Understanding Users' Preferences for Surface Gestures

Meredith Ringel Morris[1], Jacob O. Wobbrock[2], Andrew D. Wilson[1]

[1]Microsoft Research, [2]Information School, DUB Group, University of Washington

Abstract

We compare two gesture sets for interactive surfaces—a set of gestures created by an end-user elicitation method and a set of gestures authored by three HCI researchers. Twenty-two participants who were blind to the gestures' authorship evaluated 81 gestures presented and performed on a Microsoft Surface. Our findings indicate that participants preferred gestures authored by larger groups of people, such as those created by end-user elicitation methodologies or those proposed by more than one researcher. This preference pattern seems to arise in part because the HCI researchers proposed more physically and conceptually complex gestures than end-users. We discuss our findings in detail, including the implications for surface gesture design.

Keywords: Surface computing, interactive tabletops, gestures.

Index Terms: H.5.2 [Information Interfaces and Presentation]: User Interfaces — interaction styles, user-centered design.

1 Introduction

Surface computing technologies (i.e., direct-touch technologies like interactive walls and tabletops [2][6][13][17][19]) have become increasingly common in the past few years, mostly due to hardware breakthroughs that allow precise sensing through either touch or computer vision, and due to lowering costs of component technologies. These new technologies are generally operated via hand gestures; hence, gesture design will play an important role in determining the usability and success of surface computers. Accordingly, researchers have proposed a variety of hand-gesture sets for interactive surfaces [8][10][11][13][14][16][22][23].

Surface gestures typically have been designed by computer science, design, or usability professionals, and are often created to manage constraints such as ease of automatic recognition rather than ease of use (e.g., [12]). Recently, we reported results from a user-centered methodology for gesture design [20]. This method involved playing an audio description of a command to participants (e.g., "undo"), showing participants a simulation of the effect of that command, and then asking participants to perform a gesture that they felt would cause the effect just presented. After using this elicitation method with 20 people, the participants' proposed gestures were reconciled using a majority-vote formulation of agreement and removal of conflicts [21], resulting in a User-Defined Gesture Set [20] covering 22 common commands.

In this paper, we build upon our prior work by comparing the User-Defined Gesture Set to gestures produced by three HCI researchers, the authors of this and our prior paper. We describe a study where 22 participants evaluated user-authored and researcher-authored gestures. Our findings reveal that even though participants were unaware of the authorship of each gesture, they preferred user-defined gestures over the researcher-made gesture sets. Participants preferred physically and conceptually simple gestures, while HCI researchers tended to create more complex gestures, such as those with more moving parts, precise timing, or spatial dependencies. Our results indicate the importance of incorporating consensus, by end-users or groups of designers, in the creation of surface gestures, and offer evidence that HCI researchers may not always create optimal gesture designs despite their expertise.

2 Related Work

Surface computing technologies have become a focus of research and commercial interest in recent years thanks to advances in hardware that enable accurate sensing of touch input. Systems like DiamondTouch [2] and SmartSkin [13] use capacitive touch sensing, while systems like FTIR [6] or PlayAnywhere [19] rely on computer vision techniques. Part of the appeal of these surface computing systems is their ability to support direct-touch and gesture-based interactions.

A variety of hand gestures for interactive surfaces have been proposed in the research literature. For example, Wu and Balakrishnan [22] described a set of multi-finger and whole-hand gestures for manipulating furniture layouts on a DiamondTouch table. Wu et al. [23] also described a set of gestures for manipulating and editing photos on tabletop displays. Rekimoto [13] described a set of gestures for actions such as panning, scaling, and rotation that could be used with his SmartSkin system. Ringel et al. [14] proposed a set of hand gestures that could be used to invoke mouse actions and editing actions on a camera-augmented SMARTBoard wall.

Some gesture systems operate on a horizontal surface in order to control objects on an associated display. Malik et al. [8] described multi-finger gestures for use on a horizontal surface that could be used to control objects on a nearby vertical display. Moscovitch and Hughes [11] proposed multi-finger gestures for controlling objects on a computer desktop. Wigdor et al. [18] demonstrated gesturing on the underside of a table to control content appearing on the table's topside. Surface gesture systems that combine multiple sources of input are also a topic of study. Morris et al. [10] introduced "cooperative gestures", wherein the gestures of several DiamondTouch users acting in synchrony are jointly interpreted. Tse et al. [16] combined gesture input with speech in order to control applications on tabletop displays.

The gestures in the aforementioned surface systems were all designed by the system creators, usually professional HCI researchers or developers who are advanced users of technology. This differs starkly with the direct incorporation of end-user input to the design process known as participatory design [5], which is an influential method in the field of HCI. Some gesture systems are influenced by observations of user behavior "in the wild," such as the TNT gesture for combining rotation and translation on

[1]Redmond, WA, USA, [2]Seattle, WA, USA
{merrie, awilson}@microsoft.com, wobbrock@u.washington.edu

tabletop displays [7], which was inspired by observing the manner in which paper is passed among people on traditional tables. The Charade system's [1] gesture design was also influenced by observing the types of hand movements people made naturally when giving presentations. Epps et al. [3] took a more user-centered approach, asking people to demonstrate gestures in response to specific prompts and observing common trends, although they did not generate a gesture set based on their observations.

