Journal of Mobile Multimedia, Vol. 12, No.3&4 (2017) 243-264

A COMPARISON OF TECHNIQUES FOR CROSS-DEVICE INTERACTION FROM MOBILE DEVICES TO LARGE DISPLAYS

JENI PAAY, DIMITRIOS RAPTIS, JESPER KJELDSKOV,
Department of Computer Science, Aalborg University, Denmark
{jeni, raptis, jesper}@cs.aau.dk

BJARKE M. LAURIDSEN, IVAN S. PENCHEV, ELIAS RINGHAUGE and ERIC V. RUDER
Department of Computer Science, Aalborg University, Denmark

In recent years there has been an increasing interest in cross-device interaction research involving mobile computing. We contribute to this research with a comparative study of four interaction techniques for moving information from a mobile device to a large display. The four techniques (Pinch, Swipe, Throw, and Tilt) were compared through a laboratory experiment with 53 participants, measuring their effectiveness, efficiency and error size. Findings from the experiment revealed that the Swipe technique performed best on all measures. In terms of effectiveness, the Tilt technique performed the worst, and especially so with small targets. In terms of efficiency and error size, the Pinch technique was the slowest and also the most imprecise. We also found that target size mattered considerably for all techniques, confirming previous research. Based on our findings we discuss why the individual techniques performed as observed, and discuss implications for using mobile devices in cross-device interaction design.

Key words: Cross-Device Interaction, mobile devices, large displays, Interaction Techniques, Kinect, Mid-air Gestures.

1 Introduction

As the number of interactive computing devices around us continues to grow, there is an increasing need for research investigating how to best facilitate people's interaction using their mobile devices, with other devices, in concert. This has led to a growth in human-computer interaction research on "cross-device" and "digital ecosystem" interaction. This research has investigated how cross-device interaction can be applied in practice (e.g. [11]), how it is used and understood by users, (e.g. [19]), how it can be modelled conceptually (e.g. [25]), how it can be implemented (e.g. [7]), and how it can be supported by concrete interaction techniques (e.g. [14]). In terms of the latter, the opportunities and challenges of cross-device interaction using mobile devices, in particular, have inspired a renewed focus on new interaction techniques that go beyond traditional point-and-click type interactions. Instead, researchers have explored new input technologies, such as the accelerometers in smartphones [5], radio modules [14], and the use of mobile phones as styluses on touch tables [22]. The use of these interactions for mobile devices allows people to operate several devices in parallel with their personal device, and to move information and activities easily from one device to another. This has been investigated for many different cross-device combinations, for example, phones and watches (e.g. [5]), mobile devices of different sizes (e.g. [24]), phones and table tops (e.g. [23]) and phones and large displays (e.g. [3]).

Although previous investigations cover different aspects and qualities of various cross-device interaction techniques, there are limited investigations and empirical studies specifically comparing techniques used to transfer data from a personal mobile device to a large display. The reason that this area became the focus of this paper is the fact that these systems, such as large public displays that encourage passers-by to interact with them using their personal mobile devices, are becoming more prevalent, for example, in public squares and building foyers. This therefore is driving our interest in learning more about alternative interaction techniques for such systems and more importantly to evaluate techniques that transition interaction from one device to another, since a comparative study focussing on the strengths and weaknesses of different techniques in terms of their effectiveness, efficiency and accuracy would give interaction designers valuable knowledge about the performance of these techniques, useful for informing the design of cross-device systems.

In the work presented here, we have specifically investigated different gesture-based interaction techniques for facilitating interaction that begins on a handheld mobile device, and continues on a large display. Within this scope we have investigated the relative strengths and weaknesses of different techniques making use of combinations of sensors in the phone and sensors mounted under the large display. We have done this through an experiment with 53 participants, measuring the respective effectiveness, efficiency and error size of four existing interaction techniques. We begin this paper by providing an overview of related work in cross-device interaction. We then present our four interaction techniques, *Pinch*, *Swipe*, *Throw*, and *Tilt*, and the experiment comparing their use for cross-device interaction from a mobile device to a large display. We then report our findings, and discuss their implications for the design of cross-device interaction.

2 Related Work

The idea of cross-device interaction can be traced back to the work of Rekimoto [18] who envisioned what was called "multiple-computer user interfaces", and argued that dedicated interaction techniques would be needed to overcome the boundaries among devices in multiple-computer environments. Since then, a large amount of research has been conducted following this line of thinking, and investigating, among others, how cross-device interaction can be applied in practice, how it is used and understood by users, how it can be modelled conceptually, how it can be implemented, and how it can be supported by concrete interaction techniques. While exploring the notion of ubiquitous computing, Rekimoto [17] devised the pick-and-drop technique that operates between multiple devices allowing users to "pick up" an object on a display and "drop it" on another display as if manipulating a physical object. A similar approach was also presented by Dachselt et al. [6], where users could "throw" digital context to a large display from distance. This use of natural gestures has continued as a useful analogy for the design of interaction techniques that transfer data from one device to another.

