus, as touch-based interactions become more commonplace, there is a need for additional research to understand how the devices can be used by individuals of varying physical and cognitive developmental levels.

Foundational Studies

In our own work, we have employed quantitative and qualitative methods to understand the characteristics of touch and surface gesture interactions for both adults and children. In our previous work [2,8], we have presented small studies of adult and child users interacting with mobile applications that implement the interactions of real-world touch and gesture activities such as tapping, sketching, and writing [2,8]. Building on these results, we extend our previous small-scale studies here to understand whether our findings are consistent and robust for larger groups of participants. We present findings from two studies with adults and children. Participants worked with two mobile apps that we designed and implemented. Data from Study 1 have been previously published [2]. Since Study 2 uses nearly the same protocol, we group both datasets together for analysis purposes.

Participants

We conducted two touch and surface gesture interaction studies with a total of 74 participants: 44 children and 30 adults. Over both studies, the child participants ranged in age from 6 to 17 years (mean: 12.2 yrs, stdev: 2.4 yrs) and all of the adults were over the age of 18 (mean: 23.7 yrs, min: 18 yrs, max: 33 yrs, stdev: 4.0 yrs). Of the 30 adults, 12 were female, and of the 44 children, 23 were female. A large majority of the participants were right handed (61 of 74); 6 considered themselves ambidextrous, and 7 were left-handed.

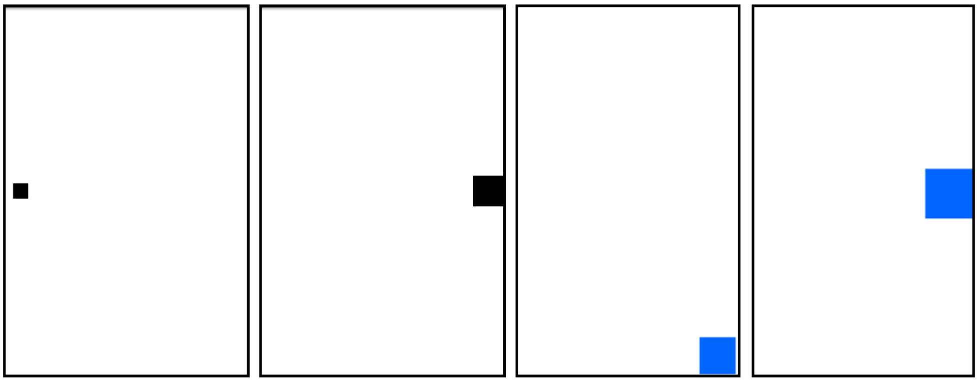
During the study, participants completed a questionnaire to ascertain their familiarity with touchscreen devices. Adults tended to self-rate their familiarity as "expert" (20 of 30, or 67%) or as "average" (10 of 30, or 33%). Children rated themselves as "beginners" (2 of 44, or 5%), "average" (18 of 44, or 41%), or "expert" (23 of 44, or 52%). Table 1 shows the percentage of our participants (adults and children) that used and owned touchscreen devices. These results reflect the rising ubiquity of touchscreen devices such as smartphones and tablets.

		Adults	**Children**
Mobile phone	Own one	83%	41%
	Use one daily	80%	66%
Tablet	Own one	33%	30%
	Use one daily	40%	39%
MP3 Player	Own one	33%	48%
	Use one daily	17%	41%
Tablet PC	Own one	10%	2%
	Use one daily	7%	7%

Table 1. Distribution of ownership and usage of touchscreen devices for the 74 participants in our studies.

Equipment

Both experiments were conducted on Samsung Google Nexus S smartphones running the Android 4.0.4 operating system. The phones measured 4.88 x



(a) **(b)** **(c)** **(d)**

Figure 1. Target task interface screens, from Study 1 with black target squares: (a) very small, with edge padding and (b) small, flush to the edge of the screen; and from Study 2 with blue target squares: (c) medium, with edge padding, and (d) large, flush to the screen edge. In both studies, all four target sizes were represented equally.

2.48 x 0.43” with a 4” diagonal screen. The display resolution was 480 x 800 pixels (233 pixels per inch (ppi) pixel density).

Target Acquisition Task

The goal of the target acquisition task was to identify the touch characteristics of adults and children and how they may differ. For details on the study design and protocol, see Anthony et al. [2]; for convenience, pertinent details are briefly described here.

The touch task application used 104 targets of four different sizes: very small (0.125”), small (0.25”), medium (0.375”), and large (0.5”). (Targets were designed to display uniformly across devices of varying screen resolution.) These sizes were chosen based on target sizes in existing applications in order to examine the range of acceptable target sizes. Targets could appear in any one of 13 pre-defined locations (selected using locations in existing apps, *e.g.,* along edges, in corners, in the center of screen). In both studies, we introduced a variable we call *edge padding*: half of the targets were slightly inset from the edge of the screen by a constant “gutter” of 10 pixels. The other half were aligned with the edge of the screen. A total of 104 targets resulted from this combination of 4 sizes x 13 locations x 2 edge-padding conditions, appearing in a set order evenly balanced for transitions.

Participants sat at a table in a user studies lab for both studies and were allowed to hold or rest the mobile device in a comfortable manner. They were instructed to touch each target that appeared onscreen. The app scored each touch event as a *hit* or a *miss*, depending on whether the touch was registered within the bounds of the target (this scoring was not visible to the participant). On a successful *hit*, the interface advanced to display the next target. Figure 1 shows an example of each target size; black targets were shown in Study 1, and blue targets were shown in Study 2 (blue was chosen to increase visual salience of targets when next to screen edges). All touch events registered by the device were logged and recorded for later data analysis, including information such as the *x-coordinate*, *y-coordinate*, time, touch pressure, and touch size of each event.