Finally, as noted above, in our prior work [20] we employed a user-elicitation methodology based on command effect prompts, and an agreement score (as defined in Wobbrock et al. [21]) to combine multiple users' gestures into a coherent, conflict-free gesture set giving maximum coverage of the set of user-proposed gestures. Micire et al. [9] used our methodology to derive a set of surface gestures specific to the domain of robot control, and Frisch et al. [4] used the method to derive surface gestures for diagram editing. We use this proposed User-Defined Gesture Set in our study to explore whether user-authored or researcher-authored gestures are more preferred by end-users of surface technology.

3 METHODOLOGY

To better understand the differences between user-defined and researcher-defined surface gestures, we conducted a lab study in which 22 participants provided feedback on 81 gestures, which were previously created by a mixture of end-users [20] and HCI researchers. This section provides more detail on the creation of the gesture sets and the methodology for preference elicitation.

3.1 Gesture Set Creation

We studied the set of 22 commands covered by the User-Defined Gesture Set [20]. This set of commands, listed in Table 1, covers a broad spectrum of tasks common to many applications, including tasks familiar from the WIMP paradigm (e.g., summoning a menu), the Web paradigm (e.g., next/previous), and direct-manipulation tasks often associated with touch-based surfaces and interactive media (e.g., rotation, scaling, zoom, panning). The results reported in [20] describe a set of 48 user-defined gestures covering these 22 commands. Note that this means some of the commands can be issued with multiple gestures.

Three HCI researchers (the authors of this paper) individually designed a one-handed and a two-handed gesture for each of the 22 commands. Each of the three researchers is an expert in the field of HCI and in the field of gesture interaction specifically. Each of the researchers has formal training in both computer science and human-computer interaction, and each has designed, implemented, and evaluated several gesture interactive systems, including gesture systems for surface computers. These three researchers did not consult with each other and did not have any exposure to the User-Defined Gesture Set before defining their own gestures. That is, the gesture set was designed *before* the previous study [20]. Each researcher's goal was to propose an intuitive set of gestures for the given list of commands.

A total of 63 distinct gestures were proposed by the three researchers. There was some overlap among the researchers' proposed gestures: 37 were proposed by only one of the three researchers, 12 were proposed by two, and 14 were proposed by all three. Some of the researcher-created gestures coincidentally overlapped with gestures from the User-Defined Gesture Set, but as noted, this set was not yet in existence. Thus, put together, we had a set of 81 gestures covering 22 commands: 30 that were proposed by both the researchers and the user-elicitation method,

Figure 1. A participant imitates a gesture for "zoom out" after viewing a video demonstration of that gesture.

which we will refer to as the "overlapping gestures," 18 proposed only by the user-elicitation technique of prior work [20], and 33 proposed only by the researchers. The Appendix depicts all 81 gestures.

3.2 Preference Elicitation

After gathering the 81 user- and researcher-authored gestures, we conducted a study to learn about users' gesture preferences. We recruited 22 participants (12 male). Participants' ages ranged from 18 to 49 years (mean = 32), and participants had a variety of occupations unrelated to computer science, design, or usability. Example occupations included pre-school teacher, lifeguard, army private, nurse, office manager, environmental engineer, minister, and homemaker. All participants were right-handed, and had no prior experience using interactive surfaces or other touch-screen technologies, including the Apple iPhone. Note that although we employ the User-Defined Gesture Set [20] in this study, none of our participants were involved in the creation of that set.

Each participant sat in front of a Microsoft Surface interactive tabletop display, with a numeric keypad located on the edge of the table (Figure 1). Participants first did a tutorial for the command "clear screen," which was not part of the command set studied. We created two gestures for the "clear screen" command for the purposes of the tutorial. The procedure for the tutorial, which is the same as the procedure used for the remainder of the study, is described in the following paragraphs.

First, the Microsoft Surface display showed a screen that portrayed the name of the current command, in this case, "clear screen." A pre-recorded audio prompt stated the name of the command and provided a brief audio definition (e.g., "clear screen: remove all on-screen objects").

Then, for each proposed gesture for the current command, in this case, the two "clear screen" tutorial gestures, the surface displayed a video showing an actor demonstrating the current gesture. The gestures were demonstrated in a "shapes world," as was done previously [20], in order to avoid any similarity with Windows or pre-existing software applications. Although the surface did not yet recognize the 81 proposed gestures, Wizard-of-Oz techniques were used in the videos to make the surface appear to respond to the actor's gestures as if the gestures were recognized. Participants could replay the video as many times as desired. Next, participants were shown the same prompt used in the video demonstrating the gesture (e.g., a field of shapes to clear), and were asked to imitate the gesture they had just seen (Figure 1) so that they would be better able to judge which motions they preferred. The system did not respond to

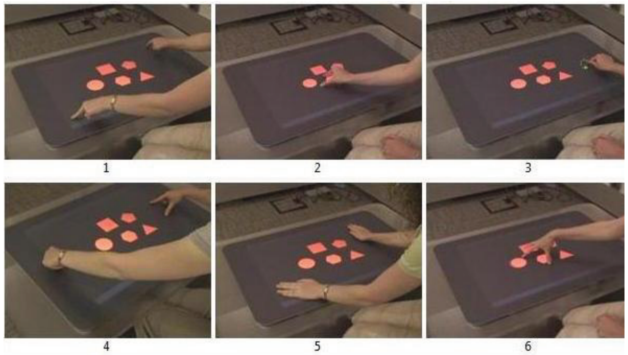


Figure 2. After viewing, imitating, and rating all of the gestures for a particular command, participants were presented a screen of thumbnail images depicting each of the proposed gestures, and were asked to select which one was the best for that command. This figure depicts the six alternatives shown for "zoom out." Any gesture could be replayed at this stage.

participants' gestures during this imitation phase. All subjects reported at the end of the study that they found this "imitation" step helpful in their decision-making process.

