Understanding Child-Defined Gestures and Children's Mental Models for Touchscreen Tabletop Interaction

Karen Rust[1], Meethu Malu[2], Lisa Anthony[3], Leah K. Findlater[1]

[1] College of Information Studies | HCIL
[2] Department of Computer Science | HCIL
University of Maryland, College Park, MD
{k579, meethu, leahkf}@umd.edu

[3] Department of CISE
University of Florida
Gainesville, FL USA 32611
lanthony@cise.ufl.edu

ABSTRACT
Creating a pre-defined set of touchscreen gestures that caters to all users and age groups is difficult. To inform the design of intuitive and easy to use gestures specifically for children, we adapted a user-defined gesture study by Wobbrock et al. [12] that had been designed for adults. We then compared gestures created on an interactive tabletop by 12 children and 14 adults. Our study indicates that previous touchscreen experience strongly influences the gestures created by both groups; that adults and children create similar gestures; and that the adaptations we made allowed us to successfully elicit user-defined gestures from both children and adults. These findings will aid designers in better supporting touchscreen gestures for children, and provide a basis for further user-defined gesture studies with children.

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

General Terms
Design, Experimentation, Human Factors.

Keywords
Children, interactive tabletop, touchscreens, user-defined gestures.

1. INTRODUCTION
Interactive touchscreens have become ubiquitous. In the United States, an estimated eight out of ten children use mobile devices regularly [7]. Larger touchscreens such as interactive tabletops are also being adopted, particularly in educational contexts such as classrooms [11] and museums [6]. However, the vast majority of research on touchscreen gestural interaction has focused on adults, for example, investigating user consistency [2], user preferences [9], and gesture intuitiveness [12]. To date, research on children's touchscreen interaction is limited, despite unique challenges that children encounter both in performing gestures and in having gestures registered and interpreted correctly by these devices [1].

A common method for gaining insight into users' mental models and preferences with respect to gestures is the *user-defined gesture* approach popularized by Wobbrock et al. [12]. In their study, adult participants created one-handed and two-handed gestures for *referents* (actions) on an interactive tabletop (e.g., opening an application, deleting an object). Participants were first shown an effect (a video clip of a change in the interface) and then asked to create a gesture that would cause that effect. Analyzing these gestures and the creation process, exposed through a think-aloud protocol, can provide insight into gesture mechanics, mental models, and user preferences. The user-defined gestures method yields gestures that are preferred to those created by designers [9]. However, this method has only been employed with children to a limited extent, to study whole-body gestural interaction [5]; touchscreen gestures have not been examined. This gap is important to address because touchscreen interaction studies with adults do not always generalize to children, who have different motor and cognitive capabilities [1].

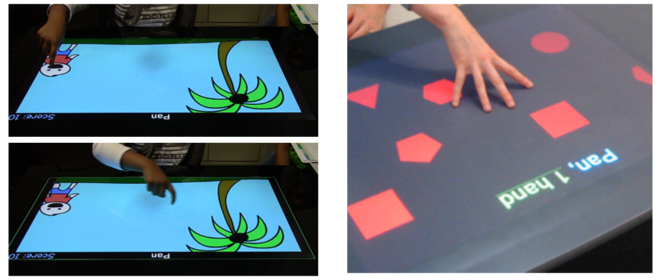

Figure 1: A sample gesture (*pan*) from our study (left) and Wobbrock et al.'s [12] study (right). Our child-centric protocol uses more concrete on-screen representations.

Even considering other approaches, only a few studies have examined children's touchscreen interaction (e.g., [1,6,8,11]). For example, Anthony et al. [1] identified differences in how children and adults articulated standard gestures on mobile devices (e.g., tapping, tracing). McKnight and Fitton [8] investigated young children's ability to understand touchscreen terminology such as "tap," "press and hold," and "slide," finding that children are more prone to accidental or unintended touches. Hinrichs and Carpendale [6] examined use of a touchscreen tabletop in a museum setting by children and adults, and found that people may use different gestures for the same action depending on age, context, and overall intention. Rick et al. [11] reported design considerations for group use of a tabletop in the classroom, such as equity of participation and use of touch space. In contrast to these studies, we consider a more open-ended gesture elicitation approach focusing on how children *design* new gesture-based interactions, allowing us to infer preferences and characterize assumptions that children bring to touchscreens.

