12-2012

Eliciting usable gestures for multi-display environments

Teddy SEYED

Chris BURNS

Mario COSTA SOUSA

Frank MAURER

Anthony TANG
Singapore Management University, tonyt@smu.edu.sg

Citation

SEYED, Teddy; BURNS, Chris; COSTA SOUSA, Mario; MAURER, Frank; and TANG, Anthony. Eliciting usable gestures for multi-display environments. (2012). *ITS '12: Proceedings of the 2012 ACM international conference on Interactive tabletops and surfaces, Cambridge, Massachusetts, USA, November 11-14*. 41-50.

Available at: https://ink.library.smu.edu.sg/sis_research/7988

Eliciting Usable Gestures for Multi-Display Environments

Teddy Seyed, Chris Burns, Mario Costa Sousa, Frank Maurer, Anthony Tang
University of Calgary, Department of Computer Science
2500 University Drive NW, Calgary, Alberta, Canada, T2N 1N4
{teddy.seyed, chris.burns, smcosta, frank.maurer, tonyt}@ucalgary.ca

ABSTRACT
Multi-display environments (MDEs) have advanced rapidly in recent years, incorporating multi-touch tabletops, tablets, wall displays and even position tracking systems. Designers have proposed a variety of interesting gestures for use in an MDE, some of which involve a user moving their hands, arms, body or even a device itself. These gestures are often used as part of interactions to move data between the various components of an MDE, which is a longstanding research problem. But designers, not users, have created most of these gestures and concerns over implementation issues such as recognition may have influenced their design. We performed a user study to elicit these gestures directly from users, but found a low level of convergence among the gestures produced. This lack of agreement is important and we discuss its possible causes and the implication it has for designers. To assist designers, we present the most prevalent gestures and some of the underlying conceptual themes behind them. We also provide analysis of how certain factors such as distance and device type impact the choice of gestures and discuss how to apply them to real-world systems.

Author Keywords
Tabletop; gestures; multi-display environments; multi-surface environments; multi-display interaction; cross-device interaction; touch; mobile devices;

ACM Classification Keywords
H.5.2. Information interfaces and presentation: User Interfaces – Interaction styles, evaluation/methodology, user-centered design

General Terms
Design; Experimentation; Human Factors

INTRODUCTION

A multi-display environment (MDE) is a system where interaction is divided over several displays, such as digital tabletops, wall displays and personal devices like tablets or mobile phones. MDEs often include heterogeneous displays to take advantage of different capabilities such as their size, position, resolution or mobility to support the task at hand.

ITS'12, November 11–14, 2012, Cambridge, Massachusetts, USA.

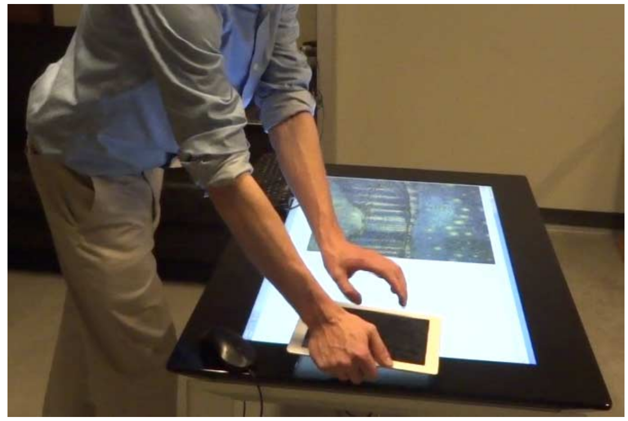


Figure 1: A user performing a gesture to move an image from a tabletop to a tablet.

Yet, a number of socio-technical questions remain for designers of MDEs: what kinds of tasks are amenable to MDEs? How should tasks be distributed among and across displays? How do we enable effective collaborative work in MDEs?

A central MDE design problem is how to allow users to move applications and content across different displays. Addressing this problem allows users to distribute tasks across displays in an ad hoc fashion as their task needs change. Designers have proposed a wide range of solutions to this problem, including considering the space as continuous to allow dragging and flicking [19,21], "docking" and physical token approaches [5,14], menu-based approaches [25], world-in-miniature approaches [2], clip-board metaphors [22], approaches that involve "publishing" personal spaces to a shared space [17] as well as approaches that consider users as conduits for information [26]. These solutions arise from attempts to resolve users' conceptual model of MDEs with the technical constraints of such systems. A more recent approach has been to consider gestural interactions that leverage the spatial layout of displays in an MDE [1].

With the increasing popularity of gestures, several researchers have explored how these can be leveraged for cross-device communication. Some approaches involve on-device gestures such as pointing and flicking [4], while others involve flicking/moving the device itself [7,8]. This gestural approach shows promise, as it seems to provide a good mapping between users' spatial model of an MDE (i.e.

that there are multiple distinct displays) to a conceptual interaction model. While recent designs have often focused on moving content between specific devices (e.g. making a tossing gesture with a mobile device towards a large display [11]) it is often unclear how well they will generalize (e.g. how does one "toss" a tabletop display?). How do we design gestures for moving content that will be meaningful across various devices in an MDE that are discoverable, and map to users' conceptual model of the space?