After imitating the gesture, the surface display presented two 7-point Likert-scale questions that the participants answered using the numeric keypad. The first question asked whether the gesture they had just imitated was a good match for the current command (i.e., would that gesture be a good way to execute that command). The second question asked whether they felt the gesture they had just tried was easy to perform (i.e., rate the difficulty of carrying out the gesture's physical action). After completing both Likert questions, the participant repeated the video-imitation-question process for all remaining proposed gestures for that command. The order of presentation of gestures for a given command was randomized for each participant. Each command had between 1 and 6 alternative gestures (Table 1).

After completing the video-imitation-question procedure for each of the current command's proposed gestures, participants saw a screen with a large thumbnail image depicting each of the proposed gestures for that command (Figure 2). In the event that they could not recall what gesture a particular thumbnail represented, they could replay the associated video demonstration. Participants used the numeric keypad to indicate which of the gestures they felt was best for the current command, i.e., which of the gestures they would want to signify that command in an actual system. Participants were told that they could consider each command in isolation, i.e., they did not need to worry about whether a gesture they chose as best for one command was similar to one they already chose for another command. This was done to lessen the cognitive and memory demands on participants.

This entire procedure (command definition; video-imitation-question for each proposed gesture; choice of preferred gesture) was repeated for all 22 commands in Table 1. For each participant, the 22 commands were randomly ordered. Participants were blind to the authorship of the gestures, and were not even aware that different gestures may have been authored by different sources. The experiment took between 60-90 minutes per user.

At the conclusion of the study, the experimenter asked each participant for any comments or feedback regarding what they had just experienced. The experimenter also took notes throughout the study on comments made by participants; all sessions were also video-recorded. Responses to Likert-scale and gesture preference questions were logged by our software.

command	**number of gestures**	**% choosing "winner"**
accept	1	100%
minimize	3	90.9%
previous	2	90.9%
select single	3	90.9%
help	3	86.4%
next	2	86.4%
open	5	86.4%
move	3	81.8%
cut	2	77.3%
rotate	4	68.2%
shrink	5	68.2%
delete	5	63.6%
pan	2	63.6%
undo	4	63.6%
select group	3	59.1%
menu	5	54.5%
paste	4	54.5%
reject	5	54.5%
enlarge	5	45.5%
zoom in	5	45.0%
duplicate	4	36.4%
zoom out	6	22.7%

Table 1. For each command studied, the number of gesture alternatives and the percent of participants who chose the gesture receiving the most "best" votes.

4 RESULTS

Overall, participants exhibited a surprising degree of consensus in their choice of preferred gestures. We found that gestures rated more highly by participants were also proposed by a greater number of gesture authors—that is, researchers and users from our previous study [20]. The gestures proposed by both users and researchers were preferred to those proposed by users only, which in turn, were preferred to those proposed by researchers only. We also found that participants greatly preferred simple, easy-to-perform gestures over more complex ones (e.g., gestures using a single finger were preferred to those using an entire hand, and gestures using one hand were preferred to bimanual gestures). The remainder of this section provides more detail on these findings.

Note that due to the subjective (and potentially non-equidistant) interpretations participants may attribute to Likert scales, we use non-parametric statistical tests when analyzing Likert scale responses; however, we include both the median and mean scores in the accompanying tables to provide the reader with a detailed overview of the data.

4.1 Preferred Gestures

The Appendix indicates the gesture for each command that received the highest number of "votes" (i.e., number of participants who chose that gesture as the best gesture for that command).

Table 1 shows the percent of participants who chose the most-preferred gesture for each command as their favorite. If participants did not exhibit commonalities in their preferences for gesture/command pairings, we would expect the distribution of votes for the winner to be distributed similar to chance, i.e., the proportion of participants voting as favorite a particular gesture for a given command would be 1/n, where n represents the number of gestures proposed for a given command. However, we found instead that there was substantial similarity in participants' choice of preferred gesture for each command. Excluding the "accept" command, since it had only one proposed gesture, the percent of participants agreeing on the most-preferred gesture for each command (mean = 66.2%, std dev = 19.4) was significantly higher than chance (mean = 29.7%, std dev = 11.5), as confirmed by a paired-samples t-test ($t(20) = 10.54$, $p < .001$).

4.2 Influence of Authorship

Gestures authored by more people were rated on Likert scales more highly by participants than those authored by fewer people.

A Kruskal-Wallis test comparing the "good match" scores for all gestures grouped by author class (user-defined, researcher-defined, or overlapping) showed significant differences: $\chi^2(2, N = 1780) = 106.10$, $p < .001$ (Table 2). Follow up pairwise Mann-Whitney U tests found that all pairwise differences were significant, with gestures authored by users-only being considered a better match for their respective commands than gestures authored by researchers-only ($z = -4.91$, $p < .001$) and gestures proposed by both groups having the highest ratings of all ($z = -4.09$, $p < .001$).

Likert-scale ratings for how easy each gesture was to perform showed a similar trend. A Kruskal-Wallis test comparing the "ease of performance" scores for all gestures grouped by author class showed significant differences: $\chi^2(2, N = 1780) = 47.82$, $p < .001$ (Table 3). Follow up pairwise Mann-Whitney U tests found that all pairwise differences were significant, with gestures authored by users-only having higher ease ratings than gestures authored by researchers only ($z = -4.01$, $p < .001$), and gestures authored by both groups having the highest ease ratings of all ($z = -1.96$, $p = .05$).

authorship	median	mean	std dev
overlapping	6	5.55	1.20
users	5	5.22	1.31
researchers	5	4.76	1.49

Table 2. Likert ratings for how good a match each gesture was for its respective command, according to gesture authorship.