To inform the design of touchscreen gestures for children, we first employed an iterative design process with six participants to adapt Wobbrock et al.'s [12] study protocol to better support children (Figures 1 and 2). We then employed this modified protocol to compare gestures created by 12 children and 14 adults on an interactive tabletop. The contributions of this paper are: (1) a set of changes to the user-defined gesture protocol to accommodate

IDC'14, June 17 - 20 2014, Aarhus, Denmark

http://dx.doi.org/10.1145/2593968.2610452.

the unique needs of children; (2) a characterization of differences and similarities between children's and adults' gestures; and (3) a comparison of mental models for gestural interaction between adults and children. These findings can be used to design better touchscreen gestural interactions for children and as a basis for further user-defined gesture studies with children.

2. ADAPTATION OF USER-DEFINED GESTURE PROTOCOL FOR CHILDREN

Experimental protocols designed for adults do not always apply to children [3,4]. We closely adapted Wobbrock et al.'s [12] original user-defined touchscreen gesture protocol to work with children. Our method is in contrast to Connell et al.'s [5], which included Wizard-of-Oz interaction that responded to children's whole-body gestures. Instead, we retained Wobbrock et al.'s [12] more open-ended approach of not providing feedback.

To adapt the protocol, we first conducted an iterative design session with six children (aged 7 to 11). We divided the children into two groups of three. Each group went through the original protocol on a Microsoft PixelSense tabletop. We noted when participants lost focus, were confused, or made suggestions. The final set of modifications, below, was derived (a) from this design session, (b) from our experience conducting studies with children, and, finally, (c) from two pilot sessions we conducted of the protocol (one adult and one child). The final adaptations were:

Facilitator. The original study used automated audio and video to guide participants. To better support child participants, we introduced a facilitator. The facilitator could replay video examples, repeat instructions, and answer questions.

Referents. Session length was a concern for child participants in terms of attention and focus; in addition, some referents in the original study were very similar or conceptually complex. We combined, renamed, or cut referents to create a more child-friendly list. We also added three referents to reflect actions that children might do on a touchscreen device—*app switch, make a note* (annotate), and *share* (e.g., by email)—and two collaborative referents to reflect that children often use tabletops in a collaborative context—*give* and *my space* (i.e., delineate physical space). These changes reduced the original set from 27 to 20 referents, including: one *practice* referent (*flip*), 17 *individual* user referents, and the two new *collaborative* referents. The *individual* referents were *app switch*, *back*, *bigger*, *copy*, *help*, *make a note*, *move*, *next*, *pan*, *pick many*, *pick one*, *remove*, *share*, *smaller*, *tools*, *turn*, and *undo*. Examples of renaming referents include *maximize* to *bigger*, and *previous* to *back*. A modified referent description example is changing "*Pan. Pretend you are moving the screen to reveal off screen content.*" to "*The gesture is called pan. You are looking at picture that is too large to fit on the screen. How would you move the picture to see the hidden part?*"

Effective Abstraction. The original study used abstract visuals, which children in the iterative design session found confusing. We redesigned them to be more concrete, without being so specific as to lead participants to perform a particular gesture (Figure 1).

Novelty Effect. The children were distracted by the novelty of an interactive tabletop and spent several minutes at first playing with it to understand what it could do. To mitigate this novelty effect, we added an initial 5-minute "play" period during which children used a drawing program before moving on to the "real" tasks.

Gesture Feedback. The original study did not show any visual feedback while participants gestured. Our child participants found this lack of feedback confusing and were unsure if their gestures were registered. We added visual feedback by displaying a green border around the edge of the screen when the user touched it.

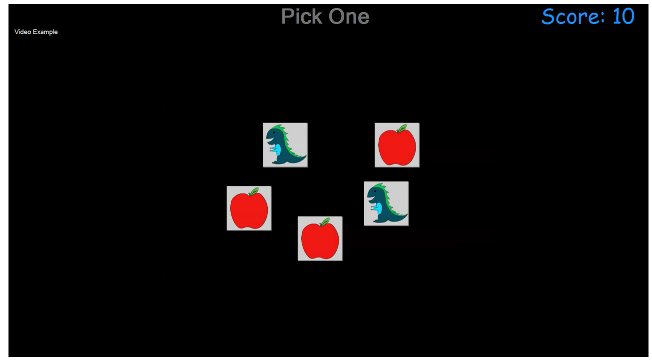

Figure 2. Screenshot of experiment software showing the *pick one* (select) referent. As a gamification element, the score in the top-right corner increased by 5 points after every referent.

