Examining the Need for Visual Feedback during Gesture Interaction on Mobile Touchscreen Devices for Kids

Lisa Anthony[1], Quincy Brown[2], Jaye Nias[2], Berthel Tate[2]

[1]UMBC
Information Systems
1000 Hilltop Circle
Baltimore MD 21250 USA
lanthony@umbc.edu

[2]Bowie State University
Computer Science Department
14000 Jericho Park Road
Bowie MD 20715 USA
qbrown@bowiestate.edu

ABSTRACT
Surface gesture interaction styles used on modern mobile touchscreen devices are often dependent on the platform and application. Some applications show a visual trace of gesture input as it is made by the user, whereas others do not. Little work has been done examining the usability of visual feedback for surface gestures, especially for children. In this paper, we present results from an empirical study conducted with children, teens, and adults to explore characteristics of gesture interaction with and without visual feedback. We find that the gestures generated with and without visual feedback by users of different ages diverge significantly in ways that make them difficult to interpret. In addition, users prefer to see visual feedback. Based on these findings, we present several design recommendations for new surface gesture interfaces for children, teens, and adults on mobile touchscreen devices. In general, we recommend providing visual feedback, especially for children, wherever possible.

Categories and Subject Descriptors
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

Keywords
Gesture interaction, surface gestures, touch interaction, interaction design, empirical studies, mobile devices, touchscreens, child-computer interaction.

1. INTRODUCTION
Touch interaction on modern mobile devices such as smartphones and tablet computers has become one of the most prevalent modes of interaction with technology for many users. These devices all support some form of surface gesture interaction, but the interaction styles used are often dependent on the platform and application. While some gestures have emerged as cross-platform standards, such as swipe, pinch-to-zoom, and drag-to-pan, there is still quite a variety of other gestures in use for specific apps. For example, the note-taking and sketching app from FiftyThree, Inc., called Paper[1], uses a counter-clockwise spiral gesture to "rewind" (e.g., undo) the user's command history. Another example is Realmac Software's Clear[2] list-keeping app, which uses a drag-and-hold gesture to create a new list item. Anthony et al. [3] and Zhai et al. [36] both include summaries of the range of gestures used in research on surface gesture interaction. In addition to using a variety of gestures, some applications show a visual trace of gesture input as it is made by the user, such as drawing / tracing games (e.g., Luck-u's Art Penguin[3]), whereas others do not, such as navigation apps (e.g., Ulmon GmbH's City Maps 2Go[4]). Evidence from cognitive and perceptual psychology literature suggests that both children and adults have more difficulty drawing and writing in the absence of visual feedback [8,32,34]. Young children may benefit even more strongly from the use of visual feedback during interaction because they are still developing the required sensorimotor coordination ability to draw without looking [34]. However, little work has been done to examine the usability of visual feedback for surface gestures in general, let alone for children or teens.

(a) with feedback (b) without feedback
Figure 1. Examples of gestures produced with and without visual feedback by one participant (a child, to scale).

In this paper, we present results from an empirical study conducted with 41 children, teens, and adults to explore characteristics of gesture interaction with and without visual feedback (Figure 1). We asked questions such as: How well are children, teens, and adults able to enter surface gestures with and without visual feedback? Does presence or absence of visual feedback affect consistency of the gestures made? If so, is automatic gesture recognition impacted by these changes in consistency? Which mode of gesture input (with or without feedback) do adults, teens, and children prefer?

IDC '13, June 24 - 27 2013, New York, NY, USA

[1] https://itunes.apple.com/us/app/paper-by-fiftythree/id506003812
[2] https://itunes.apple.com/us/app/id493136154
[3] https://itunes.apple.com/us/app/id449097181
[4] https://itunes.apple.com/us/app/city-maps-2go/id327783342

We find that the gestures generated by users of different ages with and without visual feedback diverge significantly in ways that make them difficult to interpret. For example, users tend to make gestures with fewer strokes in the absence of visual feedback. They also tend to make shorter, more compact gestures in the presence of visual feedback. In addition, based on our observations, users of all age groups we studied prefer to see visual feedback, although adults are more willing to accept lack of feedback. Based on these findings, we present several design recommendations for new surface gesture interfaces for children, teens, and adults on mobile touchscreen devices.

The contributions of this work include the following. First, we present an analysis of gesture features that change when visual feedback is present or absent during the interaction, which can be used to design better gesture sets and recognizers for one or both situations. Second, we analyze the actual impact of these feature differences on recognition by current algorithms used in user interaction research. Third, we present design recommendations based on empirical data we collected from children, teens, and adults to examine the necessity, utility, and desirability of visual feedback. The results of this work are informative to designers and researchers interested in surface gesture interaction on mobile devices for users of all ages.

2. RELATED WORK

We briefly survey related work on surface gesture interaction on mobile devices for both children and adults, as well as prior work on usability of interactions with and without visual feedback.