Gesture Interaction Task

The goal of the gesture interaction phase was to characterize how adults and children create gestures and identify differences that may cause recognition of these gestures to be challenging. Again, we provide a summary of the pertinent details here; for more in-depth treatment of our method, see Anthony et al. [2].

The gesture task application prompted the participant to draw a specific gesture. We chose a set of 20 gestures based on existing apps as well as literature from educational psychology [5] to ensure that the selected gestures were developmentally appropriate for children. After entering the prompted gesture, the participant was instructed to touch the onscreen “Done” button. Figure 2 shows screenshots of the gesture task interface for both Study 1 and Study 2, both before and after a gesture has been entered. Although we have also explored versions of the app in which no visual trace of the gesture path is shown, the data presented in this paper used apps that showed visual feedback. The complete gesture set included letters, numbers, symbols, and geometric shapes[1] (Figure 3).

[1] We did not include command gestures common today such as swipe and pinch-to-zoom, based both on initial studies finding that these gestures are difficult for children [9], and on the prevalence of tracing or handwriting practice activities in

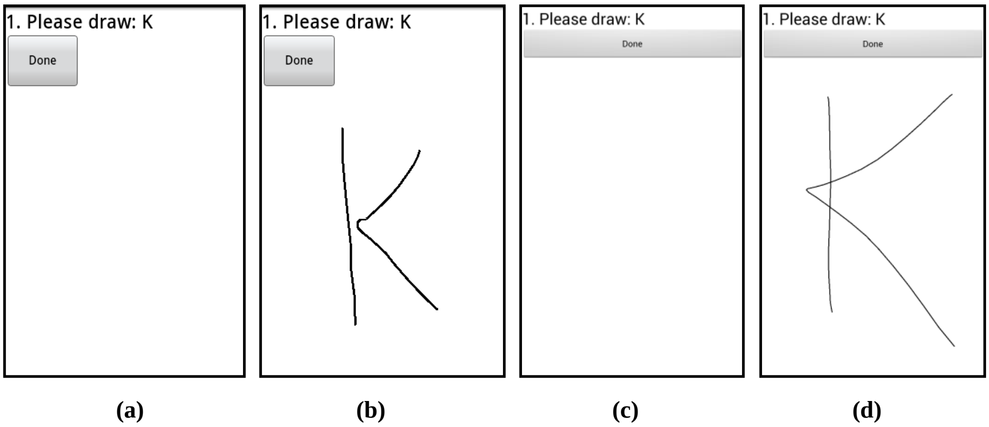


Figure 2. Gesture task interface screens, from Study 1 with a smaller Done button: (a) before drawing the gesture and (b) the trace the user sees after drawing the gesture; and from Study 2 with a larger Done button (to prevent unintentional touches after finishing the gesture): (c) before drawing the gesture and (d) after drawing the gesture. The prompt at the top of the screen tells the participant what gesture to draw.

Figure 3. The set of 20 gestures used in our studies.

Participants again sat at a table in the lab for both studies and were allowed to hold or rest the mobile device in a comfortable manner. To help young children understand the textual onscreen prompts, in Study 1, participants were given a paper sheet showing what each gesture should look like. In Study 2, we instead asked participants to hand-draw each gesture on paper to avoid priming them to gesture styles. The app prompted the participant to enter one example of each gesture in the set, and then repeated the set five times, yielding a total of 120 gesture samples per participant. All touch events were again logged and recorded. All strokes made before hitting "Done" were counted as part of one gesture and stored together. Aside from certain features of the gesture that can be calculated from these data (e.g., height, width, duration, etc.), we also conducted gesture recognition experiments with several recognizers (discussed in detail later in the paper) on the gestures collected in both studies.

Data Analysis Procedures

We have analyzed data from over 10,000 touches and nearly 7,000 gestures to characterize both adults' and children's patterns. For our analysis, we divide the children into four subgroups: 5 to 7 years old (yrs), 8 to 10 yrs, 11 to 13 yrs, and 14 to 17 yrs. These groups were chosen based on developmental psychology literature (e.g., Piaget [34]), our previous experience conducting research with children, and typical school age groupings in the United States: elementary school (5 to 10 yrs), middle school (11 to 13 yrs), and high school (14 to 17 yrs). Since we have only three 5-to-7-year-olds in our Study 1+Study 2 dataset, for this paper we group them with the 8-to-10-year-olds during data analysis. Future work will examine the 5-to-7-year-old age group in more depth.

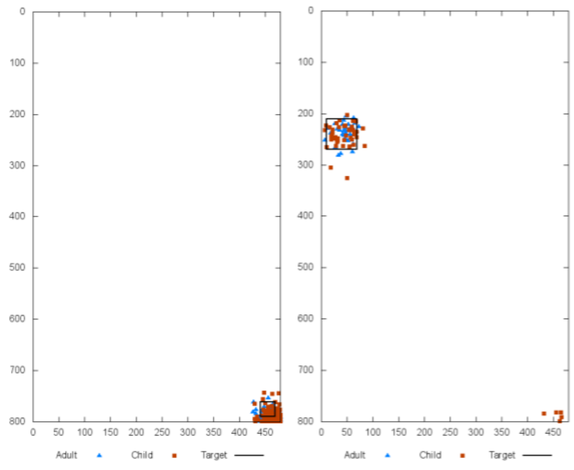
Target Acquisition Findings

Of the 74 participants across Study 1 and Study 2, some users' log data were excluded due to technical difficulties or protocol abnormalities such as not finishing the task. A total of 66 participants' data are included in our analysis of the target interaction task (37 children, 39 adults). Also, the first target for every person was counted as practice and excluded from analysis. After removing these and other individual abnormal attempts (i.e., task restarts or long pauses, less than 0.1% of data), we had just under 10,300 total touch events across the 66 participants.

Holdovers

In the touch interaction task, we observed touches that were located in the vicinity of the *previous* target, which we have termed *holdovers* [2]. These

children's education apps today [2]. Future work could examine other gestures in more depth.