Number of Hands. The original study prompted users to create both one-handed and two-handed gestures for each referent, but participants strongly preferred one-handed gestures [12]. Therefore, we chose to only prompt participants for one gesture per referent and did not specify how many hands to use.

Likert Scale. The original study used 7-point Likert scales to rate each gesture on suitability and ease of use. Such questionnaires can be difficult for children to use [10], so we employed the 5-point Smiley-o-meter scale instead [10].

Gamification. Gamification has been shown to be useful for maintaining children's engagement with empirical protocols and does not compromise data integrity [3]. As such, in our study, participants received five points per referent and a prize at the end based on total points (Figure 2).

3. METHOD

Following the iterative protocol design process, we conducted a study to compare children's versus adults' gestures on a tabletop.

3.1 Participants

Twenty-six participants (12 children and 14 adults) were recruited via email lists and word-of-mouth; none had participated in the iterative design or pilot sessions. Adults (9 females) ranged in age from 19 to 60 (M = 34, SD = 16.0), while children (4 females) were recruited to be 8 to 11 years old (M = 9.4, SD = 1.2). This age group was selected because children within this range have been shown to be clearly different than adults in terms of touchscreen use patterns [1]. Touchscreen experience was high in both groups: only two participants in each group did not have access to a touchscreen device at home, and only three adults and two children self-reported as beginner touchscreen users.

3.2 Apparatus

We adapted Wobbrock et al.'s [12] custom experiment software, written in C#, for use on a 40" diagonal Microsoft PixelSense tabletop running Windows 7 (screen resolution 1920×1080); see Figure 2. Sessions were recorded with two video cameras: one placed to the side to capture the horizontal plane of the screen and the participant from the waist up, and the second placed above the screen to capture the participant's arms and hands.

3.3 Procedure

Following the adaptations described in Section 2, the same procedure was used for adult and child participants. Participants first used a drawing application to overcome novelty and to practice the think-aloud protocol. The facilitator then opened the experiment software and walked the participant through the practice referent, after which the 17 individual referents were presented in random order. For each referent, the facilitator read a brief description, then played a video demonstrating the effect (4 to 11 seconds long). The first frame of the video was then shown again, and the facilitator prompted the participant to think aloud and envision an appropriate gesture that would cause the effect they had seen. The participant then performed their new gesture and rated it on Smiley-o-meter scales [10] for suitability ("The gesture I picked is a good match for the action.") and ease of use ("The gesture I picked is easy to do."). The two collaborative referents, which we do not report on (see Section 3.4), were presented at the end of the session and did not affect earlier tasks.

In terms of gamification, participants received a prize after completing the 17 individual referents based on points earned for completing each referent, and could trade in that prize for a different one if desired after the collaborative referents. The session ended with a survey on the participant's touchscreen experience, administered verbally for children.

3.4 Video Analysis

We analyzed 442 gestures (17 per participant); this data excludes the two collaborative gestures, which we observed to be particularly confusing for both age groups. In about 1% of cases (5 gestures), the participant became distracted by the object in the video example or went off-task despite facilitator intervention and protocol explanation. For example, P3 [child, male] focused on the texture of a tile rather than the referent's effect for *undo*. We discarded such data, leaving 437 gestures. The experiment software also logged touches on the screen, but we did not analyze this data because we were unable to accurately separate intended (hand) from unintended (arm, sleeve) touches.

We qualitatively coded the videos and think-aloud comments along five *objective* dimensions (e.g., which hands or fingers were used) and 13 *subjective* dimensions that included gesture type (e.g., tap or swipe) and mental models (e.g., rationale, and whether a menu, widget, or button was referenced). The initial code set was created based on prior work [12] and our analysis of two randomly selected videos (one child and one adult). We then used an iterative process to refine the subjective codes: two independent coders analyzed two more videos and met to discuss and refine the codes before randomly dividing and coding all remaining videos. To assess inter-rater reliability, we randomly selected two of these videos (one child and one adult) to be coded by both researchers with the final code set and calculated Cohen's kappa for each subjective coding dimension. We do not report on the dimensions for which reliability was low (kappa < 0.50); for the remaining eight dimensions, kappa ranged from 0.57 to 1.0 (M = 0.78, SD = 0.18).

4. RESULTS

Adults and children exhibited similar gesture creation patterns. For frequencies across all gestures, we report percentages out of 236 total gestures for adults and 201 total gestures for children.