2.1 Surface Gestures and Mobile Devices

Gesture-based interaction on touch-enabled surfaces have been studied extensively in the HCI literature, particularly from a usability perspective [12,14,20,21,25,28,30,33,35]. Gesture set design [25,35], multitouch gestures [12,20], accessible gestures [21], and differences between pen/stylus and finger gesture input [33] are just some of the areas that have been examined, but none of these studies have included children. From a child-computer interaction perspective, surface gestures for children especially on mobile devices have generally been neglected. Multitouch gestures for children on tabletops have been explored [14,28,30], but research typically has either included children only, or has not distinguished between adults and children, making the comparisons needed for tailored interaction design difficult. Some work recently has explicitly compared and contrasted surface-gesture interaction design for children and adults [2,6,15], but has not specifically looked at the question of feedback. As we continue to see an increase in the use of touch-based mobile technologies by children [31], further work in this area is needed.

Related work in pen-based handwriting interactions for children [27], pointing and mouse pathing interactions for children [11,16,19,29], and drag-and-drop gestures (with mice or fingers) for children [7,17,18] have found that children make less stable movements, have difficulty maintaining contact with the screen, and make more input errors overall than do adults. We predict that similar results will hold for other types of surface gestures performed on mobile touchscreen devices, and we explore this relationship in our own work.

2.2 Usability and Visual Feedback

Many researchers have examined the use of visual feedback (among other types of feedback) for various modalities such as pointing with a mouse [1], text entry [9], 3D gestures [22], and hand-tracking gestures [24]. In these cases, visual feedback is usually found to be necessary to allow users to understand that their input has had the desired effect. In Clawson et al.'s work [9], however, the visual feedback that was preferred by users during mobile text entry had the side effect of decreasing typing speed, because visible input errors distracted users. Two examples of work that explicitly seeks to reduce reliance on visual feedback are Gustafson's [13] "imaginary interfaces," which uses accelerometer-based gestures on screen-less devices, and Zhao et al.'s [37] *EarPod*, an eyes-free menu selection technique that uses auditory rather than visual feedback. In both cases, the benefit of eyes-free interaction trades off with a new burden on the user to recall required input actions without visual confirmation of their successful interaction.

Very little work has explored the use of visual feedback for touch and gesture interaction. One example is Li's [23] *GestureSearch* tool, which accepts letter gestures as shortcuts for searching, e.g., to jump to a particular alphabetic section of one's contact list. In that work, users prefer character-based gesture shortcuts for commands due to the mode switch required by text entry on mobile devices. Gesture interaction differs from other modalities in that it can support two types of visual feedback: visual feedback of the actual action being entered (e.g., the trace of a gesture), and visual feedback of the action's effect (e.g., the recognition of a gesture). Work on visual feedback in other modalities can provide design recommendations for the latter type of visual feedback. We are the first to examine the former type.

In addition, none of these studies in any modality has involved child users. Based on child development literature (e.g., [34]), we hypothesize that providing visual feedback will be even more crucial for gesture interaction design for children than for adults since children are still developing the sensorimotor coordination ability required to draw without looking.

3. EXPERIMENT METHOD

We conducted an empirical study with children, teens, and adults using mobile devices in a laboratory setting. Further work in this area will investigate more natural interaction outside of the laboratory, but for these initial explorations into the effect of visual feedback on interaction, collecting robust input data of specific types is necessary. We describe here the tasks performed by the participants and how visual feedback varied.

3.1 Participants

A total of 41 participants (25 children/teens and 16 adults) participated in this study (20 were female). A demographic breakdown of the participants is given in Table 1. In general, the children in this study were pre-teens and teens. In future work, we plan to investigate younger children as well. Furthermore, most of the participants in our sample had experience using touchscreen devices such as smartphones and tablets and rated themselves either "average" or "expert" on a demographic questionnaire. In future work, we plan to investigate less expert users.

3.2 Equipment

We used Samsung Google Nexus S smartphones running the Android 4.0.4 operating system to conduct the experiment. The phones' dimensions were 4.88" x 2.48" x 0.43", and had 4" screens, measured diagonally. Display resolution was 480 x 800 pixels (233 pixels per inch (ppi) pixel density). We created our own apps for this platform that enabled us to log all input events generated by the participants during the study session.

	Overall	Children / Teens	Adults
N	41	25	16
Gender	20 female (49%)	14 female (56%)	6 female (38%)
Age (yrs)	*M* = 17.5 *Min* = 10 *Max* = 33 *SD* = 6.6	*M* = 12.8 *Min* = 10 *Max* = 17 *SD* = 1.8	*M* = 24.8 *Min* = 20 *Max* = 33 *SD* = 4.2
Grade Levels	n/a	5th to 11th	n/a
Handedness	85% right	88% right	81% right
Expertise (self-report)	0% beginner 39% average 59% expert	0% beginner 44% average 52% expert	0% beginner 31% average 69% expert

Table 1. Demographic information for the 41 participants.