To investigate these issues we conducted a user study to elicit gestures directly from potential users of MDEs. This follows previous work on tabletop gestures [9,10,27] and previous work regarding MDE gestures by [15,16]. We asked 17 participants to complete various commands or tasks in an MDE environment; they were also shown the outcome of the command across displays. We then asked them to create gestures they felt were appropriate to produce that outcome. These commands made use of three different displays (tablet, tabletop, wall display), and were illustrated using a Wizard of Oz approach. Participants were given a wide scope to create imaginative and creative gestures; in return, they generated 816 total gestures for the 16 commands. We then coded these gestures to identify common conceptual themes which underlie the gestures that users produced. We found that the gestures could be collected into four general themes: contact and closeness gestures, moving physical object gestures, identification gestures and borrowed metaphors gestures. These themes can be used by designers to derive new gestures or as a starting point for future research.

This paper makes several contributions. We confirm quantitatively that gestures elicited for MDEs show a low level of agreement. We consider the potential causes for this and the implications for applying gestures. The most prevalent gestures for each command and a convergence metric are presented. Second, we outline some conceptual themes behind the gestures and provide them as tools for designers to generate new gestures with. Third, we provide an analysis of how various factors such as the devices used and the distance involved in an interaction affected the gestures elicited from our participants. Finally, we provide a discussion of how a designer can use our findings in creating gestures for an MDE.

RELATED WORK

A main cocnern in MDEs is providing users with the ability to move content and applications across and between displays. Several researchers have proposed methods and system designs addressing this problem. An early approach was to consider all displays in an environment as being connected in a continuous space [14,19]. Other researchers considered a "bridging" interaction where a digital artifact would be bound to physical objects that could then be physically moved to other displays [23]. Other approaches took into account the ad hoc nature of the displays in the environment (in particular, mobile devices, such as tablets and PDAs). Rekimoto focused on extending the clipboard metaphor with a *pick-and-drop* interaction technique, whereby users could transfer digital items between displays by first picking it with a pen gesture, and then tapping on a target display to drop the content [22]. Hinckley explored a mechanism to recognize when multiple devices had been physically bumped together [13]. This bumping, combined with a simple gesture allowed devices to be temporarily linked for content transfer.

Proxemics for Cross-Device Information Transfer

Several recent approaches have considered how proxemics, or the spatial relationships between people and devices, can be leveraged for cross-device information transfer in MDEs. Generally, proxemics considers several factors such as position, orientation, movement and identity [13]. Voida et al., proposed a set of flicking, throwing and pointing gestures to move content between surfaces in an augmented MDE [24]. Wilson and Benko, explored how people could act as conduits for information: users could physically cup a digital photo and carry it with them to another display [26]. In another gestural interaction, users could touch an image displayed on the tabletop and create a 'bridge' between the two devices to transfer that content. In *Code Space,* a series of interactions are proposed for moving data to a wall display or other tablets, mainly through in-air pointing and on-device swiping gestures [4].

Some systems have also incorporated 'simulated proxemic' interactions, where the action a user performs in the system incorporates certain proxemic dimensions, even when the system itself is not able to track all these proxemic values. Several researchers have considered how rotating or moving the device itself can form a part of the gestural dialogue for transferring data [7,8,11]. For example, Dachselt et al. propose gestures for sending and retrieving content from a large display where users point their device towards it and mimic 'throwing' content towards the display [7].

Designing Gestures

Many of these interactions, particularly those that incorporate proxemics, are gestural. While these gestures may be effective and interesting, they were all developed by system designers, who may be unduly influenced by the underlying technical constraints of the system. At issue is that the conceptual models which designers form, may not match those of the users.

Nielsen et al. [20] propose a procedure for developing user defined gestures. Tasks for a system are first defined by the system designers. Gestures are elicited for these tasks from the users and are then extracted into a gesture vocabulary for the system. Benchmarking is then done to validate the gestures chosen for the system.

Wobbrock et al. employ the gesture elicitation procedure to construct a gesture set for general actions on interactive tabletops (e.g. move, copy, paste) [27]. Here, participants are exposed to the visual result of such an action, called a referent, and then asked to produce a gesture that would result in the outcome. A relatively high level of agreement was found among the gestures produced for a given task. This convergence allowed for a gesture set to be created. The gestures in this set were then evaluated in work by Morris et al. who showed that the elicited tabletop gestures were preferred by users, compared to those created by expert designers [18].

Frisch et al. [11] employ this gesture elicitation procedure to elicit a set of pen and touch gestures for diagram editing on tabletops. These gestures were then evaluated in a follow up work by Heydekorn et al [12], where the gestures were built into a real system and then evaluated with a usability study. It was found that implementing elicited gestures in a real world system requires careful consideration of how these gestures are applied.

Several researchers have applied a similar type of approach within the context of MDEs. For instance, Kray et al. explored how mobile phones could be used in a MDE environment comprised of a tabletop and a wall display, focusing on a wide range of tasks and content types (e.g. voting for content, synchronizing devices and rewinding interactions) [15]. In this work the elicited gestures were not categorized or labeled so no agreement statistics could be produced, however, the authors observed a wide variety of novel and diverse gestures.