(a) (b)

Figure 4. Holdovers in the bottom right corner visible after the transition from target (a) to target (b).

touches occurred because the application advances to the next target after a successful touch; holdovers occur possibly before the participant is able to notice that the interface has changed. Out of the over 10,000 touch events, 153 were classified as holdovers (1.5%). Although this is only a very small proportion of all touch events, we believe holdovers are a distinct class of unintentional interaction with potentially negative consequences (e.g., launching or canceling a process, changing a setting, or entering or deleting data without the user noticing) that can easily be detected and avoided. We classified a touch event as a *holdover* if its (*x*,*y*) coordinates were within a certain threshold from the *previous* onscreen target. This threshold was set to 117 pixels, the width of the largest target size on the study devices. (This formula did not suffer from any false positives because subsequent target stimuli were always much more than 117 pixels apart.) The classifications were verified visually by examining images of all touch events recorded for each target across the 2 studies (in which holdovers are visible clearly, for example, see Figure 4). Holdovers occurred on 40 of the 103 (38.8%) targets.

Analysis revealed that the holdover phenomenon occurred more often in data from children (81.0% of all holdovers), and was roughly evenly split between the two youngest age groups (35.3% and 38.6% for 7 to 10 and 11 to 13, respectively), with only 7.2% for 14 to 17 and 19.0% for adults. A one-way ANOVA on per-user frequency of holdovers by *age group* found a significant difference ($F_{3,62}=7.28$, $p<0.01$). Children ages 7 to 10 (M = 3.93%, SD = 2.88%) had higher holdover frequency than children ages 11 to 13 (M = 1.59%, SD = 2.14%), 14 to 17 (M = 0.69%, SD = 0.88%) and adults (M = 0.56%, SD = 1.32%).

Misses

Across all data, 78.3% of the targets were successfully hit on the first attempt, while 21.7% of the targets showed multiple attempts before success (mean number of attempts per user per target: M = 1.53; SD = 1.79). The maximum number of attempts for a single target (for a single user) was 37.

Not counting holdovers, children in general missed more targets than adults. A repeated measures ANOVA was performed on the per-user target miss rate (e.g., proportion of targets missed on the first attempt) with a within-subjects factor of *target size* (large, medium, small, very small) and a between-subjects factor of *age group* (7 to 10, 11 to 13, 14 to 17, or adults 18+). Tests of within-subjects effects with a Greenhouse-Geisser correction indicate a significant interaction between *age group* and *target size* ($F_{7.4,152.9}=2.52$, $p<0.05$). All participants experienced the most difficulty with the "very small" targets, but the youngest children (7 to 10) experienced more difficulty with the next smallest target size ("small") than the other age groups (Figure 5). Figure 6 shows examples of touch event distribution patterns in the dataset we collected; a representative target for each size is drawn within the context of the screen size, and all touches that were recorded for that target for both children and adults (both *hits* and *misses*) are shown. The many more misses that children clearly make on small targetscompared to adults create a challenge for determining the intended target, when multiple interactive targets are onscreen.

The edge padding variable also contributed to the proportion of misses per target. In the presence of edge padding, the target miss rate nearly doubled (for both children and adults) compared to no edge padding (Figure 7). A repeated measures ANOVA was performed on the per-user target miss rate with a within-subjects factor of presence of *edge padding* (yes or no) and a between-subjects factor of *age*

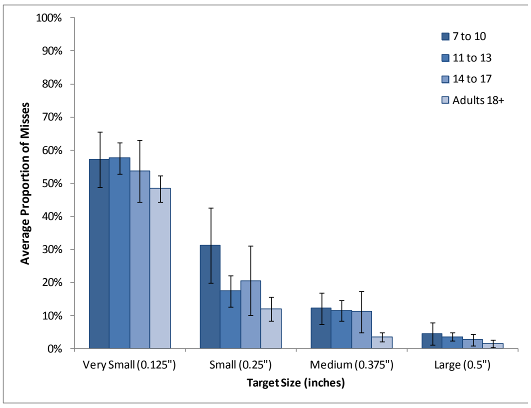
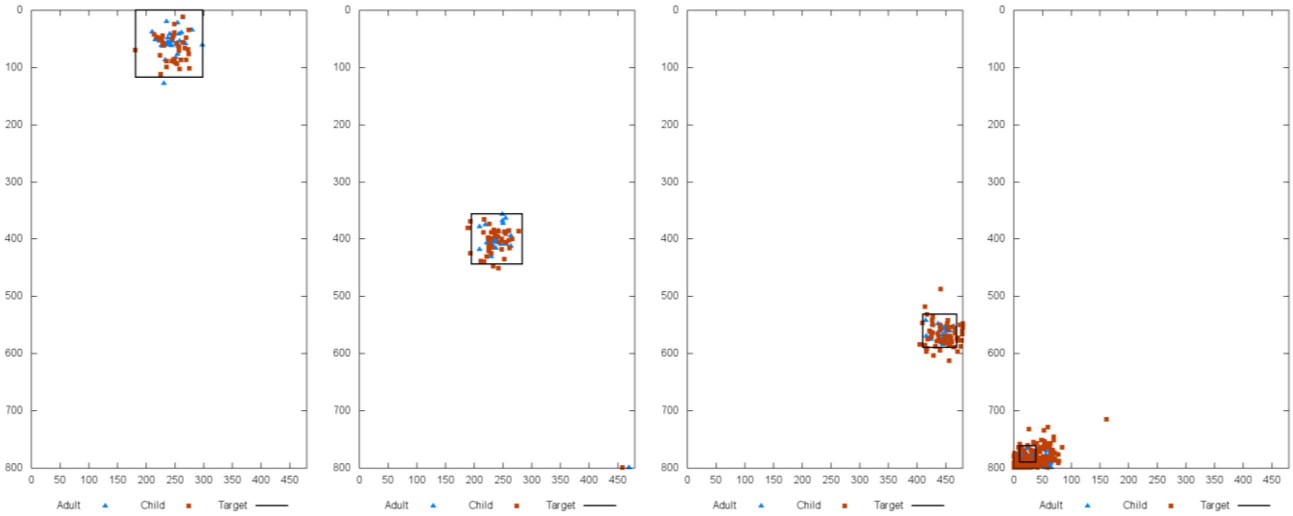


Figure 5. Average proportion of misses overall. Errors bars indicate the 95% confidence interval.