3.3 Procedure

Participants came to a research laboratory to participate in the study. Up to three people could participate at one time (children/teens or adults only, no mixed sessions were conducted). Sessions lasted about one hour. During these sessions, participants engaged in a variety of tasks on mobile touchscreen devices. For the purposes of this paper, we present only one of these: the Gesture Task, which included both a Feedback and a No-Feedback condition, indicating the presence or lack of visual feedback. Participants were compensated $10.

3.4 Gesture Input Task

During the study, the participants drew gestures onscreen with their finger one at a time. There were 20 gestures used in the study, created based on existing mobile device apps as well as educational psychology literature about developmentally-appropriate gestures for children [5]. (A similar task and gesture set has been used in prior work on gesture interaction for children [2].) The gesture set (Figure 2) included letters, numbers, symbols, and geometric shapes[5]. Participants saw a prompt onscreen as to which gesture to enter (Figure 3a). To test the impact of visual feedback on gesture input, we included both a Feedback and a No-Feedback condition. In the *Feedback* condition, a trace was shown as the participant gestured (Figure 3b), but in the *No-Feedback* condition, no trace was shown. After entering the gesture, the participant touched the onscreen "Done" button to move on to the next gesture.

During the study, participants sat at a table in the lab and were allowed to hold the phone in a manner comfortable to them (e.g., handheld, resting on the table, etc.). Before doing the gesture input task on the phone, participants drew one sample of each gesture by hand on a sheet of paper. This activity helped ensure all of the gestures were familiar to the participants by name, since the app's prompt was textual (Figure 3). The app prompted the participant to enter one example of each gesture in the set one at a time, and then repeated this five times, yielding a total of 120 gesture samples (240 across both conditions per participant). Order of presentation of Feedback and No-Feedback tasks was

[5] Command gestures such as swipe and pinch-to-zoom were not included for two reasons: (1) studies find these gestures are difficult for children [7], and (2) many current children's educational apps use tracing or handwriting activities [2].

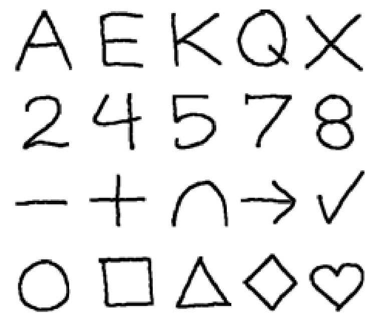


Figure 2. The set of 20 gestures used in our study, which we borrow from prior work on gesture interaction for kids [2].

counterbalanced across sessions (all participants in any one session completed them in the same order).

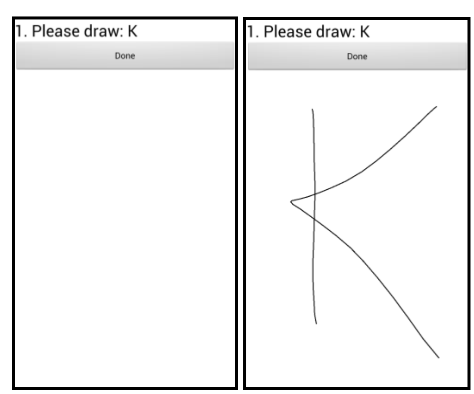
3.5 Measures

As participants dragged their finger across the device screen to enter each gesture, touch events were registered by the hardware and recorded by our app software for later data analysis. These touch events include information such as the *x*-coordinate, *y*-coordinate, timestamp, touch pressure, and touch size of each event. A *gesture* might consist of multiple strokes; one *stroke* consists of all touch events registered between the time a finger-down and a finger-up event are registered. We used these data to calculate geometric properties of the gestures that were generated by each user, as well as to feed the stroke data into gesture recognition software to analyze recognition accuracy.

4. ANALYSIS AND RESULTS

In our study, we collected 9840 gestures across 41 participants. The first round of gestures in each condition was considered practice, and therefore is not included in our analysis, leaving a total of 8200 gestures. In our work, we typically consider 5 sub-groups of children/teens, based on age: 0 to 4 years, 5 to 7 years, 8 to 10 years, 11 to 13 years, and 14 to 17 years. These groups are derived from the following sources:

- developmental psychology literature (e.g., Piaget [26]);
- typical school age groupings in the United States (e.g., elementary school (5 to 10 years), middle school (11 to 13



(a) (b)

Figure 3. Screenshot of the study app's interface: (a) before drawing the gesture, and (b) after drawing the gesture, Feedback condition. (In No-Feedback, after drawing the gesture, the screen looked the same as (a).)

Gesture Feature	How Computed
No. strokes (S)	Total number of finger-down to finger-up periods registered during a gesture.
No. points (N)	Total number of touch events registered during a gesture, cumulatively over all strokes.
Gesture length	Cumulative path distance from the first touch event registered for the gesture to the last.
Gesture height	Height of the smallest bounding box that contains the gesture ($max_y - min_y$).
Gesture width	Width of the smallest bounding box that contains the gesture ($max_x - min_x$).
Gesture area	Gesture height * Gesture width.
Gesture duration	Time elapsed while drawing the gesture, e.g., time of the last touch event registered for the gesture minus time of the first touch event, including breaks between strokes (milliseconds, or ms).
Gesture pressure	Average pressure registered over all the touch events belonging to a gesture (pressure / N).
Gesture speed	Average speed registered over all the touch events belonging to a gesture (duration / length).