Similarly, Kurdyukova et al. elicited a set of gestures from users specifically for transferring content between a tablet and fixed devices such as a wall display and tabletop [16]. The work presented here also focuses on the use of tablets rather than mobile devices, as these are likely to be used in a meeting/business scenario. The study found a number of interesting gestures and suggested that distance impacts the type of gestures, but this is presented as a qualitative observation and is not analyzed statistically.

Our approach extends this previous work by focusing on data transfer in an MDE, exploring how factors such as distance-to-target-display and impacts the nature of the gestures being performed. Finally, rather than producing a patchwork of unrelated gestures, we were interested in understanding the conceptual models that users were working with, thereby providing guidance to designers.

USER STUDY

Our primary interest in designing our user study was to understand users' conceptual model of interaction in an MDE. To what extent would this model be based on the spatial characteristics of the environment, or the characteristics/affordances of particular devices?

We focused primarily on tasks related to data transfer between the various components of an MDE. We based our tasks on those described in Voida et al., which all relate to moving content between components at varying distances [24]. We were interested in looking at gestures for transfer directions (sending and retrieving), different source and destination devices (tablet, tabletop and wall display) and at different distances (near and far). To ensure that the study could be completed in a reasonable amount of time, we selectively chose combinations that we thought would occur most commonly in a meeting/war-room scenario. Table 1 lists the final set of command combinations that we used in the study.

Apparatus

The study was conducted using an Apple iPad, Microsoft Surface 2, and a SMART Board Display. A specialized Node.js web application was written to visualize the outcomes of the commands. This application allowed us to simulate the transfer of images between devices. This was by showing the image disappearing from the originating device and, after a delay, appearing on the target device.

Video recording was used to capture each performed gesture along with notes taken by the authors. To consider distance consistently between studies, specific locations were marked on the floor and participants were asked to move to these locations for certain commands.

Participants

Seventeen paid volunteers participated in the study (10 male, 7 female). Participants were recruited using posters, email and word of mouth. The participants chosen had a variety of backgrounds, such as economics, education, business and engineering. However, participants were not excluded from the study based on experience with multi-touch systems, motion tracking systems or tabletops.

Procedure

At the beginning of the session, the participants were informed of the purpose of the study and given some background information on the type of gestures that were possible within an MDE. The participants were told they could produce gestures by rotating or manipulating their device, performing actions on the screen of the device or by performing gestures using their hands, arms or even their whole body.

At the start of each command, the intended outcome was displayed for each participant. Verbal instructions confirming the purpose of the task were also provided. Participants could also ask questions if they were unclear about what was required. They were then asked to create three distinct gestures for each command as a means of stimulating creativity. Participants were also encouraged to provide gestures they felt were appropriate without considering technological feasibility such as gesture recognition or positional tracking. This was done to ensure that users were not influenced into choosing gestures because they seemed easier to track or recognize. At this stage, participants were asked to explain their gestures while they were performing them and we also asked them to choose a preferred gesture.

Commands				
No	**Task**	**Source**	**Target**	**Distance**
Single Image Transfer				
1	Send	Tablet	Tablet	Near
2	Send	Tablet	Tablet	Far
3	Retrieve	Tablet	Tablet	Near
4	Retrieve	Tablet	Tablet	Far
5	Send	Tablet	Wall Display	Near
6	Send	Tablet	Wall Display	Far
7	Retrieve	Tablet	Wall Display	Near
8	Retrieve	Tablet	Wall Display	Far
9	Send	Tablet	Tabletop	Near
10	Send	Tablet	Tabletop	Far
11	Retrieve	Tablet	Tabletop	Near
12	Retrieve	Tablet	Tabletop	Far
13	Send	Tabletop	Wall Display	Fixed
14	Send	Wall Display	Tabletop	Fixed
Multiple Image Transfer				
15	Retrieve	Tablet	Wall Display	Near
16	Retrieve	Tablet	Wall Display	Far

Table 1: Commands for user study.

Classification of Gestures

The gestures collected were classified and guided by the *Descriptive Labeling* suggested by Nielsen [20]. In this sense, we described the gestures based on their motions and actions rather than their semantic meaning. Guided by work by Buxton [6] on chunking and phrasing, we looked for distinct phrases delineated by periods of tensions and relaxation. Such a phrase was considered an *atomic* gesture while several together were considered a *compound* gesture. Overall, this classification was motivated by the desire to find useful gestures for real world systems. This meant describing the gestures in a way that emphasized the physical action involved, thus allowing designers to build gestures easily recognizable by systems utilizing motion capture or vision sensing.

RESULTS

From our 17 participants, we collected a total of 816 gestures (17 participants' × 16 commands × 3 gestures). These were classified based on the previously described strategy.

What was striking about these gestures is that our participants, completely unprompted, frequently composed gestures that we considered to be *compound* gestures. That is, these gestures were clearly comprised of one or more *atomic* gestures performed sequentially. For example, in Figure 1 the participant is moving content from the table to his tablet. Here, he performed a *Position Down* gesture (moving the iPad close to the table) followed *by Swipe Away* gesture (swiping from the centre of the iPad to the edge, in a direction away from the iPad).