Among the three researchers, who individually designed their gesture sets without consulting each other, there was some overlap in proposed gestures. Thus, some of the gestures in the "researcher-only" authorship category were proposed by all three researchers, some by only two of the researchers, and some by only a single researcher. When considering only the researcher-only gestures, the trend still holds that gestures proposed by more people were more highly rated. A Kruskal-Wallis comparing the median Likert scores for "good match" for the researcher-only gestures, grouped by number of researchers that proposed that gesture (1, 2, or 3), showed significant differences $\chi^2(2, N = 550) = 11.31$, $p = .004$ (Table 4). Follow-up pairwise Mann-Whitney U tests found that gestures proposed by either two researchers ($z = -2.79$, $p = .005$) or three researchers ($z = -2.25$, $p = .025$) had significantly higher ratings than those proposed by only one researcher.

authorship	median	mean	std dev
overlapping	6	5.84	1.24
users	6	5.71	1.23
researchers	6	5.32	1.49

Table 3. Likert ratings for how easy each gesture was to perform, according to gesture authorship.

authorship	median	mean	std dev
3 researchers	5	5.05	1.40
2 researchers	5	5.15	1.37
1 researcher	5	4.65	2.29

Table 4. Likert ratings for how good a match each researcher-only gesture was for its respective command, according to the number of researchers proposing that gesture.

4.3 Influence of Simplicity

In general, participants preferred simpler gestures to more complex ones. By simple, we mean gestures that were physically easier to perform and/or demanded less cognitive effort. For instance, one-handed gestures were preferred to two-handed, and gestures using only a single finger were preferred to those using multiple fingers or an entire hand. Conceptually simpler gestures (i.e., based on physical analogies rather than abstract mappings) were also preferred.

The preference for simple gestures is reflected in the correlation between participants' Likert-scale ratings of how easy a gesture was to perform and their ratings of whether that gesture was a good match for its command—there was a positive correlation between these two ratings ($r(1778) = 0.59$, $p < .001$). Additionally, the gestures voted as "best" for each command had significantly higher "ease of performance" ratings (median = 6, mean = 6.40, std dev = 0.71) than those not voted best (median = 6, mean = 5.31, std dev = 1.43), as confirmed by a Mann-Whitney U test ($z = -15.65$, $p < .001$).

One-handed gestures were rated more highly than two-handed gestures, in terms of both the goodness of match between gesture and command and in terms of the ease of performing the motions (Table 5). Mann-Whitney U tests showed that one-handed gestures received significantly better "good match" scores than two-handed gestures ($z = -5.91$, $p < .001$) and that one-handed gestures received significantly better "ease of performance" scores than two-handed gestures ($z = -8.04$, $p < .001$).

Gestures using only a single-finger were rated more highly than those using more than one finger (Table 6). Mann-Whitney U tests showed that single-fingered gestures received significantly better "good match" scores ($z = -4.88$, $p < .001$) and that single-fingered gestures received significantly better "ease of performance" scores ($z = -8.55$, $p < .001$).

rating type	hands	median	mean	std dev
good match	1	6	5.29	1.35
good match	2	5	4.88	1.43
performance ease	1	6	5.79	1.27
performance ease	2	6	5.22	1.47

Table 5. Participants preferred one-handed gestures to two-handed, rating one-handed as significantly better in terms of match between gesture and command and in terms of ease of performance.

Our prior work [20] proposed a taxonomy of surface gestures, classifying a gesture's "nature" as either physical, symbolic, metaphorical, or abstract. We classified the 81 gestures from our current study according to this taxonomy in order to see whether a gesture's nature impacted it's preference by end-users. We found that gestures with conceptually simpler natures (those based on analogy to the physical world, and those using common symbols)

rating type	fingers	median	mean	std dev
good match	1	6	5.33	1.40
good match	>1	5	5.04	1.38
performance ease	1	6	5.91	1.23
performance ease	>1	6	5.39	1.41

Table 6. Participants preferred single-finger gestures to multi-finger or whole-hand gestures, rating single-finger gestures better in terms of match between gesture and command and in terms of ease of performance.

were preferred by our participants to those with more conceptually complex natures (those based on metaphorical or abstract mappings). We performed Kruskal-Wallis tests to compare "good match" and "ease of performance" ratings for all gestures, grouped by "nature" into the four categories above. The tests showed no significant effect of nature on "ease of performance" ratings. However, there was a significant effect of nature on "good match" scores: $\chi^2(3, N = 1780) = 20.14, p < .001$ (Table 7). Follow-up pairwise Mann-Whitney U tests showed that there were neither any significant differences in "good match" ratings between physical and symbolic gestures, nor between abstract and metaphorical gestures. However, all other pairwise differences were significant, with physical gestures having higher "good match" scores than abstract ($z = -3.90, p < .001$) and metaphorical ($z = -2.46, p = .014$), and symbolic gestures also having higher scores than abstract ($z = -3.11, p = .002$) and metaphorical ($z = -2.21, p = .027$).

nature	median	mean	std dev
symbolic	6	5.29	1.41
physical	6	5.28	1.30
metaphorical	5	5.00	1.48
abstract	5	4.94	1.47

Table 7. Participants rated conceptually simpler gestures (those with symbolic or physical natures) as being significantly better matches for their respective commands than those based on more complex (metaphorical or abstract) mappings.