Table 2. Gesture features analyzed in this paper.

years), and high school (14 to 17 years)); and,
- our experience conducting research with children and teens.

In this study, children as young as 10 years of age participated, so our analyses are based on the following groups: 10 years (2 children), 11 to 13 years (16 children), 14 to 17 years (7 children), and adults (18+ years, 16 adults).

4.1 Gesture Features

Table 2 shows a list of the nine gesture features we analyzed in this paper: (a) *Number of (No.) strokes*, (b) *Number of (No.) points*, (c) *Gesture length*, (d) *Gesture height*, (e) *Gesture width*, (f) *Gesture area*, (g) *Gesture duration*, (h) *Gesture pressure*, and (i) *Gesture speed*; as well as how they were computed. These features are geometric features that may be expected to impact recognition accuracy by making the gestures 'look' different to the recognizer. While this list is by no means exhaustive, we believe it covers the most commonly used features and those most likely to affect interpretation of gesture input. Also, they have been used in similar prior work on gesture interaction, especially for children [6].

We analyzed each feature for gestures created by children and adults in the presence and absence of visual feedback. We conducted a series of ANOVA tests to determine where differences may lie for each feature. In all cases, we conducted a univariate ANOVA with *participant age group* and *visual feedback?* as fixed factors. Because each participant entered multiple gestures, we included *participant* as a random factor[6]. Because the study design was nested (e.g., participants could only be in one age group), we constructed a model with the main effect terms for *participant age group* and *visual feedback?*, a nested term for *participant*(*participant age group*), and an interaction term for *participant age group* x *visual feedback?*. Table 3 summarizes the significant effects; specific findings for each feature are discussed below.

No. strokes. The number of strokes showed no significant differences based on *participant age group* ($F_{3,37} = 1.00$, *n.s.*), *visual feedback?* ($F_{1,8154} = 3.72$, *n.s.*), or their interaction ($F_{3,8154} = 0.42$, *n.s.*). Because the interaction was not significant, we re-ran the ANOVA without the interaction in the model; this time there was a significant main effect of *visual feedback?* ($F_{1,8157} = 9.01$, $p < .01$). Thus, although the number of strokes was not a distinctive feature between age groups, it was responsive to the presence or absence of visual feedback. Users tended to generate gestures with fewer strokes with no visual feedback.

No. points. The number of points sampled during a gesture showed a significant interaction between *participant age group* and *visual feedback?* ($F_{3,8154} = 15.47$, $p < .01$). We see a shift from the younger children to the adults, in which 10 year olds and 11 to 13 year olds tend to make gestures more quickly in the *absence* of visual feedback, but 14 to 17 year olds show no difference, and adults tend to make them more quickly in the *presence* of visual feedback.

Gesture length. There was a significant interaction between *participant age group* and *visual feedback?* ($F_{3,8154} = 4.03$, $p < .01$). All age groups tended to make shorter (length) gestures in the presence of visual feedback, but the youngest children and adults showed a smaller difference than the middle age groups.

Gesture height. There was a significant interaction between *participant age group* and *visual feedback?* ($F_{3,8154} = 15.14$, $p < .01$), in which most age groups made shorter (height) gestures in the presence of visual feedback, except the 10-year-olds.

Gesture width. The width of the gestures generated showed a significant interaction between *participant age group* and *visual feedback?* ($F_{3,8154} = 14.10$, $p < .01$). Most age groups made narrower gestures in the presence of visual feedback, but what varied was the degree of difference (smaller for adults).

Gesture area. As gesture area is a composite of gesture height and gesture width, it is perhaps unsurprising that this feature

[6] A *random* factor's levels have been chosen at random and might change when doing the study again (e.g., participants drawn from the population). It is an accepted practice to use participant as a random factor for repeated measures when the number of samples per participant is very many or not equal [10] (p. 630).

Gesture Feature	Significant Results *Participant Age Group* Main Effect	 *Visual Feedback?* Main Effect	 *Participant Age Group* x *Visual Feedback?* Interaction
No. strokes	-	-	-
No. points	-	-	++
Gesture length	-	++	++
Gesture height	-	++	++
Gesture width	-	++	++
Gesture area	-	++	++
Gesture duration	-	++	++
Gesture pressure	-	++	++
Gesture speed	-	++	++

Table 3. Significant interactions and main effects for each gesture feature of *participant age group* and *visual feedback?*. + indicates *p* < .05 and ++ denotes *p* < .01, - indicates *p* > .05.

showed the same relationship again: a significant interaction was found between *participant age group* and *visual feedback?* ($F_{3,8154} = 25.84$, $p < .01$).