To address this complexity, we first consider the *atomic* gestures. We group these atomic gestures into a set of four conceptual themes that address users' mental models of the space. We then discuss how participants grouped these atomic gestures into *compound* gestures for the commands. Finally, we use participants' *preferred* gestures to explore the impact of distance and device type on the gesture set.

Agreement Scores and Prevalent Gestures

After classifying the gestures we determined the amount of convergence among different participants for each command. An elicited gesture set should have a large amount of agreement among the gestures as it shows that common gestures are independently created by different participants for the same command. We calculated agreement scores using the favorite gestures chosen by participants and applying the same metric used by Wobbrock et al. [27]. This equation is defined in Eq 1.

$$A = \frac{\sum_{c \in C} \sum_{Pi \subseteq Pr} \left(\frac{|P_i|}{|P_r|}\right)^2}{|R|} \qquad (1)$$

The agreement scores calculated using our elicited gestures were substantially lower than those found by Wobbrock et al [27] in their study of gestures for tabletop computing. The overall agreement scores for the favorite gestures was calculated to be $A_{Favorite}$=0.16 compared to A_{1H}=0.32 and A_{2H} =0.28 for one-handed and two-handed gestures respectively. Our results demonstrate that participants showed very little convergence in the gestures they produced for the tasks. This

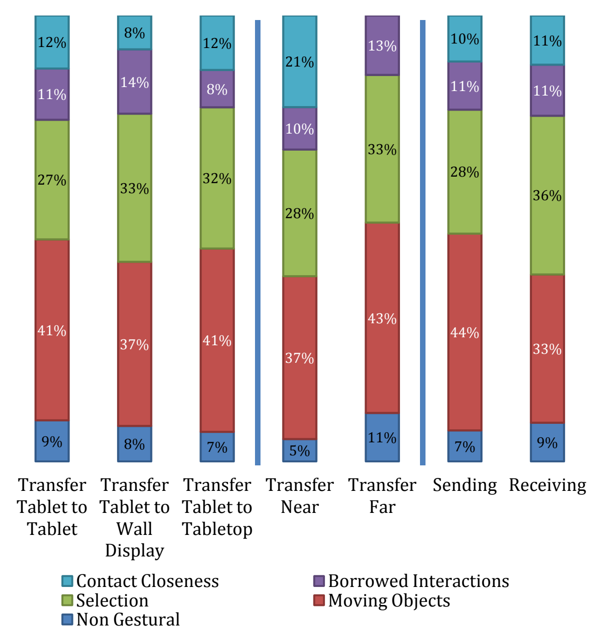


Figure 2: Percentage breakdown for metaphors for device conditions

result confirms qualitative suggestions made by earlier elicitation studies specifically for MDEs [16]. However, elicitation studies for tabletop gestures (e.g. [27]) saw a higher level of convergence and were able to build gesture sets based on that convergence.

While we do not have high enough convergence levels to propose a user-defined gesture set, we present the most prevalent gestures elicited for each command (see Figure 3). This set provides examples of gestures and the agreement scores calculated for each command gives a relative sense of the convergence for that command. To determine this set, we removed all gestures with $n < 3$ observations for each command. In some cases two gestures are presented for each command and we discuss the implications of gesture overloading and aliasing in the next section.

Conceptual Themes around Atomic Gestures

We first coded each atomic gesture for how the users positioned themselves relative to the devices, how devices were positioned/repositioned, the kinds of motions that were being made, how many fingers were used (if the device surface was contacted), etc. Through an iterative process of refinement, we arrived at four distinct conceptual themes that related the gestures with one another:

1. *Close contact:* Gestures based on close position or contact.
2. *Moving object*: Gestures based on the metaphor of moving of physical objects.
3. *Selection*: Gestures based on the act of uniquely identifying or selecting an object.
4. *Borrowed interactions*: Gestures based on existing interactions with other technologies.

These conceptual themes group the kinds of atomic gestures that we observed. Figure 2 illustrates how these different themes were used across commands. Table 2 describes the break down for the popular gestures by theme, and we illustrate some of these gestures in Figure 3.

Close Contact Theme

Users created many kinds of gestures that incorporated physical contact or physical closeness. Some users felt that positioning two devices close together or even touching together should cause data to be transferred. Atomic gestures specifically involving closeness had many variations. The *Position Above* gesture occurred when users placed one device directly over top of another device. A *Bump* gesture occurred whenever a user made temporary physical contact between two devices. These gestures were quite popular, being the underlying metaphor in 10% of users' preferred gestures. Because these gestures can only be performed when the user is physically close to their target device they are not generalizable across distance.

Moving Objects Theme

Data transfer lends itself naturally to the idea of moving object in physical space. People have a wide variety of experience with this because we live in a world filled with physical objects. Because there are many ways to move an object such as sliding, flicking or throwing, we saw considerable variety in this theme. These gestures were sometimes performed using hands, such as in the *Throw* gesture where a user would mimic the act of throwing an object. In other instances these gestures were performed using the device itself, such as in the *Chucking* gesture where the user would simulate the act of tossing the device.

Gestures in the *Moving Objects* theme has the largest number of variations with 24 distinct gestures, making up 40% of total gestures. While these gestures were popular, some users mentioned that repeatedly performing them would become fatiguing.