Participants' preference for simple gestures was also reflected by their comments during and after the study. Although participants were not explicitly asked questions regarding simplicity, their comments revealed five main reasons why they preferred simpler gestures. These are explained below.

Desire to Use One Hand for Other Tasks: Six participants mentioned that they preferred one-handed gestures because they imagined that they may not always have two hands available. For example, one participant mentioned that he prefers to lean on one hand when seated around a table, while five participants mentioned that they might want to hold other items (such as beverages) with one hand while touching the surface with the other.

Familiarity with Legacy Applications and Mice: Six participants attributed their fondness of one-fingered gestures to their familiarity with desktop PCs and mice. Gestures drawing on GUI metaphors, such as double-tapping to invoke "open," evoked comments such as, "It's just like on the computer, so that makes sense, it's like what I'm already used to," and "[that gesture is] one of the best… it reminds me of double-clicking with the mouse at home." One participant reflected on his preference for mouse-like gestures by noting, "I think I'm kind of stuck in legacy."

Precision: Three participants mentioned that a subjective sense of precision played a role in their preference for single-fingered gestures. For example, one noted that when she used her entire hand for a gesture, she felt more likely to accidentally touch on-screen objects that were not the target of her action. Another echoed this sentiment; he described whole-hand gestures as "overwhelming," but called single-finger gestures "accurate."

Efficiency of Frequent Actions: Four participants felt that gestures using multiple fingers and especially multiple hands would become tiring, and time-consuming, if they were to use them with any frequency. For example, one complained that two-handed gestures took too much "coordination" and "energy" to perform. Describing a gesture for "undo" that involved rubbing the hand back and forth, another participant expressed her desire for a simpler motion by pointing out that the need to rub back and forth several times "takes too long," and described her displeasure with a gesture for invoking a menu by drawing the letter "M", noting that "it should be much simpler for things that I use all the time." Another user noted that a researcher-authored "select single" gesture (scooping up an item with one's hand) had more initial appeal than an alternative where tapping with a single finger selects an object. She observed that it would be more interesting to watch someone else use the scooping gesture (making an analogy to the movie *Minority Report*, which features gesture technology), but concluded that for her own everyday use she would rather use the simpler motion.

Physical Discomfort: Two participants also mentioned that multi-finger and multi-hand gestures were uncomfortable to perform. Both mentioned that they felt contacting the surface with multiple fingers (compared to with a single finger) made the table dirty and made their hands sweaty. One also noted that gestures requiring two hands to perform made it "seem like you would get shoulder pain after a while." These comments are consistent with participants' Likert-scale ratings indicating that one-handed gestures were easier to perform than two-handed, and that single-finger gestures were easier to perform than multi-finger or whole-hand gestures.

5 DISCUSSION

Our results showed that our study participants generally exhibited similar gesture preference patterns, and that these preferences tended toward physically and conceptually simple gestures, and towards gestures that had been designed by larger sets of people, even though our participants did not know how many gesture authors proposed the gesture, nor did they know the expertise of the gesture authors. In this section, we discuss differences in the types of gestures proposed by the user-defined methodology [20] and those proposed by the HCI researchers that may have resulted in the different preferences for these two authorship groups. We then discuss the broader implications of our findings for the design of surface gesture systems, and for the process of design itself.

5.1 Differences in User and Researcher Gestures

Our study found that participants gave higher ratings to gestures from the user-defined set than to those authored by the HCI researchers. The researcher-authored gestures tended to be more physically and conceptually complex than the user-authored gestures, which contrasted with the desire for simplicity espoused by our participants. For example, the user-authored gestures were more likely to use only a single finger (65.6% of the one-handed gestures) than the researcher-authored gestures (58.1% of the one-handed gestures). The user-authored gestures were also more likely to be conceptually simpler (i.e., symbolic or physical, at 66.7%) than the researcher-authored gestures (at 58.1%).

In general, it seemed that the researcher-authored gestures were often more creative and visually appealing. For example, nearly all participants laughed or smiled when they saw the demonstration video of a researcher-authored "help" gesture in which the actor struck the table in frustration in order to summon a help system; however, only 2 of the 22 participants selected this as the preferred gesture for the "help" command. Although people were entertained by these "clever" gestures, they ultimately preferred simplicity. P11 captured this sentiment when she noted that the gestures using two hands or the whole hand were more "exciting," and that she would prefer to watch someone else perform those gestures, but for doing them herself she liked using just one finger. This finding may help explain Ryall et al.'s [15] informal observations of DiamondTouch users, noting that people used only a single finger when interacting with the tabletop, even though multi-finger interactions were available.

Trying the gestures themselves, rather than merely watching the video demonstrations, seemed to influence participants to revise their preferences if they found an action was effortful to perform. P7 articulated this best when she commented that imitating the gestures helped her decide which ones she didn't like. Having participants physically mimic each gesture was therefore an important part of our study methodology, emphasizing the kinesthetic influences gestures can exert on users' preferences.

5.2 Implications for Gesture Design

Gestures authored by larger groups of people received better ratings in our study. Gestures proposed by both the user-defined methodology and by the researchers were rated most highly, followed by those proposed by the user-defined methodology only, followed by those proposed by researchers only. Even within the researcher-only gestures, gestures proposed by multiple researchers were preferred to those that were proposed by only a single researcher. This seems to make a strong case for employing participatory design when creating gesture sets.