Gesture duration. The amount of time taken to draw a gesture showed a significant interaction between *participant age group* and *visual feedback?* ($F_{3,8154} = 5.80$, $p < .01$). The younger children (10 year olds and 11 to 13 year olds) tended to take more time to draw gestures in the presence of visual feedback vs. absence of visual feedback than did 14 to 17 year olds and adults.

Gesture pressure. The average pressure exerted by the participant's finger during a gesture also showed a significant interaction between *participant age group* and *visual feedback?* ($F_{3,8154} = 31.75$, $p < .01$). All age groups exerted less pressure in the presence of visual feedback, but this difference was more pronounced for the 10-year-olds.

Gesture speed. The average speed of a gesture is related to the length and the duration, so unsurprisingly, this feature showed the same relationship: a significant interaction was found between *participant age group* and *visual feedback?* ($F_{3,8154} = 26.03$, $p < .01$). All age groups tended to draw gestures faster in the absence of visual feedback, but for adults, this effect was less pronounced.

Age Group	Condition	Mean	SD	N
10 yrs	**Feedback**	77.1%	7.8%	2
	No Feedback	77.6%	6.4%	2
11 to 13	**Feedback**	80.7%	7.8%	16
	No Feedback	84.2%	5.4%	16
14 to 17	**Feedback**	87.6%	5.4%	7
	No Feedback	87.2%	7.8%	7
Adults (18+)	**Feedback**	90.8%	6.4%	16
	No Feedback	90.8%	4.9%	16

Table 4. Recognition accuracy of $N-Protractor on gestures by age group and presence or absences of visual feedback.

All of these features show a general trend that users are more careful when generating gestures in the presence of visual feedback (shown by gesture duration and gesture speed). This behavior could be due to an increased fluency in entering gestures when users can see what they are doing; the visual feedback provides a "check" on the sensorimotor feedback they get while drawing a gesture, increasing confidence. The generation of gestures with fewer strokes in the absence of visual feedback was unexpected, but when one considers the challenge of joining strokes without visual feedback once a finger has been lifted, it becomes more clear. Finally, the shorter and more compact gestures that were made in the presence of visual feedback could also be due to a fluency effect: being able to see one's trace visually can increase confidence in finer-grained movements.

In most cases, the younger the children, the greater degree of variation they exhibited between gestures they created in the presence or absence of visual feedback. This effect is a strong indicator that children struggle without visual feedback. However, even adults show some variation between the two cases, and as a result, we recommend that visual feedback always be provided during surface gesture interaction with these types of input gestures. Users of all ages can benefit from this accommodation.

4.2 Gesture Recognition

Differences in how children and adults make gestures in the presence or absence of visual feedback may not actually be relevant to the design of gesture interaction on mobile devices if it is equivalently easy (or difficult) to recognize these gestures. To understand the impact on recognition of the significant differences we discovered in these gesture features, we ran the gestures through the $N-Protractor recognizer [4], an accurate, open-source, trainable recognizer used by gesture interaction researchers and mobile app developers. We conducted *user-dependent* training, in which we first trained the gesture recognizer on a small set of one user's gestures (evenly sampling from all gesture types) and then tested the recognizer on the remainder of that user's samples. We repeated this procedure for all users, computing average per-user recognition accuracy for each of the age groups. Note that we conducted separate tests for gestures generated in the presence of visual feedback and those generated in its absence. Table 4 shows all recognition results by age group separated by presence or absence of visual feedback.

We conducted a repeated-measures ANOVA on the per-user recognition accuracy on the within-subjects factor of *visual feedback?* and the between-subjects factor of *participant age group*. We found no significant difference in accuracy based on presence or absence of visual feedback ($F_{1,37} = 0.56$, *n.s.*), nor was the interaction with participant age group significant ($F_{3,37} = 1.45$, *n.s.*), but we did find a significant main effect of participant age group alone ($F_{3,37} = 7.38$, $p < .01$).

Supporting prior work on recognition of children's gestures with current recognizers [2,6], we find that recognizing children's gestures is still harder than recognizing adults' gestures. The younger the children are, the lower recognition accuracy is. However, we do not find support for the hypothesis that recognition accuracy would be lower for gestures made in the absence of visual feedback, in spite of the differences already noted in the features that make up these gestures. This finding

	Adults	Children
Presence of Visual Feedback	Triangle A E	Diamond Triangle Arrowhead
Absence of Visual Feedback	Triangle Rectangle Arrowhead	Rectangle Triangle A

Table 5. Top three worst recognized gestures for children and adults, with and without visual feedback.

could be due to a lack of correlation between the features that differed and the features that the recognizer uses to classify gestures. Further exploration of other gesture features or other gesture recognizers would help continue to characterize this space. In the meantime, it seems that the significant differences in gesture generation that we have found do not reliably impact gesture recognition. This finding brings the verdict on the question "feedback or no feedback?" back to neutral, because interaction success from the system's standpoint is not affected.