(a) (b) (c) (d)

Figure 6. Examples of the distribution of touch events for adults (blue triangles) and for children (red squares) for targets in our studies: (a) large, (b) medium, (c), small, and (d) very small.

group (7 to 10, 11 to 13, 14 to 17, or adults 18+). Again, a significant interaction was found between *age group* and *edge padding* by multivariate tests ($F_{3,62}$=3.93, $p<0.05$). The youngest children had more trouble with edge-padded targets than did older children or adults.

Interestingly, nearly all (99%) misses that occurred on edge-padded targets were located within the "gutter" (the space between the target and the edge of the screen), consistently for all age groups in our dataset. This finding contradicts the utility of interactive targets that are inset from the screen, as seen in some mobile apps, since it is difficult for users not to hit that area by mistake.

Gesture Interaction Findings

In total we collected 8,880 gesture samples from 74 people, who each provided 6 samples of 20 gestures.

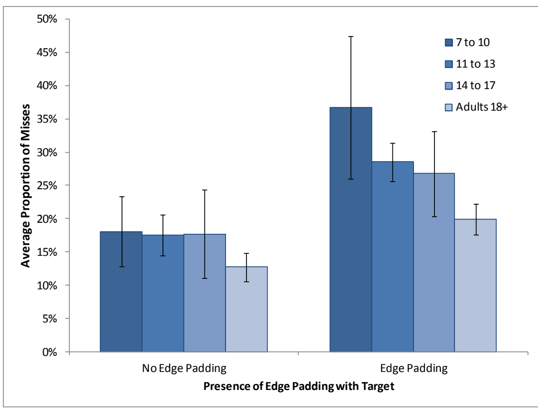


Figure 7. Average proportion of misses for edge padded targets only. Errors bars indicate the 95% confidence interval.

For the same reasons noted for the target task above, we excluded log data from some participants due to technical difficulties and protocol abnormalities. A total of 69 participants' data are included in the analysis for the gesture interaction task (39 children, 30 adults). The first round of gestures for all participants was considered practice, leaving 6,900 gesture samples for our analysis. Some data from Study 1 had also been removed from analysis due to data collection abnormalities [2], but after a slight re-design of the gesture app (Figure 2), this issue did not occur in Study 2. In total, we include just over 6,700 gestures in the analysis presented in this paper.

Recognition Accuracy

Our past work has shown that children's gestures are significantly less accurately recognized by modern recognizers compared to gestures by adults: 81% for children on average compared to 90% for adults [2] for *user-dependent* testing, and 34% for children compared to 64% for adults for *user-independent* testing [8][2]. That testing was done primarily with just one recognizer, the $N-Protractor recognizer [3], which is an open-source, trainable recognizer currently used by gesture interaction researchers and mobile app developers. We are interested in whether other recognizer types might actually perform better than $N-Protractor for children, since $N-Protractor tends to work best with more consistent input [3]. As

[2] *User-dependent* testing refers to recognition testing in which the recognizer is trained on samples of the same person's writing on which it is to be tested. *User-independent* testing refers to training and testing on samples only from different users.

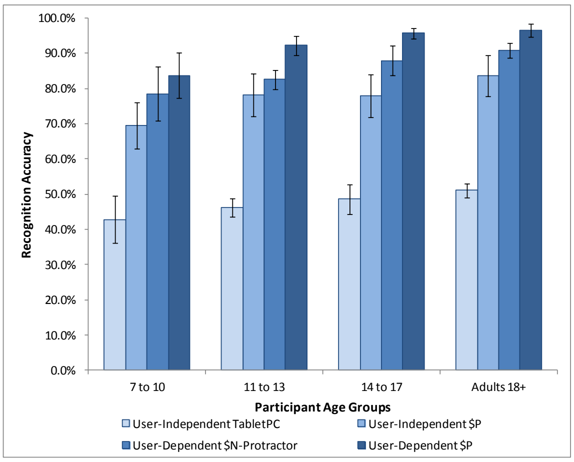


Figure 8. Recognition performance by age group for each of the four recognizer configurations tested. Error bars indicate the 95% confidence interval. User-dependent $P is the best performer but still performs quite low, around 84%, for young children.

our prior work showed, children are much less consistent than are adults in how they make gestures [8]; for example, they tend to use more strokes and make larger gestures. Therefore, recognizers that are more robust to number of strokes and other execution details may be more successful for children's gestures. $P [45], another open-source and trainable recognizer, is a recent advance over $N-Protractor in terms of robustness to gesture execution styles, and we hypothesized it may do better on children's gestures.

For this paper, we tested our corpus of gestures with both $N-Protractor and $P, as well as the industry standard Microsoft Tablet PC recognizer (non-trainable) for comparison's sake[3]. We compared children to adults based on the same four age groups we have been using throughout this paper. To get a fuller picture of expected recognition accuracy first out of the box, and then after learning a particular child's gesture patterns over time, we compare both *user-independent* (out-of-the-box accuracy, for $P and Microsoft Tablet PC recognizer) and *user-dependent* (per-child accuracy, for $N-Protractor and $P) recognition accuracy.