4.1 Gesture Types

For the majority of gestures, both adults and children employed standard touchscreen gestures (tap, drag, swipe, pinch, rotate), rather than creating entirely new gestures. These standard gestures were used in 96% of all gestures, and the pattern of gesture types employed was similar for each participant group (Figure 3). Besides these standard gestures, participants also created symbolic gestures, such as drawing an undo arrow, writing an "X" for remove, and drawing a question mark for help. Ten adults performed a symbolic gesture compared to only four children. This result suggests that gestures for children should include direct physical manipulation rather than being symbolic.

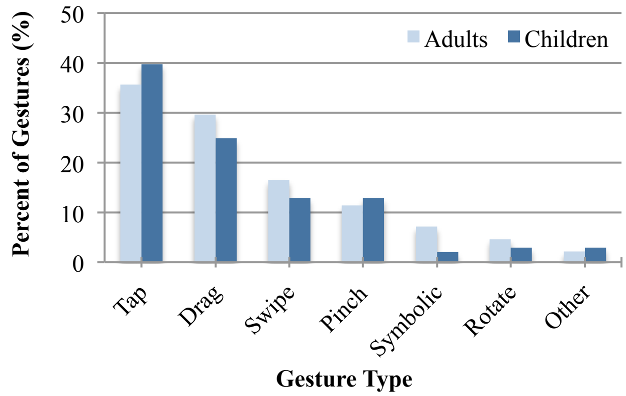

Figure 3. Types of gestures created by adults and children, showing similar patterns across both groups. Complex gestures could be coded as exhibiting more than one type (*N* = 236 adult gestures and 201 child gestures).

4.2 Gesture Mechanics

Each gesture was coded for the number of hands, number of fingers, and individual fingers used. Almost all gestures were performed with one hand: 93% of adult gestures and 92% of child gestures. For these one-handed gestures, children tended to use only one finger more often than adults: 75% of child gestures and 62% of adult gestures (Table 1). For both groups, one-handed gestures with two fingers were much less frequent, accounting for only 15% of adult gestures and 7% of child gestures. Of the adults, 86% (12/14) used two fingers for at least one one-handed gesture, while 75% (9/12) of children did so.

There were differences in individual finger usage across adults and children as well. Adults used their middle finger in 28% of gestures, compared to 18% of child gestures. Though the percentage of gestures using a particular finger was similar for both adult and child gestures, more individual adult participants used the middle, ring, and pinky fingers than child participants did. More adult participants used a wider range of fingers than child participants did. Overall, users performed one-handed gestures using the index or middle fingers the most.

4.3 Mental Models

We inferred mental models of touchscreen interaction from think-

Number of Fingers			Individual Finger Usage		
	Adults	Children		Adults	Children
1	62%	75%	Index	94%	95%
2	15%	7%	Middle	29%	18%
3	6%	2%	Ring	16%	12%
4	3%	2%	Pinky	11%	8%
5	6%	6%	Thumb	17%	15%

Table 1. Percent of all one-handed adult and child gestures in terms of number of fingers and individual fingers used. Adults were more likely than children to use more than one finger. (*N* = 236 adult and 201 child gestures).

aloud comments. As mentioned above, the majority of gestures created were based on standard touchscreen interactions. This reliance, however, was not always evident in the think-aloud comments, perhaps because the mapping from participants' past touchscreen experience to the study tasks was immediate enough not to compel elaboration. Participants verbally referenced a touchscreen device for only 5% of adult gestures and 7% of child gestures. As P20 [adult, male] said, "*It's all kind of going back to what touchscreen things [are] already out there. Because they've [touchscreens] been out for enough now, it's kind of just become the prevalent idea that everything has the same sort of function*."

We also examined references to WIMP (windows, icons, menus, pointers) interfaces. Note that five referents included a WIMP element in the video example, though not in the static image over which participants gestured (*app switch*, *share*, *tools*, *make a note*, and *help*; the last three are similar to [12]). In total, for 27% of child gestures and 15% of adult gestures the participant mentioned using a WIMP element. While creating a gesture for *copy*, for example, P24 [child, male] said, "*Normally, there's, on my computer at school, when I move my mouse over to that corner, if then, then I pull it out it'll turn the whole screen blue and then when I right-click it it'll normally say 'What do you want to do?' I'll press 'copy,' 'take' then 'make two'*." The top three referents for which participants mentioned using a WIMP element were *tools* 63% (15/24 instances), *app switch* 58% (15/26 instances), and *help* 52% (13/25 instances); note that two gesture instances for *tools* and one for *help* were excluded due to user confusion.