4.3 Qualitative Observations

We observed anecdotally that participants seemed confused by the absence of visual feedback while they were performing the gesture task. Participants commented that they could not see their finger markings to help them enter the gestures, and that they didn't like not being able to see what was being drawn. These comments were especially common if the participants had done the Feedback condition first. We noted that several of the 14 to 17 year old participants expressed pleasure when they could see their gestures and were happy with how they appeared. This effect diminished in later sessions when we warned participants that they would not see their gestures some of the time, but the gestures they created in the absence of visual feedback remained poorer from a qualitative perspective (as well as the quantitative feature differences already discussed). Figure 4 shows a few comparative examples of gestures drawn by children in the Feedback and No-Feedback conditions. Although these differences did not cause a decrease in recognition rates (probably because the recognizer was trained on No-Feedback gestures before being tested on them, and the same was true for Feedback), they are clearly visually different, demonstrating the gesture feature differences we found (e.g., gesture length, height, width, etc.), as well as others (e.g., corner joining, line straightness).

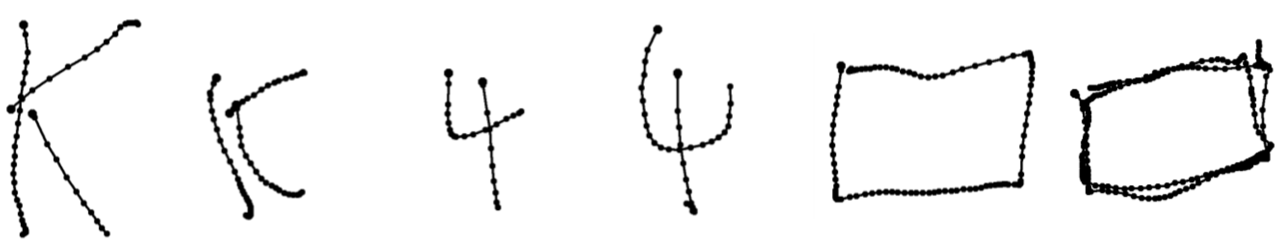
Although eventually all users completed all gestures in both the presence and absence of visual feedback, users' frustration in the latter case did not seem to diminish as the study session went on. Especially given the lower quality of gestures in the No-Feedback condition, we suggest that these observations indicate that gesture interaction without visual feedback does not feel comfortable to users and is not recommended.

4.4 Discussion

In general, we have found evidence that both supports and contradicts the use of visual feedback during gesture interactions. In terms of geometric properties of the gestures drawn by our participants, all age groups (children, teens, and adults) made different gestures in the presence of visual feedback than they did in its absence. However, these differences did not impact the ability of a gesture recognizer ($N-Protractor [4] in this case) to classify the gestures. When we consider users' opinions on using interfaces with and without visual feedback, though, we find further reason to recommend the inclusion of visual feedback.

We here consider in more detail why, in spite of the differences in gesture features, recognition tests did not show an effect of presence or absence of visual feedback. First, we examine which gestures were the most challenging for the recognizer in our tests. Table 5 lists the three gestures that were the least accurately recognized when considering *participant age group* (just adults vs. children) and *visual feedback?*. The lists are very similar, which could explain why, even though the gestures were made with inconsistent features, the recognizer had a consistent amount of difficulty with classification. We suggest that these gestures are either particularly difficult for participants to draw, or they are particularly challenging in general for $N-Protractor. Anecdotally, we did observe during the study that gestures like the arrowhead and diamond were not as familiar to the participants as the other types of gestures. Furthermore, the recognizer could be expected to have difficulty distinguishing between triangles and diamonds, which are very similar geometrically. More exploration of the types of errors made by the recognizer (e.g., which gestures are confused for each other most often) is needed to answer this question sufficiently. With such analysis, it could be possible to design gesture sets to ensure consistency by users, and to use only gestures that are well-recognized by the system.

Also, some prior work has examined differences between features of surface gestures generated by children and adults, in one case finding a difference [6], and in another case not finding one [2]. We believe the work we report in this paper can settle the discrepancy between these two prior studies. In the study by Anthony et al. [2], in which no gesture feature differences were found between adults and children for a similar gesture input task, only a *feedback* condition was tested. In the study by Brown et al. [6], in which differences *were* found (number of strokes, gesture height, gesture duration, and gesture pressure), they used a gesture input task with *no visual feedback*. Neither of these studies tested both the presence and absence of visual feedback, as we have done here. When we consider both of these similar prior studies and the interactions we have found between *participant age group* and *visual feedback?* in this study, we can conclude that the primary factor contributing to gesture



(a) with feedback (b) without feedback (c) with feedback (d) without feedback (e) with feedback (f) without feedback
Figure 4. Examples of gestures produced with and without visual feedback by three different children (to scale).

generation differences among children and adults is whether or not there is visual feedback provided. When visual feedback is used, participants are more comfortable and generate more consistent gestures. When it is not used, participants' input behaviors are less consistent, and this effect is magnified for children over adults. Therefore, we believe that the cumulative evidence across these three studies favors use of visual feedback for these types of gestures during surface gesture interaction.