One of the earliest operational cross-device applications, presented by Myers, in 2001 [15], was the Pebbles Slide Show Commander, which utilized Personal Device Assistants (PDAs) to control a PowerPoint presentation running on a separate computer or laptop. Participants could remotely initiate a move between slides by enacting it on the PDA. Additionally, annotations made on the PDA screen would be shown on the presentation screen for the audience. This was an early investigation of the use of personal mobile devices to interact with fixed displays.

In more recent research, Boring et al. [2] built a cross-device application to explore the implications of different approaches to transferring data from a large public display onto a mobile device. One approach was to use the camera on the smartphone. The user simply had to take a picture of the information they wanted to transfer to their phone. A content server then visually analysed the picture to determine which content the user was interested in and then sent that content to their phone. In this study, Boring et al. confirmed the need for enabling data exchange between mobile devices and public displays. In a follow up study, Boring et al. [3] compared three different interaction techniques, Move, Tilt and Scroll, to continuously control a pointer on a large screen using a mobile device. They found that Move and Tilt enabled a faster selection time compared to Scroll, but at the cost of higher error rates. Furthermore, other researchers moved beyond performance measures and also explored the impact of such interactions on the human body, such as arm fatigue [8].

While Boring et al. [3] focused on passing data between public and personal displays, research by Nielsen et al. [16] explored techniques for aligning multiple mobile devices of varying sizes to present common content. Their collaboration surface combines multiple devices, which then appear as one larger collaborative workspace. Two or more mobile devices aligned along an edge can be “pinched” together to show a single image, which adjusts seamlessly across all screens as additional devices are added or taken away. In this way, the interaction area can expand or contract depending on the number and size of devices that make up the surface. This pinching motion, like the pick-and-drop motion proposed by Rekimoto [17], indicates the successful application of a technique that uses a natural motion of putting things together, to achieve the digital equivalent.

Also creating a common workspace, Schmidt et al. [22] proposed a cross-device interaction style for mobiles and surfaces where multiple phones could be used to interact with a digital surface. They considered the role of natural user interaction when using multiple devices and stated that “natural forms of interaction have evolved for personal devices that we carry with us (mobiles) as well as for shared interactive displays around us (surfaces) but interaction across the two remains cumbersome in practice”. To make the interaction more natural, they proposed the use of mobile phones as tangible input devices on the surface in a stylus like fashion. This indicates the need for investigation into identifying the kinds of interactions that might work more effectively between mobiles and surfaces.

Marquardt et al. [14] and Bragdon et al. [4] also investigated cross-device interaction in relation to natural user interactions. Marquardt et al. studied cross-device interaction on tablets using natural modes of communication, involving spatial information through proxemics. Based on the constructs of f-formation, micro-mobility, and co-present collaboration, they built a prototype for document transfer that supports fluid and minimally disruptive interaction. Bragdon et al. proposed *Code Space*, a system using a combination of mid-air gestures and touch to support co-located, small group developer meetings. Their interaction techniques used a combination of in-air pointing and touching with precise gestures and mobile devices. This allowed a group of users to interact, using air and touch gestures to control and share items across multiple personal devices such as smartphones and laptops to a large multi-touch display. Although this research explores different interaction options using mobile devices and a shared display, they do so in a qualitative way, with a pilot user study collecting user feedback on the various techniques used in *Code Space*. In fact, in looking at future work, the authors themselves support the value of a quantitative study of these techniques to inform further knowledge about cross-device interactions.

Skov et al. [24] compared six different cross-device interaction techniques in a quantitative study both in the lab, and in the field. They used the case of card playing to study techniques. In their experiment, a player could see their “hand” of cards on their phone and use three different “play” techniques to play a card to a tablet that represented the common play area. Players could also draw a card from the play area (tablet) using one of three “draw” techniques. The study found differences in time and error rates between techniques, e.g., the swipe gesture caused significant interaction errors while trying to swipe a card from mobile phone to the shared tablet. More interestingly to our study, they found that those techniques that mimicked the natural gesture of playing cards were slower than others and participants found them less useful when playing an actual game. This appears to be in contrast to other studies we have looked at, where the more natural gestures were the recommended ones.

In an effort to formalize knowledge about data transfer between devices, Hamilton and Wigdor [7] created *Conductor*, a prototype framework for cross-device applications, which acts as an exemplar for the construction of cross-device applications. *Conductor* was designed to provide a set of generic interaction techniques, generalized across different applications, allowing several forms of cross device interaction and enabling users to work simultaneously across multiple devices and by which users can easily share information. The set of inter-device communication mechanisms were inspired by mechanisms and transfer concepts found in related research. They also present an example usage scenario as well as a user study to evaluate Conductor. A qualitative evaluation was based on observed behaviours and interviews with participants as feedback on the success of *Conductor*, in respect to which mechanisms they chose to use to complete the given task.