Borrowed Interactions Theme

Most of the previously described themes are based on activities and concepts from human interaction and everyday life. However, we also saw some gestures that were clearly borrowed from other experiences with existing technologies. In some cases, these interactions with technology can become a metaphor on which gestures are based. This was seen with two gestures, the *Camera* gesture and *Mirror* gesture.

In the *Camera* gesture, a user would hold their device up similar to using a camera. This gesture was overwhelmingly used for retrieving data (similar to [3]). Another gesture based on this idea was the *Mirror* gesture, where a user would reverse their tablet and expose the screen to their target device. Interestingly, this gesture was used for both sending and retrieving data.

Selection Theme

Gestures in this theme were often created by users for data transfer tasks. This was sometimes as simple as touching, either with a single finger or a whole hand. Another example of this identifying theme appeared in the *Point* gesture, which involved the user pointing with their finger or the *Devices Point* gesture, where users would use a device to point at a target. Often these interactions were used in conjunction with another gesture, an issue we discuss in the next section.

Compound Gestures

About 43% of the gestures we saw for commands were some kind of *compound* gesture. By this we mean that the gesture had several distinct phrases in the terminology of Buxton [6]. Each phrase can be considered as an *atomic* gesture, and taken all together as a *compound* gesture. In such a gesture the phrases are performed sequentially—one following another temporally. In some cases, users suggested that the command should be performed by two users, each performing a phrase simultaneously. We also considered these to be compound gestures and they clearly emphasize the social aspect of the task.

Gesture Name	Description	# times
Close Contact Theme		
Bump	Device is physically contacting another device for a brief period of time.	36
Position On	Device is placed down on top of a target.	26
Position Near	Device is placed very close to the target device, but not above or exactly beside or on top of.	12
Position Beside	Device is placed exactly beside target, such that the two devices line up.	11
Position Above	Device is placed over top or directly above its target.	10
Selection Theme		
Touch Five	Five fingers touch the device.	117
Point	User air-points with a single finger at target.	106
Single Tap	User taps device with one finger.	63
Point Device	Device is pointed toward target.	29
Touch Hand	Device is touched with a flat hand (palm or back of hand).	20
Double Tap	User taps device twice with a finger.	19
Borrowed Interactions Theme		
Camera	Device is held vertically, screen towards the user, mimicking the action of taking a picture with a camera.	52
Mirror	Device is held vertically, and inverted, exposing the screen to the target.	51
Lasso	The user traces and closes a distinct loop in the air with a finger.	11
Moving Objects Theme		
Throw	Performed in the air, a gesture mimicking the action of tossing or throwing an object.	99
Swipe Up	Performed on a device, with an arbitrary number of fingers, in the upwards direction.	96
Grab	Performed in the air, a gesture where the user mimics grabbing an object.	67
Swipe Down	Performed on a device, with an arbitrary number of fingers, in the downwards direction.	45
Shake Device	Performed with a device which is shaken in a rapid manner.	30
Swipe Device	Performed over two devices, a swipe which continues across a second device.	21
Rocking Device	Device is moved up and down in a rocking motion along its horizontal axis.	16
Chucking Device	User mimics a throwing action with the device.	12

For example a user might rotate their tablet and then use their tablet to point at a target device. We considered these as two individual atomic gestures (rotation, and the pointing). In one instance, a user suggested a *Touch Simultaneous* gesture, where two devices would be placed near one another, and then one (or two) users could tap the device screens simultaneously.

Pick & Drop Interaction Pattern

One particular pattern of compound gestures was very common, and following the pick-and-drop metaphor [21] very closely. These compound gestures involved two distinct atomic gestures, one for selecting content on the source device and the other for selecting the target device. This pattern was used for both the sending and retrieval of data, with users varying the individual gestures, depending on distance. For example, when a user was far away from the target they would perform *a Five Finger Touch* gesture to select an image on their tablet, followed by a *Throw* gesture in the direction of their target. But if the same user was close to a target device, they would again perform a *Five Finger Touch* gesture, but now would follow it with another *Five Finger Touch* gesture.

Two-Party Interactions

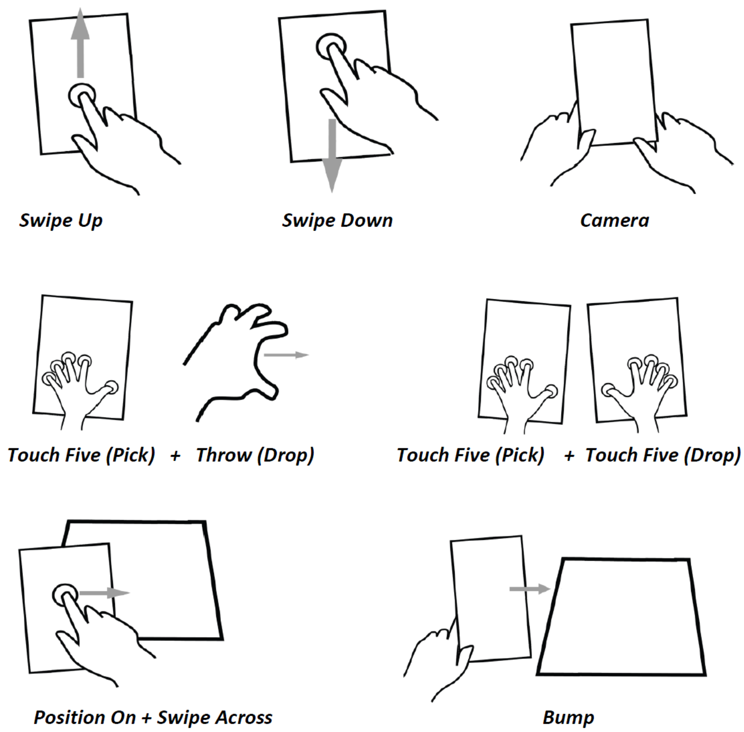
When a participant was asked to perform a gesture for tasks that involved another personal device (e.g. tablet to tablet), they would sometimes ask another user to perform a gesture on their device as well.