5. DESIGN IMPLICATIONS

Based on the findings from the study presented in this paper, we outline three new design recommendations for surface gesture interaction on mobile devices for children, teens, and adults.

DO provide visual feedback for surface gesture interaction on mobile devices. We found evidence that users' gestures are made differently in the presence than in the absence of visual feedback. Although in this study it did not impact recognition results, users expressed dissatisfaction with surface gesture interactions without visual feedback. Allowing users to see the trace of their finger's path along the device screen can improve carefulness and confidence in their input. Although this recommendation can improve interaction for users of all ages, it is particularly relevant to interaction design for children. Children's mental agility in imagining their finger's path is less well-developed than that of adults, and therefore visual feedback can aid them in developing this hand-eye coordination skill as they mature.

DON'T include gestures unfamiliar to users. When designing gesture sets for new applications, it is risky to use new gestures that users may not already know how to draw. More commonly used shapes that users encounter outside of their interactions with a given application will be more comfortable for them, increasing the consistency with which they generate gestures. In turn, these gestures will be more easily recognized by the system. This consideration applies to users of all ages, but is especially critical for interaction design for children. Children have less experience with technology, less schooling and exposure to the range of possible letters and shapes [5], and less developed fine-motor control, which impacts the dexterity of this population. Designers of application gesture sets should consider both the requirements of the algorithms along with the cognitive abilities of their users.

DO test new gesture sets with the target recognizer in advance. When designing gesture-based interaction, the recognition approach can make a difference in how well users' gestures are understood. We have tested just one common approach, $N-Protractor [4]. It may be the case that other recognizers will have less difficulty with some of the basic shapes tested in this study that $N-Protractor classified more poorly (e.g., triangle, diamond, rectangle). A key design recommendation for surface gesture interaction, especially with children, is to use iterative rapid prototyping that can expose conflicts (either from the user's or system's perspective) in the gesture set early.

6. FUTURE WORK

This work is the first to explore the impact of visual feedback on surface-gesture input for children, teens, and adults, and as such represents a foundational study in this space. Many other factors may also be relevant to successful gesture-based interaction design for children, and we briefly list a few that we have identified as promising areas of future work. First, we have included a wide age range of children in this study, from 10 to 17 years old (and adults from 20 to 33 years old). This work characterizes the impact of visual feedback on gesture generation for older children who are fairly comfortable with writing and drawing activities, and it may be informative to extend this work to younger children who are just starting out in school (ages 5 to 9) or even pre-school-aged children (ages 1 to 4). We anticipate that the impact of visual feedback will be more pronounced for these younger children. We also think that validating these results with children, teens, and adults of varying levels of experience with mobile touchscreen devices and gesture interaction will also be important to fully explore this space.

Second, we have examined a fairly abstracted task, in which the participants were entering samples of the gestures without a goal for using that gesture to do anything (e.g., to launch a task or respond to a query). We do not yet know how a change in the user's goal might interact with the user's input with or without visual feedback. In some handwriting practice activity apps for children that exist today (e.g., Jaloby's AlphaCount[7]), the interface may be only a little more embellished than our app to prompt the child for a gesture to draw. Thus, we believe that this abstracted task makes a good foundation, and plan to extend it to contextualized tasks in future work. We expect to see similar patterns, but predict a decrease in the impact of the absence of visual feedback for tasks where there is important information onscreen that the gesture might otherwise obscure.

7. CONCLUSION

We have presented the results of an empirical study examining the impact of the presence or absence of visual feedback during surface gesture input on a mobile device. Our findings from data from 41 children, teens, and adults indicate that gestures generated with and without visual feedback differ significantly in ways that make them difficult to interpret. In addition, users of all age groups we studied prefer to see visual feedback, although adults are more willing to accept lack of feedback. Based on our findings, we present design recommendations for new surface gesture interfaces for children, teens, and adults regarding the use of visual feedback. The results of this work will be informative to designers and researchers interested in surface gesture interaction on mobile devices for all ages.

8. ACKNOWLEDGMENTS

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. We would also like to thank our collaborators at Bowie State and UMBC, including Robin Brewer, Thaddeus Brown, Germaine Irwin, Shreya Mohan, Monique Ogburn, and Chiamaka Okorohoa for project support.

9. REFERENCES

1. Akamatsu, M., MacKenzie, I.S., and Hasbroucq, T. A comparison of tactile, auditory, and visual feedback in a pointing task using a mouse-type device. *Ergonomics 38*, 4 (1995), 816–827.
2. Anthony, L., Brown, Q., Nias, J., Tate, B., and Mohan, S. Interaction and Recognition Challenges in Interpreting

[7] https://itunes.apple.com/us/app/alphacount/id359046783

Children's Touch and Gesture Input on Mobile Devices. *Proc. ITS 2012*, ACM Press (2012), 225–234.