Figure 8 shows a graph of the recognition accuracy means by age group and recognizer. The best out-of-the-box performer is $P: accuracy ranges from nearly 84% for adults, to about 70% for the youngest children in our study (7 to 10). Prior work [28] has indicated that adults will tolerate recognition accuracy only as low as 97%, whereas child-specific studies [37] found children more tolerant, accepting about 91% accuracy. Thus, the user-independent configuration of $P is not accurate enough for good out-of-the-box performance for new children using a gesture-based interactive system. The per-child recognition accuracy of the user-dependent configuration of $P is the best performer overall, reaching nearly 97% for adults and 96% for teenagers, but still performs quite low, around 84%, for young children. These findings point to a need for tailored age-specific gesture recognition for children that can be more accurate on the types of gestures expected from them.

		7 to 10	11 to 13	14 to 17	Adults 18+
User-Indep. TabletPC	***M***	42.7%	46.2%	49.1%	50.7%
	SD	9.1%	6.4%	6.7%	5.4%
	N	7	22	10	30
User-Indep. $P	***M***	69.6%	78.2%	78.0%	83.7%
	SD	8.3%	7.6%	7.7%	7.2%
	N	6	6	6	6
User-Dep. $N-Prot.	***M***	75.7%	82.4%	86.4%	90.4%
	SD	10.3%	6.5%	6.8%	6.1%
	N	7	22	10	30
User-Dep. $P	***M***	83.7%	92.2%	95.7%	96.6%
	SD	8.7%	6.3%	2.4%	5.0%
	N	7	22	10	30

Table 2. Descriptive statistics for the four recognizers tested by age group. Recognition accuracy was highest with $P, but lowest overall for children ages 7 to 10.

A repeated-measures ANOVA was performed on per-user average gesture recognition accuracy with a within-subjects factor of *recognizer* (Microsoft Tablet PC, $N-Protractor, and $P) and a between-subjects factor of *age group* (7 to 10, 11 to 13, 14 to 17, and adults), with a Greenhouse-Geisser correction. No significant interaction was found between *recognizer* and *age group* ($F_{4.98,107.8}$=1.56, *n.s.*), but there is a significant main effect of both *age group* ($F_{3,65}$=13.75, p<0.01) and *recognizer* ($F_{1.66,107.8}$=1168.7, p<0.01). As mentioned, recognition accuracy was highest with $P, but lowest overall for young children ages 7 to 10 (Table 2). These findings indicate that the more robust recognition approach of $P is in fact more suited to children's gestures than the other recognizers. Still, children's gestures are less well recognized than adults' even by $P, according to a one-way ANOVA on *age group* for just the $P user-independent recognition results ($F_{3,20}$=3.42, p<0.05).

As in our prior work [2], we hypothesized that there might be a direct relationship between recognition accuracy and a child's age, that is, we expected that accuracy on children's gestures would be lower for younger children. Indeed, there was a strong positive

[3] Note that the Tablet PC recognizer is only suited for handwriting gestures, so we removed arch, arrowhead, checkmark, diamond, heart, rectangle and triangle from the results; circle, line and plus have keyboard symbol equivalents and were kept, along with all the numbers and letters in the corpus.

correlation between age (in years) and average recognition accuracy by $P for children in our dataset (*r*=.60, *p*<0.01, *N*=39). Because of this correlation, we expect accuracy for gestures made by children even younger than those in our studies to continue to deteriorate. Future work must investigate recognition approaches that can be tailored to the types of gestures we expect children of all ages to generate.

Implications for Design

This work, based on data from two studies with nearly 70 adults and children so far, has indicated that there are reliable differences in how children and adults use touchscreens that impact successful touchscreen interaction. Many of the technical challenges we have identified in previous reports on these and similar studies [2,8] involve understanding and interpreting the user's intended target or gesture, especially for children. We review these implications, as well as new ideas for future work based on these cumulative findings.

Prevent Unintended Touch Interactions

Many of our recommendations center on challenges of understanding and interpreting what the user intended to do in ambiguous interaction modalities such as touch and gesture. These challenges are even more impactful for children than for adults, given the range of inconsistent input children tend to generate according to our studies. Systems that have been programmed or trained to expect certain patterns of input will find it difficult to handle children's input, which may not have been predicted at design time.

In our prior work on Study 1 alone [2], we identified several design recommendations that could be incorporated into such systems to allow them to better adapt to children's input. We reiterate them here succinctly for reference, and add to them. Table 3 summarizes them for quick and easy reference.

Use timing and location of touches and interface widgets to identify and ignore holdover touches (corollary: improve interface responsiveness to prevent holdovers). Our extended dataset shows the holdover effect is robust, occurring in the same patterns as for our previous report on Study 1: mostly the youngest children, mostly the smallest targets. As we mentioned, although holdovers are a small subset of the overall touch event data we collected (1.5%), their consequences are potentially severe. Holdovers can be easy to prevent using a simple heuristic that ignores any touches registered in the same place as the previously-activated widget within a short time threshold.

Use consistent, platform-recommended target sizes. Our extended dataset supports our initial finding from Study 1 that the "very small" targets, 0.125" square, presented the greatest challenge for both children and adults: approximately half were missed on the first try, more than double the miss rate for the other sizes. In the extended dataset presented in this paper, we see that the youngest children (ages 7 to 10) miss an average of 30% of the "small" (0.25") targets as well. This finding contradicts our original hypothesis that children's smaller fingers might enable them to use smaller touch targets. Rather, our data indicate that, the younger the child, the larger the targets are actually necessary to be.

Increase active area for interface widgets to allow slightly out-of-bounds touches to register and activate the intended widget. In our data, missed touches tend to be close to the intended target—not many touches are registered in other areas of the screen. When more than one target is onscreen at once, an intelligent interaction layer could probabilistically determine which interactive widget

Interaction	Strategy	Design Implication
Touch and Pointing	**Prevent unintended touch interactions**	Use timing and location of touches and interface widgets to identify and ignore holdover touches (corollary: improve interface responsiveness to prevent holdovers).
		Use consistent, platform-recommended target sizes.
		Increase active area for interface widgets to allow slightly out-of-bounds touches to register and activate the intended widget.
		Align targets to edge of screen, or count edge touches.
Gesture	**Tailor gesture recognition for children**	Train age-specific recognizers to improve accuracy on children's gestures.
		Design gestures and gesture sets that make conceptual sense to children and are easy for them to execute.
		Develop child-specific recognizers from the ground up.
		Allow recognizers to learn over time and adapt to an individual child's gestures.