This was most prevalent in tasks where the user was asked to retrieve an image from the other user. Participants indicated that this was because actions that involved others' devices should be negotiated, rather than be unilateral actions.

Impact of Factors

In addition to the conceptual themes we wanted to investigate how the kinds of gestures users chose would change depending on certain factors. These factors were the device involved in the task, the distance at which the task was performed at and the action that the user was attempting to accomplish.

To focus our analysis, we considered only the gestures that participants identified as their favorite (recall that they



Task	Gesture 1	Gesture 2	Agreement Score
Send to Tablet (Near)	Swipe Up	-	0.26
Send to Tablet (Far)	Swipe Up	Touch Five (Pick) + Throw (Drop)	0.33
Retrieve from Tablet (Near)	-	-	0.11
Retrive from Tablet (Far)	Swipe Down	-	0.13
Send to Wall Display (Near)	Swipe Up	Bump	0.20
Send to Wall Display (Far)	Swipe Up		0.18
Retrive from Wall Display (Near)	-	-	0.07
Retrieve from Wall Display (Far)	Swipe Down	-	0.10
Send to Tabletop (Near)	Swipe Up	-	0.21
Send to Tabletop (Far)	Swipe Up	-	0.27
Retrieve from Tabletop (Near)	Position On + Swipe Across	-	0.09
Retrieve from Tabletop (Far)	Swipe Down	-	0.08
Send from Tabletop to Wall Display	Touch Five (Pick) + Touch Five (Drop)	-	0.08
Send from Wall Display to Tabletop	Camera	-	0.10
Retrieve from Wall Display with Multiple Images (Near)	Touch Five (Pick) + Throw (Drop)	-	0.10
Retrieve from Wall Display with Multiple Images (Far)	Touch Five (Pick) + Throw (Drop)	-	0.12

Figure 3: Prevalent gesture for each command. Generated by removing all gestures in a command with fewer than three observations.

performed three gestures for each command before selecting one as their favorite). Due to longitudinal nature of the study, we analyzed the data using the Generalized Estimating Equations (GEE) method using dummy variables for each of the categories in our proposed metaphor set. The data was aggregated and we used a Poisson distribution for the analysis.

Impact of Distance

A primary goal of our study was to investigate how the gestures users chose would change, depending on whether they were close to or far from a target device. This is important because an MDE can sometimes encompass an entire room and displays will not always be close to a user.

We found that far distances have a statistically significant impact on certain gesture types. Two gesture themes are significantly & positively predicted by far distances, *Borrowed Interactions* (B = .253, p = 0.045) and *Non-Gestural* (B = 0.799, p = 0.001). The *Close Contact* theme however was significantly & negatively predicted by far distances (B=-3.773, p<.001) as it was not possible to achieve contact at that distance. This suggests that using gestures which work on all displays might remove some gestures which users prefer.

Impact of Devices

Some of the gestures in our study were device specific and others could be used on any of the component devices. In addition to looking at how the preferred gestures were impacted by distance, we wanted to consider how they would be affected by the type of device.

Analyzing the data we found that only the *Close Contact* gesture theme was statistically significantly impacted by the overall device factor(χ^2 (2)=15.144, p=0.001). Since there are three categories of the devices, therefore two dummy variables (i.e. one for the iPad & one for the wall display) are created & the tabletop is used as the reference category). This theme was negatively but not significantly predicted by the iPad condition (B=-3.26, p=0.117) and significantly & negatively predicted by the wall display condition (B=-0.622, p<0.001). This might suggests that contact gestures could be used for only the appropriate device types.

DISCUSSION

We designed our study around the procedure outlined in previous work [27,20] to develop a gesture set for moving content in multi-display environments. Our results show that unlike previous studies using this procedure, very little convergence was observed. This is an interesting result in itself and the implications of it are discussed in this section.

In spite of this lack of convergence our study produced several contributions which will be useful to system designers. The most prevalent favorite gestures were presented along with agreement scores to provide some concrete examples for designers. In addition the atomic gestures which participants created were classified into several thematic categories. These categories can serve as a basis for creating new gestures.

In this section, we discuss the direct implications of our results—that is, how designers can apply the thematic categories and example gestures to their systems. We then consider offering multiple gestures for each command or gesture aliasing. Finally we consider the notion of modifiers and feedback, issues that our participants brought up during the study.