Although it may seem obvious that more people will prefer a gesture independently suggested by more people, this is in contrast to typical practices for designing surface gesture systems. Indeed, usability and design professionals go through extensive training to become experts, and such experts are usually the creators of interaction techniques, which may then be refined via user testing and iterative design. Our finding suggests that participatory design methodologies [5], even those where participants are not so much actively "designing" as they are "proposing," should be applied to gesture design, such as the user-centered gesture elicitation methodology [20]. The use of end-user elicitation methodologies for gesture design could be a time- and cost-efficient method of reducing the number of iterative design cycles that might otherwise be needed to refine a gesture set, especially if the software for capturing people's proposals can be uploaded to and hosted on the Web.

In the event that an end-user based design is not possible, design professionals should strive to work in teams when developing gesture sets, since multi-author gestures were preferred to single-author gestures. Additionally, HCI professionals can improve their gesture designs by being aware of the tendency to create gestures that are more physically and conceptually complex than end-users seem to prefer.

5.3 Limitations

This study represents a first step toward verifying the utility of end-user gesture design methodologies; understanding the value of these methods is important, as they have recently been used by several research groups in order to create surface gesture sets for various application areas (e.g., [3][4][9][20]). The study described in this paper provides valuable insights into participants' initial reactions to end-user and researcher-authored gesture sets, as well as an understanding of the differences between the gestures proposed by these two groups. However, additional studies investigating the value of end-user gesture design are still warranted to address issues that are beyond the scope of this study.

For example, our study measured users' preferences based only on brief interactions (observing a video of the gesture and imitating the gesture). Understanding these initial impressions is important, as many surface computing systems are designed for walk-up-and-use scenarios, such as lobby kiosks (e.g., [15]), for which all users are novices and their interactions with the system are brief. However, there is also value in conducting further work to understand how long-term use affects preferences; for example, the ease of learning a gesture set may be an important factor in shaping a user's preferences over time. Similarly, our study measured users' preferences in the absence of a specific application context. Such an approach is appropriate for understanding general differences between user- and researcher-authored gestures, and for evaluating cross-application gesture sets (as the User-Defined Gesture set is intended to be [20]); however, understanding how application context influences gesture preference is a valuable area for further study.

The participants in our study were very accustomed to the WIMP paradigm; although we used the "shapes world" in order to discourage this effect, the influence of WIMP's legacy was clear both in the gestures produced by the end-user methodology and in the preferences of participants in our study. However, as gesture interfaces become more common, it is possible that a post-WIMP generation of users will emerge. This new generation of users may have a different set of biases and expectations, which may change both the nature of gestures produced via end-user methods (perhaps making them more similar to the researchers' gestures), as well as change the factors influencing users' preferences.

The choice of "experts" to contrast with the end-users is also an important issue to consider. In this study, we used three computer science and HCI researchers; these researchers came from corporate and university settings, and have designed many gesture-based systems (including commercial, open-source, and research systems). Regardless of the origin of the "expert" gestures, this study provides detailed insights into the factors affecting users' gesture preferences. However, understanding how the gestures proposed by design professionals from beyond the research world compare to either the end-user designs or the researchers' designs would certainly be an important and interesting extension to this work.

6 Conclusion

In this paper, we described a study of 81 hand gestures for interacting with surface computing systems. These gestures were obtained from two distinct sources: the end-user elicitation process, described in [20], and from HCI researchers. When 22 participants evaluated these gestures, they exhibited similarity in their preference patterns, preferring gestures with more consensus in their authorship, such as user-authored gestures and, to a lesser extent, gestures proposed by multiple researchers. These preferences seemed to arise mostly to physical and conceptual simplicity—ease of performance and understanding.

Direct-manipulation interactive surfaces are becoming increasingly prevalent, and gesture design will play an important role in determining the success of these technologies. Our findings contribute concrete suggestions for improving surface gesture design, such as utilizing user-elicitation processes or large design teams, and creating simple gestures, particularly those using only a single hand, or, better yet, a single finger.

REFERENCES

[1] Baudel, T. and Beaudouin-Lafon, M. Charade: Remote Control of Objects Using Free-Hand Gestures. *Communications of the ACM*, 36(7), 28-35.

[2] Dietz, P. and Leigh, D. DiamondTouch: A Multi-User Touch Technology. *UIST 2001*, 219-226.

[3] Epps, J., Lichman, S. and Wu, M. A Study of Hand Shape Use in Tabletop Gesture Interaction. *CHI 2006 Ext. Abstracts*, 748-753.

[4] Frisch, M., Heydekorn, J., and Dachselt, R. Investigating Multi-Touch and Pen Gestures for Diagram Editing on Interactive Surfaces. *Tabletop 2009*, 167-174.

[5] Greenbaum, J. and Kyng, M. (1991) *Design at Work*. Hillsdale, NJ: Lawrence Erlbaum.

[6] Han, J. Low-Cost Multi-Touch Sensing through Frustrated Total Internal Reflection. *UIST 2005*, 115-118.

[7] Liu, J., Pinelle, D., Sallam, S., Subramanian, S. and Gutwin, C. TNT: Improved Rotation and Translation on Digital Tables. *Graphics Interface 2006*, 25-32.

[8] Malik, S., Ranjan, A. and Balakrishnan, R. Interacting with Large Displays from a Distance with Vision-Tracked Multi-Finger Gestural Input. *UIST 2005*, 43-52.

[9] Micire, M., Desai, M., Courtemanche, A., and Yanco, H.A. Analysis of Natural Gestures for Controlling Robot Teams on Multi-touch Tabletop Surfaces. *Tabletop 2009*, 45-52.