Table 3. Design recommendations for touch and gesture interaction on mobile devices that we suggest based on our studies with children and adults.

was intended based on the location of the touch event compared to the visual boundary of the widget (e.g., [13,15]). This will increase the number of successful interactions a child will have while using the device.

Align targets to edge of screen, or count edge touches. The proportion of missed touches which occur in the “gutter” between the target and the edge of the screen (99% of touches for all age groups) remained consistent from Study 1 to the extended dataset in Study 2. We note that, if targets cannot be aligned to the screen edge for some reason (the ideal solution), touches registered in the “gutter” should activate the nearest interactive widget using similar probabilistic reasoning as suggested in the previous recommendation.

Tailor Gesture Recognition for Children

Understanding what gesture a user meant to input is the main challenge for gesture interaction. Based on our dataset of gestures made by children and adults froma wide range of age groups, consistency in gesture execution patterns among different users is not guaranteed, impacting gesture recognition accuracy. Therefore, we focus our design recommendations for gesture interaction on the recognition itself. Those recommendations that were initially proposed in our prior work on Study 1 alone [2] are reiterated here succinctly for reference, and new ones are added.

Train age-specific recognizers to improve accuracy on children’s gestures. Our extended dataset confirms our finding from Study 1 that children’s gestures are more difficult than adults’ to recognize and interpret, even by very different recognizers. Furthermore, recognizers have the most trouble with the youngest children’s gestures: recognition accuracy was strongly positively correlated with age. For best out-of-the-box performance for new children using a device, we recommend training specific recognizer models per age (e.g., a “7-year-old” recognizer, a “10-year-old” recognizer, etc.).

Design gestures and gesture sets that make conceptual sense to children and are easy for them to execute. Since prior work has indicated that children have trouble executing drag-and-drop gestures [9], we recommend exploring other interactive gestures that may be more suited to the cognitive and physical development of children. Development psychology literature [5] has established clear guidelines about the types of gestures and shapes children can be expected to execute smoothly based on their age. Therefore, we have explored mostly shape-based gestures in our studies, and find that typically the children participating do understand these gestures and are able to make them. We recommend using similar gesture types for all interaction, rather than relying on complex abstract gestures such as drag-and-drop or pinch-to-zoom.

Develop child-specific recognizers from the ground up. Our earlier work has found significant differences in the ways children and adults execute gestures; for example, children make taller gestures with more strokes [8]. In this extended dataset of Study 1 and 2, we did not find further support for such differences. However, $P [45] was the most successful recognizer we tested on our dataset, partly because its recognition approach is more robust to execution differences. Still, recognition accuracy continues to be lower for children, pointing to as-yet-unknown ways in which children’s gestures differ from adults. Beyond simply training existing recognizers to samples of children’s input, we also suggest the development of new recognition approaches developed from the ground up with children’s input in mind.

Allow recognizers to learn over time and adapt to an individual child’s gestures. User-dependent recognition, in which the recognizer is trained on samples of the same user’s data on which it is being tested, was the most accurate configuration on the extended dataset in this paper, supporting similar comparisons in our prior work [2,8]. At first, all users are first-time, “out-of-the-box” users (*user-independent*). With every new gesture seen, however, the recognizer can update its model to become more *user-dependent* over time, increasing accuracy by adapting to an individual child. Of course, accuracy in this configuration is still quite low for the youngest children in our data (less than 84%). Therefore, it is likely that a combination of this approach along with child-specific recognition algorithms will be required for accurate, successful gesture interaction for children.

Extensions of This Work

This paper presents a set of design recommendations for touch and gesture interaction for children that is based on findings proven to be robust over two studies with nearly 70 adults and children of various ages. We are currently moving this work forward in three key ways: (a) working with even younger children, (b) exploring larger-screen devices, and (c) investigating educational contexts. We briefly describe our initial explorations into these areas and how they may lead to more specific design recommendations, and educational applications.

Working with Younger Children

In the studies presented in this paper, the children who participated were six years old and older. We are currently investigating whether children ages 5 through 7 exhibit similar patterns of input behavior as the older participants in our studies. We have conducted a small pilot study with 5 to 7 year olds, using the same two tasks in this paper. From this pilot, we have identified potential challenges for collecting controlled laboratory data from children of this age group [7]. In the gesture task, we observed these young children having difficulty completing the entire gesture set. They often appeared bored or fidgety, and several children requested to end the session early. In the target task, we noticed that young children often had to repeatedly touch the targets and/or modify how they were holding the phone. During the pilot, only one child completed all of the gesture and target tasks.

Based on our observations in this pilot study, we have been exploring modifications to address the low task completion rates. These modifications include giving points for completing portions of the tasks which children can redeem at the end of the session for rewards, for example, small items like candy, toys, or games. The points are nothing more than a progress-meter, but we hope the points (and rewards) will encourage children to complete the entire session. This would allow us to compare these young children with our existing data. We anticipate such motivating details will apply equally well to older children and to adults, maintaining comparability across new studies we will run.

Exploring Larger-Screen Devices

The work presented in this paper focuses on smartphones, which are only one type of touchscreen device platform available today. Smartboards, tabletops, and tablet devices also support touch input, but on larger screens. While we anticipate that similar findings will hold for larger device sizes, the studies we have so far conducted cannot definitively answer that question. We plan to explore tablet devices in the near future, as the next most common device being used today by children and teens, and then move toward even large form-factors such as smartboards and tabletops. In the end, we will have validated empirical evidence across a variety of touchscreen device form factors that can be used to design tailored interactive apps for children.