Implications of Low Convergence

Wobbrock et al. present a metric of convergence called an agreement score which is a quantitative measure of the variability of gestures produced for a specific referent or command [27]. Low agreement scores imply a low level of convergence, which means that participants tended to produce different gestures for the same command. With sufficiently low agreement scores it is difficult to justify the creation of a definitive user-define gesture set for the commands being investigated.

Our results demonstrate that participants showed very little convergence in the gestures they produced for the tasks. This confirms some of the qualitative observations made by previous MDE gesture elicitation studies [15,16] but contradicts the results found in tabletop gesture elicitation studies [9,27]. While we cannot explain exactly why our participants' gestures did not converge, there are a number of possible explanations: the space of possibilities afforded by the MDE environment, the varying backgrounds of our participants, and possibly the lack of prior experience in an MDE space.

The previous elicitation studies for tabletop gestures have generally considered a single display surface - accordingly, most gestures occur on or near the display in a planar fashion. Our work explores a multi-display context meaning that the space of possibilities for gestures is much greater. Not only are some of the displays physically mobile, but there is also the space “between” displays that can be considered as part of the interaction space. That is, some of the implied "bounds" of what an appropriate gesture are taken away, freeing people to an even wider space of possible gestures.

Another explanation could be that participants, who did not have meaningful experiences with MDEs, accordingly did not have a strong and consistent mental model upon which to draw to create gestures. We strongly suspect that this is the case and our work in developing categories attempts to uncover the several different mental models which may be underlying the gestures. Rather than presenting a specific gesture set, we provide these categories. In so doing, we aim to give designers something to draw on to develop new gestures.

Applying the Gestures & Themes

Aliasing Gestures

Supporting multiple gestures for a specific command is one way of handling the divergence our study has found in the elicited gestures. Since we did not see any instances where gestures conflicted with one another in terms of their meaning, designers could provide aliases, where multiple gestures result in the same action. For instance, *Position {On, Near, Beside, Above}* were all expected to produce a similar result. Similarly several gestures, such as *Chucking*, *Swiping*, *Throwing*, etc. were expected to produce similar results, even though the specific nature of the gesture might be very different (with-device, on-device, in-the-air)—these can all be designed as aliases of one another.

Gestures and Themes

The themes that we identified for atomic gestures relate to the nature of these gestures. These themes can provide conceptual guidance for designers: gestures following these kinds of themes will likely be more easily discoverable as they map to users' conceptual model of how an MDE works. Of particular note, many of the conceptual models borrowed from an implicit understanding of the spatial relationships between devices in the room (i.e. proxemics [1] - that devices are considered as independent from one another, that they occupy certain parts of space, and so forth). One aspect that we did not consider deeply in this study was how these relationships might be changed if there are other people in such an environment. We saw indications that people would take social relationships into account (with the tablet-to-tablet transfers); however, this was not the central focus of our work.

Gestures for Distance & Device

Our analysis of the impact that the distance and device factors had on the type of gesture produced is useful when choosing gestures. Users often chose gestures of the *Borrow Interaction* type for far distances and a designer could pick a gesture such as *Camera* for those circumstances. Likewise *Close Contact* gestures were negatively associated (and not possible) with far distances and designers will need to provide an alternative for those situations. A gesture which can only work at a close distance such as *Bump* could be included along with a distance independent gesture such as *Swipe Up.* However *Close Contact* gestures are negatively predicted when a wall display is the target device, therefore designers might opt for different gestures if a wall display is a major component in their system.

Limitations and Future Work

While we encouraged participants to freely create gestures as they felt would be appropriate, it seems clear that their existing and current experiences with other technologies, limited (to some extent) the kinds of gestures that we saw. These experiences might include, but are not limited to, actual experiences with touch technologies, and also video games or movies. This may have resulted in a sub-optimal set of gestures; however, we are encouraged that the gestures they produced correspond to considerable prior work in this space, as well as to the spatial relationships between devices in the space. This suggests that proxemics may be a very fruitful way of thinking about MDEs.

In a future study, we would like to follow on work performed by Morris et al. by evaluating the metaphor set we have proposed in this paper [18]. This could be done by working with designers to create new gestures or by evaluating in a traditional way, a selection of the most common gestures found in our study. We also would like to work on evaluating complete MDE systems which can recognize and incorporate these kinds of gestures into their interactions.

CONCLUSION

In this paper, we designed a study to elicit gestures from potential users of MDEs. We looked specifically at gestures for data transfer tasks because of the importance of the problem in the MDE research space. We quantified the level of agreement between the gestures produced by participants for specific commands and found it to be low. We considered the potential causes for this low agreement and the implications for designers in creating systems. We classified the types of atomic gestures we saw into four conceptual themes that provide basis from which designers can design new gestures. Notably, these gestures often implicitly made use of the spatial relationships between devices in the environment. We also suggested that designers consider providing several different gestures for a single command (i.e. gesture aliasing). As multi-display environments become a reality, users will increasingly need to deal with the issue of moving content from one display to another. The results from this paper inform designers how to design gestural interactions that should be easily discoverable and usable by these users.