3. Anthony, L., Vatavu, R.-D., and Wobbrock, J.O. Understanding the Consistency of Users' Pen and Finger Stroke Gesture Articulation. *Proc. GI 2013*, Canadian Information Processing Society (2013), to appear.
4. Anthony, L. and Wobbrock, J.O. $N-protractor: a fast and accurate multistroke recognizer. *Proc. GI 2012*, Canadian Information Processing Society (2012), 117–120.
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. Brown, Q. and Anthony, L. Toward Comparing the Touchscreen Interaction Patterns of Kids and Adults. *ACM SIGCHI EIST Workshop 2012*, (2012), 4pp.
7. 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.
8. De Bruyn, D. and Davis, A. Visual feedback is not essential for children to make the perpendicular bias. *Perception ECVP Abstract Supplement 37*, (2008), 149.
9. Clawson, J., Lyons, K., Starner, T., and Clarkson, E. The Impacts of Limited Visual Feedback on Mobile Text Entry for the Twiddler and Mini-QWERTY Keyboards. *Proc. ISWC 2005*, IEEE (2005), 170–177.
10. Dean, A. and Voss, D. *Design and Analysis of Experiments*. Springer-Verlag, New York, NY, USA, 1999.
11. 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.
12. Frisch, M., Heydekorn, J., and Dachselt, R. Investigating multi-touch and pen gestures for diagram editing on interactive surfaces. *Proc. ITS 2009*, ACM Press (2009), 149–156.
13. Gustafson, S. Imaginary interfaces: touchscreen-like interaction without the screen. *Ext. Abstracts CHI 2012*, ACM Press (2012), 927–930.
14. 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.
15. Hinrichs, U. and Carpendale, S. Gestures in the wild. *Proc. CHI 2011*, ACM Press (2011), 3023–3032.
16. 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.
17. Inkpen, K.M. Drag-and-drop versus point-and-click mouse interaction styles for children. *ACM ToCHI 8*, 1 (2001), 1–33.
18. 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.
19. Jones, T. An Empirical Study of Children's Use of Computer Pointing Devices. *Journal of Educational Computing Research 7*, 1 (1991), 61–76.
20. Kammer, D., Wojdziak, J., Keck, M., Groh, R., and Taranko, S. Towards a formalization of multi-touch gestures. *Proc. DIS 2010*, ACM Press (2010), 49–58.
21. Kane, S.K., Wobbrock, J.O., and Ladner, R.E. Usable gestures for blind people: understanding preference and performance. *Proc. CHI 2011*, ACM Press (2011), 413–422.
22. Kratz, S. and Ballagas, R. Unravelling seams: improvoing mobile gesture recognition with visual feedback techniques. *Proc. CHI 2009*, ACM Press (2009), 937–940.
23. Li, Y. Gesture search: a tool for fast mobile data access. *Proc. UIST 2010*, ACM Press (2010), 87–96.
24. Lin, E., Cassidy, A., Hook, D., and Baliga, A. Hand tracking using spatial gesture modeling and visual feedback for a virtual DJ system. *Proc. ICMI 2002*, IEEE (2002), 197–202.
25. Morris, M.R., Wobbrock, J.O., and Wilson, A.D. Understanding users' preferences for surface gestures. *Proc. GI 2010*, Canadian Information Processing Society (2010), 261–268.
26. Piaget, J. Piaget's Theory. In P. Mussen, ed., *Handbook of Child Psychology*. Wiley & Sons, New York, NY, USA, 1983.
27. 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.
28. 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.
29. 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.
30. 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.
31. Shuler, C. *Pockets of Potential: Using Mobile Technologies to Promote Children's Learning*. Joan Ganz Cooney Center at Sesame Workshop, New York, NY, 2009.
32. Smyth, M.M. Visual control of movement patterns and the grammar of action. *Acta psychologica 70*, 3 (1989), 253–265.
33. Tu, H., Ren, X., and Zhai, S. A comparative evaluation of finger and pen stroke gestures. *Proc. CHI 2012*, ACM Press (2012), 1287–1296.
34. Vinter, A. and Meulenbroek, R. The role of manual dominance and visual feedback in circular drawing movements. *Journal of Human Movement Studies 25*, (1993), 11–37.
35. Wobbrock, J.O., Morris, M.R., and Wilson, A.D. User-defined gestures for surface computing. *Proc. CHI 2009*, ACM Press (2009), 1083–1092.
36. Zhai, S., Kristensson, P.O., Appert, C., Anderson, T.H., and Cao, X. Foundational Issues in Touch-Surface Stroke Gesture Design — An Integrative Review. *Foundations and Trends in Human–Computer Interaction 5*, 2 (2012), 97–205.
37. Zhao, S., Dragicevic, P., Chignell, M., Balakrishnan, R., and Baudisch, P. Earpod: eyes-free menu selection using touch input and reactive audio feedback. *Proc. CHI 2007*, ACM Press (2007), 1395–1404.