Investigating Educational Contexts

In our foundational studies thus far, we have been exploring interaction with abstract tasks, focusing on the basic atomic input events we might expect to see from children and adults, irrespective of the task at hand. We hypothesize that, as the user’s goal shifts in different contexts, input behaviors may also shift. Therefore, we are next planning to explore whether the same input patterns occur in a game-like app, in preparation for moving to more complex environments such as intelligent tutors. If we determine that similar patterns are observed, general design recommendations can apply across contexts. If they do not hold, context-specific interaction paradigms may be necessary. We believe studying the differences between contextualized and abstract apps can transfer to all age groups.

We have already begun design and development of in-context apps that are modeled after our abstract task applications, but add elements of game-like activity to change the user’s focus during the task. As an example, the new target task application, ‘Patty the Penguin’, uses the same targets as the original target app. However, instead of a white background, a penguin occupies the center of the screen and, instead of blue squares, small fish appear. Participants have to help Patty Penguin eat by touching the fish to capture them for her. We might expect to see different patterns of touch events for such a context. For example, perhaps children will miss even more often or experience more holdovers as they become engrossed in the metaphor of the game. Designing different interactive layers for different contexts may be required. We anticipate exploring educational applications as well, once the basic contextual factors are understood.

Conclusion

This paper has presented our investigations into mobile touch and gesture interaction for children and how children might differ from adults. In data from two studies with nearly 70 users, we have found robust evidence that children have more trouble successfully acquiring onscreen targets and having their gestures recognized than do adults, especially the youngest age group (7 to 10 years old). We also have examined the *holdover* phenomenon, which children exhibit far more often than do adults. In next-generation classrooms and educational environments, mobile touchscreen devices will enable new means of motivating and engaging students in active learning. Based on our findings, we identify key design and implementation challenges of supporting children’s touch and gesture interactions, and we suggest ways to address them. The contributions of this work will assist designers and developers who are creating interactive educational apps for surface computing platforms, aiming to decrease student distraction due to failed interactions and increase student success.

Acknowledgements

We would like to thank Chiamaka Okorohoa, Thaddeus Brown, Monique Ogburn, and Shreya Mohan for their support of this research. This work was partially supported by Department of Education HBGI Grant Award #P031B090207-11 and National Science Foundation Grant Awards #IIS-1218395 / 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.

References

1. Accot, J. and Zhai, S. More than dotting the i's --- foundations for crossing-based interfaces. *Proc. CHI 2002*, ACM Press (2002), 73–80.
2. Anthony, L., Brown, Q., Nias, J., Tate, B., and Mohan, S. Interaction and Recognition Challenges in Interpreting Children's Touch and Gesture Input on Mobile Devices. *Proc. ITS 2012*, ACM Press (2012), 225–234.
3. Anthony, L. and Wobbrock, J.O. $N-protractor: a fast and accurate multistroke recognizer. *Proc. GI 2012*, Canadian Information Processing Society (2012), 117–120.
4. Baudisch, P. and Chu, G. Back-of-device interaction allows creating very small touch devices. *Proc. CHI 2009*, ACM Press (2009), 1923–1932.
5. Beery, K., Buktenica, N., and Beery, N.A. *The Beery-Buktenica Developmental Test of Visual-Motor Integration, 5th Edition*. Modern Curriculum Press, New Jersey, 2004.
6. Benko, H., Wilson, A.D., and Baudisch, P. Precise selection techniques for multi-touch screens. *Proc. CHI 2006*, ACM Press (2006), 1263–1272.
7. Brown, Q., Anthony, L., Brewer, R., Irwin, G., Nias, J., and Tate, B. Challenges of Replicating Empirical Studies with Children in HCI. *ACM SIGCHI RepliCHI Workshop 2013*, (2013), 5pp.
8. Brown, Q. and Anthony, L. Toward Comparing the Touchscreen Interaction Patterns of Kids and Adults. *ACM SIGCHI EIST Workshop 2012*, (2012), 4pp.
9. Brown, Q., Hatley, L., Bonsignore, E.M., and Druin, A. Mobile Natives: Unlocking the Potential of Educational Technology. *ACM SIGCHI 2nd Workshop on UI Technologies and Their Impact on Educational Pedagogy 2011*, (2011), 4pp.
10. Chen, Y.S., Kao, T.C., and Sheu, J.P. A mobile learning system for scaffolding bird watching learning. *Journal of Computer Assisted Learning 19*, 3 (2003), 347–359.
11. Cordova, D.I. and Lepper, M.R. Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *Journal of Educational Psychology 88*, 4 (1996), 715–730.
12. Donker, A. and Reitsma, P. Aiming and clicking in young children's use of the computer mouse. *Computers in Human Behavior 23*, 6 (2007), 2863–2874.
13. Findlater, L. and Wobbrock, J.O. Personalized input: improving ten-finger touchscreen typing through automatic adaptation. *Proc. CHI 2012*, ACM Press (2012), 815–824.
14. Fitts, P.M. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology 47*, 6 (1954), 381–391.
15. Goodman, J., Venolia, G., Steury, K., and Parker, C. Language modeling for soft keyboards. *Proc. IUI 2002*, ACM Press (2002), 194–195.
16. Harris, A., Rick, J., Bonnett, V., Yuill, N., Fleck, R., Marshall, P., and Rogers, Y. Around the Table: Are multiple-touch surfaces better than single-touch for children's collaborative interactions? *Proc. CSCL 2009*, ISLS (2009), 335–344.
17. Henze, N. and Poppinga, B. Measuring Latency of Touch and Tactile Feedback in Touchscreen Interaction Using a Mobile Game. *ACM MobileHCI Workshop on Research in the Large 2012*, (2012).
18. Henze, N., Rukzio, E., and Boll, S. 100,000,000 taps: Analysis and Improvement of Touch Performance in the Large. *Proc. MobileHCI 2011*, ACM Press (2011), 133–142.
19. Hinrichs, U. and Carpendale, S. Gestures in the wild. *Proc. CHI 2011*, ACM Press (2011), 3023–3032.
20. Holz, C. and Baudisch, P. Understanding touch. *Proc. CHI 2011*, ACM Press (2011), 2501–2510.
21. Hourcade, J.P., Bederson, B.B., Druin, A., and Guimbretière, F. Differences in pointing task performance between preschool children and adults using mice. *ACM Transactions on Computer-Human Interaction 11*, 4 (2004), 357–386.
22. Hurley, A. Cognitive Development: An Overview. *Tufts Open Course Ware*, 2010. http://ocw.tufts.edu/Content/35/lecturenotes/375938.