ACKNOWLEDGMENTS

We would like to thank our colleagues Abraam Mikhail, Daniel Sabourin, Abhishek Sharma, Rojin Vishkaie, Nadder Makhlouf, Patrick King and Nicole Sultanum for their useful discussions, edits and advice. We also thank the anonymous reviewers for their careful and valuable comments and suggestions. This research was partially funded by the NSERC/AITF/Foundation CMG Industrial Research Chair in Scalable Reservoir Visualization, and by the NSERC SurfNet Research Network.

We would also like to specially thank our colleague, Dr. Tak Fung, for his valuable help and steering us on the correct path with many discussions and support.

REFERENCES

1 Ballendat, T, Marquardt, N, and Greenberg, S. Proxemic interaction: designing for a proximity and orientation-aware environment. In Proc. of ITS '10,121-130.

2 Biehl, J T and Bailey, Brian P. ARIS: an interface for application relocation in an interactive space. In Proc. of GI '04, 107-116.

3 Boring, S., Baur, D., Butz, A., Gustafson, S., Baudisch, P.

Touch projector: mobile interaction through video. In Proc of CHI 2010,2287-2296.

4 Bragdon, A., DeLine, R., Hinckley, K., Morris, M. Code space: touch + air gesture hybrid interactions for supporting developer meetings. In Proc ITS 2011, 212-221.

5 Brignull, H, Izadi, S, Fitzpatrick, G, Rogers, Y, and Roddem, T. The introduction of a shared interactive surface into a communal space. In Proc. of CSCW 04, 49-58.

6 Buxton, W. Chunking and phrasing and the design of human-computer dialogues. In Human-computer interaction, Morgan Kaufmann Publishers In, pp. 494–499.

7 Dachselt, R., Bucholz, R. Natural throw and tilt interaction between mobile phones and distant displays. In Proc of CHI EA 2009, 3253-3258.

8 Döring, T, Shirazi, A, and Schmidt, A. Exploring gesture-based interaction techniques in multi-display environments with mobile phones and a multi-touch table. In Proc. of AVI '10, 419-419.

9 Epps, J.; Lichman, S. & Wu, M. A study of hand shape use in tabletop gesture interaction, Ext. Abstracts CHI 2006, 748-753.

10 Frisch, M., Heydekorn, J., Dachselt, R.: Investigating Multi-Touch and Pen Gestures for Diagram Editing on Interactive Surfaces. In Proc. of ITS 2009,149-156

11 Hassan, N., Rahman, M., Irani, P., Graham, P. Chucking: A One-Handed Document Sharing Technique. In Proc of Interact 2009, 264-278.

12 Heydekorn, J., Frisch, M., Dachselt, R.: Evaluating a User-Elicited Gesture Set for Interactive Displays, In Proc. of Mensch und Computer 2011, 191-200.

13 Hinckley, K. Synchronous gestures for multiple persons and computers. In Proc of UIST 2003 , 149-158.

14 Johanson, B, Fox, A, and Winograd, T. The Interactive Workspaces project: experiences with ubiquitous computing rooms. IEEE Pervasive Computing, 1, 2 (Apr-Jun 2002), 67-74.

15 Kray, C, Nesbitt, D, Dawson, J, and Rohs, M. User-defined gestures for connecting mobile phones, public displays, and tabletops. In Proc. of MobileHCI '10, 239-248.

16 Kurdyukova, E, Redlin, M, and André, E. Studying user-defined iPad gestures for interaction in multi-display environment. In Proc. of IUI '12, 93-96.

17 MacKenzie, R, Hawkey, K, Booth, K S, Zhangbo, L, Perswain, P, and Sukhveer, D S. LACOME: a multi-user collaboration system for shared large displays. In Proc. of CSCW '12 ,267-268.

18 Morris, M R, Wobbrock, J O, and Wilson, A D. Understanding users' preferences for surface gestures. In Proc. of GI '10, 261-268.

19 Nacenta, M A, Sakurai, S, Yamaguchi, T et al. E-conic: a perspective-aware interface for multi-display environments. In Proc. of UIST '07, 279-288.

20 Nielsen, M., Störring, M., Moeslund, T.B. and Granum, E. A procedure for developing intuitive and ergonomic gesture interfaces for HCI. In Proc. Int'l Gesture Workshop, 409-420.

21 Rekimoto, J., Saitoh, M. Augmented surfaces: a spatially continuous work space for hybrid computing environments. In Proc of CHI 1999 , 378-385.

22 Rekimoto, J. Pick-and-drop: a direct manipulation technique for multiple computer environments. In Proc. UIST'97, 31-49.

23 Streitz, N.A., Geibler, B. i-LAND: an interactive landscape for creativity and innovation. In Proc. CHI 2002, 120-127.

24 Voida, S, Podlaseck, M, Kjeldsen, R, and Pinhanez, C. A study on the manipulation of 2D objects in a projector/camera-based augmented reality environment. In Proc. of CHI '05, 611-620.

25 Wallace, J, Ha, V, Ziola, R, and Inkpen, K. Swordfish: user tailored workspaces in multi-display environments. In Proc. of CHI '06, 1487-1492.

26 Wilson, A D and Benko, H. Combining multiple depth cameras and projectors for interactions on, above and between surfaces. In Proc. of UIST '10, 273-282.

27 Wobbrock, J O, Morris, M R, and Wilson, A D. User-defined gestures for surface computing. In Proc. of CHI '10, 1083-1092.