[10] Morris, M.R., Huang, A., Paepcke, A. and Winograd, T. Cooperative gestures: Multi-user Gestural Interactions for Co-located Groupware. *CHI 2006*, 1201-1210.

[11] Moscovich, T. and Hughes, J.F. Multi-Finger Cursor Techniques. *Graphics Interface 2006*, 1-7.

[12] Nielsen, M., Störring, M., Moeslund, T.B. and Granum, E. A Procedure for Developing Intuitive and Ergonomic Gesture Interfaces for HCI. *International Gesture Workshop 2003*, 409-420.

[13] Rekimoto, J. SmartSkin: An Infrastructure for Freehand Manipulation on Interactive Surfaces. *CHI 2002*, 113-120.

[14] Ringel, M., Berg, H., Jin,Y., and Winograd, T. Barehands: Implement-Free Interaction with a Wall-Mounted Display. *CHI 2001 Ext. Abstracts*, 367-368.

[15] Ryall, K., Morris, M.R., Everitt, K., Forlines, C., and Shen, C. Experiences with and Observations of Direct-Touch Tabletops. *Tabletop 2006*, 89-96.

[16] Tse, E., Shen, C., Greenberg, S. and Forlines, C. Enabling Interaction with Single User Applications through Speech and Gestures on a Multi-User Tabletop. *AVI 2006*, 336-343.

[17] Wellner, P. (1993) Interacting with Paper on the DigitalDesk. *Communications of the ACM*, 36 (7), 87-96.

[18] Wigdor, D., Leigh, D., Forlines, C., Shipman, S., Barnwell, J., Balakrishnan, R. and Shen, C. Under the Table Interaction. *UIST 2006*, 259-268.

[19] Wilson, A.D. PlayAnywhere: A Compact Interactive Tabletop Projection-Vision System. *UIST 2005*, 83-92.

[20] Wobbrock, J.O., Morris, M.R., and Wilson, A. User-Defined Gestures for Surface Computing. *CHI 2009*, 1083-1092.

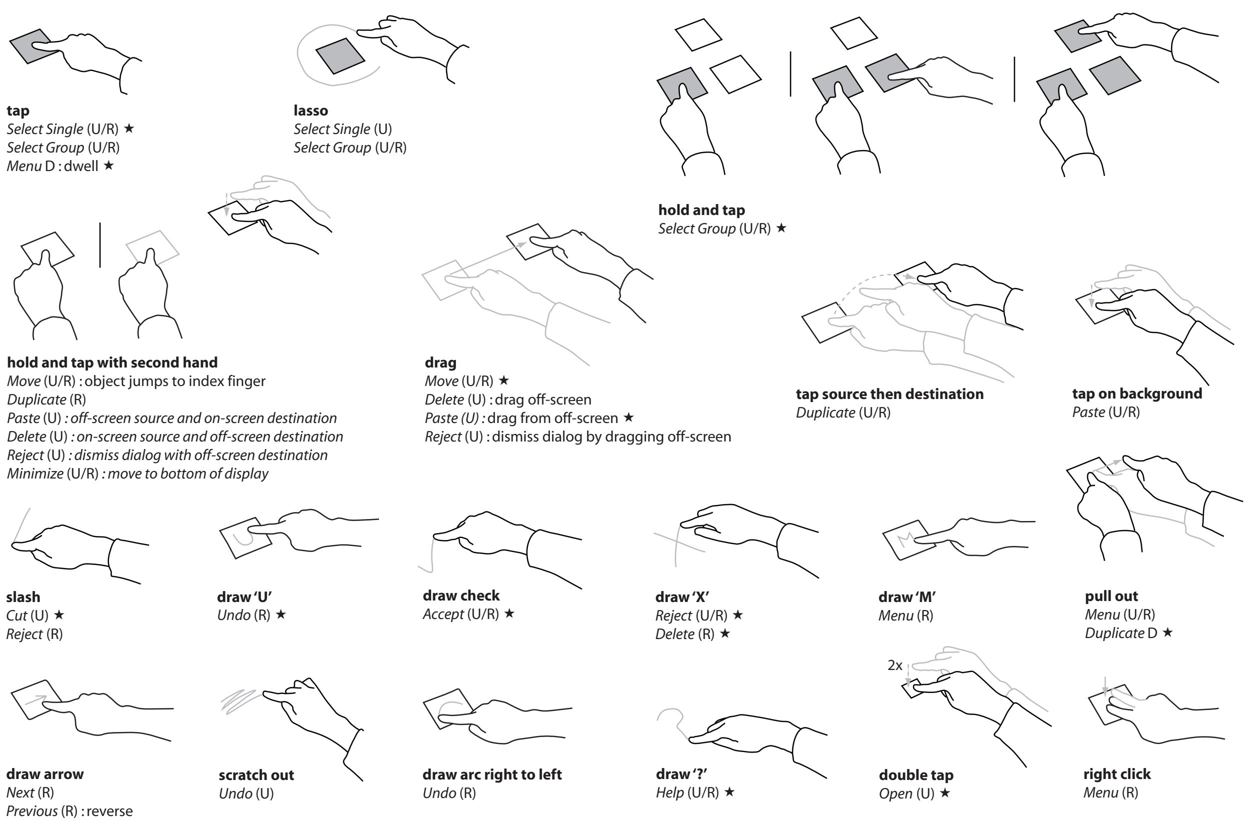
[21] Wobbrock, J.O., Aung, H.H., Rothrock, B. and Myers, B.A. Maximizing the Guessability of Symbolic Input. *CHI 2005 Ext. Abstracts,* 1869-1872.