23. Inkpen, K.M. Drag-and-drop versus point-and-click mouse interaction styles for children. *ACM ToCHI 8*, 1 (2001), 1–33.
24. Joiner, R., Messer, D., Light, P., and Littleton, K. It is best to point for young children: a comparison of children's pointing and dragging. *Computers in Human Behavior 14*, 3 (1998), 513–529.
25. Jones, T. An Empirical Study of Children's Use of Computer Pointing Devices. *Journal of Educational Computing Research 7*, 1 (1991), 61–76.
26. Kaaresoja, T. and Brewster, S. Feedback is... late: measuring multimodal delays in mobile device touchscreen interaction. *Proc. ICMI-MLMI 2010*, ACM Press (2010), Article No. 2.
27. Kane, S.K., Avrahami, D., Wobbrock, J.O., Harrison, B., Rea, A.D., Philipose, M., and LaMarca, A. Bonfire: a nomadic system for hybrid laptop-tabletop interaction. *Proc. UIST 2009*, ACM Press (2009), 129–138.
28. LaLomia, M. User acceptance of handwritten recognition accuracy. *Ext. Abstracts CHI 1994*, ACM Press (1994), 107–108.
29. Lave, J. and Wenger, E. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, Cambridge, UK, 1991.
30. Lin, M., Price, K.J., Goldman, R., Sears, A., and Jacko, J. Tapping on the move-Fitts' Law under mobile conditions. *Proc. Info. Res. Mgmt. Assoc. Intl. Conf. 2005*, (2005), 132–135.
31. McKnight, L. and Cassidy, B. Children's Interaction with Mobile Touch-Screen Devices. *International Journal of Mobile Human Computer Interaction 2*, 2 (2010), 18pp.
32. Norris, C. and Soloway, E. Envisioning the Handheld-Centric Classroom. *Journal of Educational Computing Research 30*, 4 (2004), 281–294.
33. Parhi, P., Karlson, A.K., and Bederson, B.B. Target size study for one-handed thumb use on small touchscreen devices. *Proc. MobileHCI 2006*, ACM Press (2006), 203–210.
34. Piaget, J. Piaget's Theory. In P. Mussen, ed., *Handbook of Child Psychology*. Wiley & Sons, New York, NY, USA, 1983.
35. Potter, R.L., Weldon, L.J., and Shneiderman, B. Improving the accuracy of touch screens: an experimental evaluation of three strategies. *Proc. CHI 1988*, ACM Press (1988), 27–32.
36. Read, J.C., MacFarlane, S., and Casey, C. Pens behaving badly-usability of pens and graphics tablets for text entry with children. *Adj. Proc. UIST 2002*, ACM Press (2002), 21–22.
37. Read, J.C., MacFarlane, S., and Casey, C. 'Good enough for what?': acceptance of handwriting recognition errors by child users. *Proc. IDC 2003*, ACM Press (2003), 155–155.
38. Rick, J., Harris, A., Marshall, P., Fleck, R., Yuill, N., and Rogers, Y. Children designing together on a multi-touch tabletop: An analysis of spatial orientation and user interactions. *Proc. IDC 2009*, ACM Press (2009), 106–114.
39. Rösblad, B. Reaching and eye-hand coordination. In A. Henderson and C. Pehoski, eds., *Hand Function in the Child: Foundations for Remediation*. Mosby Elsevier, St. Louis, MO, USA, 2006.
40. Ryall, K., Morris, M.R., Everitt, K., Forlines, C., and Shen, C. Experiences with and observations of direct-touch tabletops. *Proc. Tabletop 2006*, IEEE (2006), 8pp.
41. Sesame Workshop. *Best Practices: Designing Touch Tablet Experiences for Preschoolers*. New York, NY, 2012.
42. Sharples, M., Amedillo Sanchez, I., Milrad, M., and Vavoula, G. Mobile learning: small devices, big issues. In N. Balacheff, S. Ludvigsen, T. de Jong and S. Barnes, eds., *Technology Enhanced Learning: Principles and Products*. Springer, Heidelberg, 2009, 233–249.
43. Shuler, C. *Pockets of Potential: Using Mobile Technologies to Promote Children's Learning*. Joan Ganz Cooney Center at Sesame Workshop, New York, NY, 2009.
44. Thomas, J.R. Acquisition of motor skills: information processing differences between children and adults. *Research quarterly for exercise and sport 51*, 1 (1980), 158–173.
45. Vatavu, R.-D., Anthony, L., and Wobbrock, J.O. Gestures as point clouds: a $P recognizer for user interface prototypes. *Proc. ICMI 2012*, ACM Press (2012), 273–280.
46. Vatavu, R.D., Vogel, D., Casiez, G., and Grisoni, L. Estimating the perceived difficulty of pen gestures. *Proc. INTERACT 2011*, Springer (2011), 89–106.
47. Vogel, D. and Balakrishnan, R. Direct Pen Interaction with a Conventional Graphical User Interface. *Human–Computer Interaction 25*, 4 (2010), 324–388.