4.4 Ease of Use and Suitability

Participants rated each gesture on a scale from 1 (low) to 5 (high) for suitability and ease of use. For suitability, the average ratings were 4.1 (*SD* = 1.0) and 4.0 (*SD* = 0.9) for children. For ease of use, the average ratings were 4.5 (*SD* = 0.8) for adults and 4.0 (*SD* = 1.0) for children. These ratings also support the main message of this study: that both age groups are similar in their gesture creation patterns and mental models of gestures.

5. DISCUSSION

We adapted Wobbrock et al.'s [12] user-defined gesture protocol for use with children, including reducing and renaming referents, introducing a facilitator to guide participants through the protocol, and incorporating gamification elements. In general, these adaptations allowed us to successfully elicit user-defined gestures from both children and adults. One modification for future studies, however, would be to use more of an interview dynamic rather than only think-aloud protocol, especially for children. For example, some children were far less likely to elaborate unprompted than adults, commenting, e.g., "I would do this [gesturing]" (P52 [child, female]), without explaining further.

Many similar findings exist between our study and the Wobbrock et al. [12] study (e.g., references to WIMP interfaces). However, participants in their study were novice touchscreen users, and, since then, touchscreens have become ubiquitous. All participants in our study had touchscreen experience, and the influence of this experience was evident in the gestures they created. This finding suggests that touchscreen interaction is now mature enough that users have internalized certain standardized gestures. Designers should capitalize on these standards wherever possible.

Participants performed standard navigational and manipulative gestures (e.g., tap, swipe) and overwhelmingly used one hand and one finger to complete gestures. These findings show that, despite differences between adults and children, simple one-handed, one-finger gestures are the most frequently created for a variety of actions. However, there are times when both groups used a one-handed gesture with two-fingers for gesture types such as rotate or pinch. Designers should work to design systems that use simple, direct manipulation gestures using one or two fingers.

Overall, in our study, both adults and children tended to create gestures based on existing touchscreen interactions and created simple gestures that were repurposed for a variety of tasks.

6. LIMITATIONS AND FUTURE WORK

This study provides a basis for further user-defined gesture studies with children, but more work remains to be done. One limitation of our study is the focus on one age group (8 to 11 year olds). While this age group is the oldest group still clearly different from adults in terms of touchscreen use [1], future work should include younger children and teenagers to uncover differences between age groups. Another limitation is that participants in both groups had high touchscreen experience. Finally, while interactive tabletops are used in classrooms and public spaces, the results from our study may not reflect interactions on smaller smartphone or tablet devices. For example, screen size may affect direct manipulation gestures, such as resizing an object. Future studies should investigate this potential influence on children's gestures.

7. REFERENCES

1. 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. ACM ITS'2012*, ACM Press (2012), 225–234.
2. 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), 87–94.
3. Brewer, R., Anthony, L., Brown, Q., Irwin, G., Nias, J., and Tate, B. Using gamification to motivate children to complete empirical studies in lab environments. *Proc. IDC'2013*, ACM Press (2013), 388–391.
4. Brown, Q., Anthony, L., Brewer, R., Irwin, G., Nias, J., and Tate, B. Challenges of Replicating Empirical Studies with Children in HCI. *Proc. RepliCHI'2013*, CEUR Workshop Proceedings (2013), 53–57.
5. Connell, S., Kuo, P.-Y., Liu, L., and Piper, A.M. A Wizard-of-Oz elicitation study examining child-defined gestures with a whole-body interface. *Proc. IDC'2013*, ACM Press (2013), 277–280.
6. Hinrichs, U. and Carpendale, S. Gestures in the wild: studying multi-touch gesture sequences on interactive tabletop exhibits. *Proc. CHI'2011*, ACM Press (2011), 3023–3032.
7. Kang, C. Survey: For young children, mobile devices such as tablets, smart phones now a mainstay. *The Washington Post*, 2013. http://www.washingtonpost.com/7e386f3c-3f1f-11e3-a624-41d661b0bb78_story.html.
8. Mcknight, L. and Fitton, D. Touch-screen Technology for Children : Giving the Right Instructions and Getting the Right Responses. *Proc. IDC'2010*, ACM Press (2010), 238–241.
9. 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.
10. Read, J.C. and MacFarlane, S. Using the fun toolkit and other survey methods to gather opinions in child computer interaction. *Proc. IDC'2006*, ACM Press (2006), 81–88.
11. 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.
12. Wobbrock, J.O., Morris, M.R., and Wilson, A.D. User-defined gestures for surface computing. *Proc. CHI'2009*, ACM Press (2009), 1083–1092.