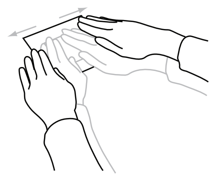
[22] Wu, M. and Balakrishnan, R. Multi-Finger and Whole Hand Gestural Interaction Techniques for Multi-User Tabletop Displays. *UIST 2003*, 193-202.

[23] Wu, M., Shen, C., Ryall, K., Forlines, C. and Balakrishnan, R. Gesture Registration, Relaxation, and Reuse for Multi-Point Direct-Touch Surfaces. *Tabletop 2006*, 185-192.

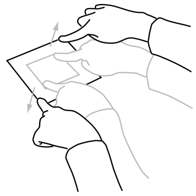
7 APPENDIX

Each diagram and accompanying description on this and the following page illustrates a gesture from our study. Underneath each diagram is a list of the commands for which that gesture was proposed, along with any descriptions of command-specific variations on the basic motion depicted. The designations “U”, “R”, and “U/R” indicate whether the user-authored gesture set, the researcher-authored sets, or both, included that gesture/command pairing. A star indicates that a particular gesture/command pairing received the most votes as the preferred gesture for its command.

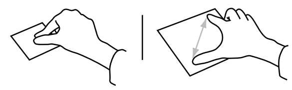




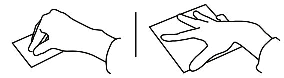
pull apart with hands
Enlarge (U/R)
Shrink (U) : reverse
Open (U/R)



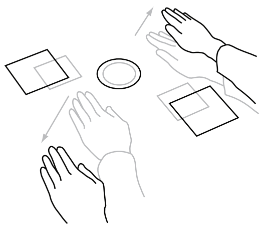
pull apart with fingers
Enlarge (U)
Shrink (U/R) : reverse
Open (U)
Zoom in (U/R) : on background
Zoom out (U) : reverse, on background



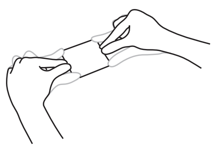
reverse pinch
Enlarge (U/R) ★
Shrink (U/R) : reverse ★
Open (U)
Zoom in (U/R) : on background ★
Zoom out (U/R) : reverse, on background ★



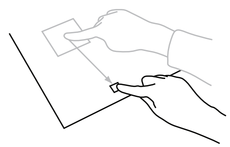
spread fingers
Enlarge (U/R)
Shrink (U/R) : reverse
Open (U/R)
Zoom in (U) : on background
Zoom out (U) : reverse, on background



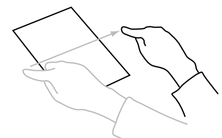
pull apart with hands
Zoom in (U/R)
Zoom out (U/R) : reverse



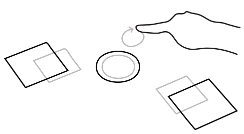
double pinch
Minimize (R)
Zoom out (R)



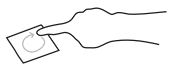
drag to bottom of display
Minimize (U/R) ★



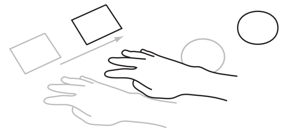
draw line left to right across object
Next (U/R) ★
Previous (U/R) : reverse ★



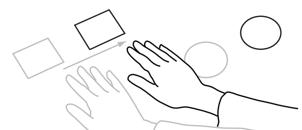
scroll ring
Zoom in (R)
Zoom out (R) : reverse



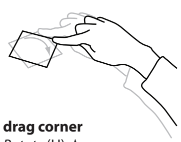
scroll ring
Enlarge (R)
Shrink (R)



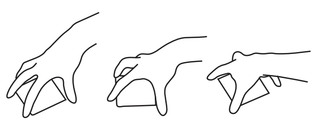
drag four fingers
Pan (R)



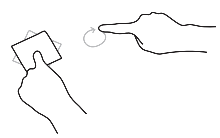
drag whole hand
Pan (U/R) ★



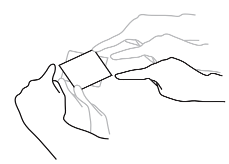
drag corner
Rotate (U) ★



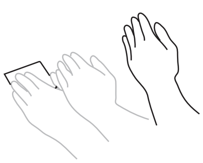
twisting grasp
Rotate (R)



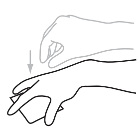
hold and scroll ring
Rotate (R)



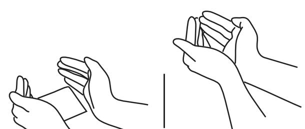
drag both corners
Rotate (R)



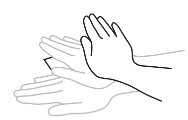
throw
Move (R)



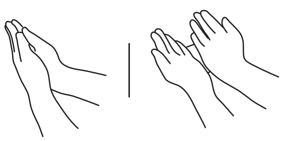
put down
Paste (R)
Cut (R) : reverse
Duplicate (R) : reverse (pick up) at source, put down at destination



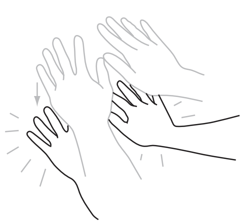
scoop up
Delete (R)
Select (R)



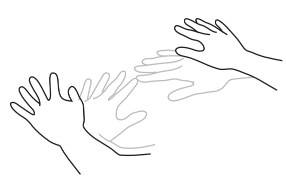
erase
Reject (R)
Undo (R)
Delete (R)



open book
Menu (R)



strike surface with two hands
Help (R)



turn hands outward
Help (R)