3 Experimental Method

From our literature review we found that quantitative research on cross-device interactions for transferring data from a mobile device onto a large display is currently limited in HCI. To be able to contribute new knowledge about the comparative strengths and weaknesses of different cross-device interaction techniques, we devised an experiment to empirically measure the hit success rate, the time taken to transfer data and the distance of error for missed targets of four different techniques (*Pinch, Swipe, Throw* and *Tilt*) in a laboratory situation. We chose to conduct our study as a laboratory experiment in order to be able to better control possible confounding variables that could influence our results.

3.1 Interaction Techniques

The four techniques implemented in this experiment were designed with respect to techniques presented in previous research as being of interest to the cross-device interaction problem. By combining different qualities of the techniques studied we were able to make a diverse and yet representative set of interaction techniques. The four selected techniques of our study were designed to be different to each other in respect to several qualities, to evaluate them against each other in a controlled situation.

There are many different selection criteria that could have been used, but we decided to use criteria raised with respect to the gestures used in studies found in the literature. This resulted in looking at existing techniques based on the following attributes: number of hands used to perform the technique, and using hands or mobile phones as a pointer to the large screen. After careful consideration of existing techniques, we decided that two of our techniques should use one hand (*Swipe, Tilt*) and the other two should require the use of both hands (*Pinch, Throw*). The method of moving the cursor on the large display would use the hand or finger as a pointer for two techniques (*Pinch, Throw*) and the phone as a pointer for the other two (*Swipe, Tilt*). Furthermore, we were also interested in related work that looked at natural gestures and therefore wanted two gestures that were mobile device-centred (*Swipe, Tilt*) to be compared against two others that used natural bodily gestures (*Pinch, Throw*) in an analogous way.

In the end, our four techniques, *Pinch*, *Swipe*, *Throw* and *Tilt*, were designed to represent existing attributes of techniques, and at the same time by making hybrid gestures, uncover new opportunities and challenges in this design space. The following sub-sections detail the studies these techniques were inspired by and why this technique was interesting to us.

3.1.1 Pinch

The *Pinch* technique (Figure 1) was used by Ikematsu and Siio [10] as part of a drag-and-drop method for moving data objects between devices. Chen et al. [5] also use a pinching gesture for cross-device interaction between a smartphone and a smartwatch to control volume. Benko and Wilson [1] used a Pinch technique for interacting with omni-directional visualizations in a dome. Nielsen at al. [16] used it to indicate a joint seam between mobile devices. Our technique is a combination of these techniques, and the technique used by Scheible et al. [21] to share mobile multimedia art on large public displays.

The *Pinch* technique was included to mimic the natural action of picking up a real object, e.g., a piece of paper, and moving it to another location. With *Pinch* we have a two handed technique which requires the user to perform a series of steps and therefore has high complexity compared with other techniques tested. *Pinch* is performed by:

1) Holding the phone in one hand and making a pinching gesture on the phone's screen with the other hand (fig. 1a), subsequently closing the hand. This step in the process is a selection step where users inform the system which piece of information they want to transfer to the large screen.
2) Then using the pinched hand to point at a target location on the large display (fig. 1b). This step allows users to explore where in the large display to place the selected piece of information.
3) Finally, opening the hand completes the data transfer (fig. 1c) and the piece of information is placed on the selected area on the large display.

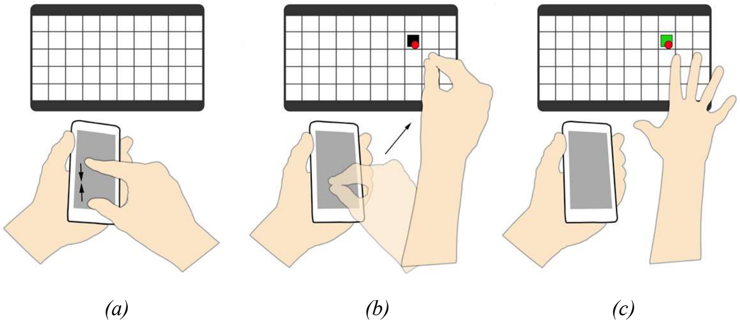

Figure 1. The Pinch technique.

3.1.2 Swipe

The *Swipe* technique (Figure 2) is used in the *Code Space* system [4]. They describe the technique as, "*cross-device interaction with touch and air pointing*" and the swipe motion is described as "*flicking up on the touch screen*". This technique is also similar to the swipe technique used by Skov et al. [24] to send cards from a mobile device to a shared tablet in a card playing game.

This technique was chosen to complement *Pinch*, with its low level of complexity and single-handed interaction. It is mobile device-centred and is similar with the familiar gesture used on smartphone touch screens to scroll content. The amount of time required to execute this technique is relatively low. *Swipe* is performed by:

1) pointing at an area on the large display with the phone in an out stretched hand (fig. 2a),
2) making a forward swipe motion with the thumb onto the phone's screen in order to move a piece of information from the phone to the selected area (figs. 2b,c).

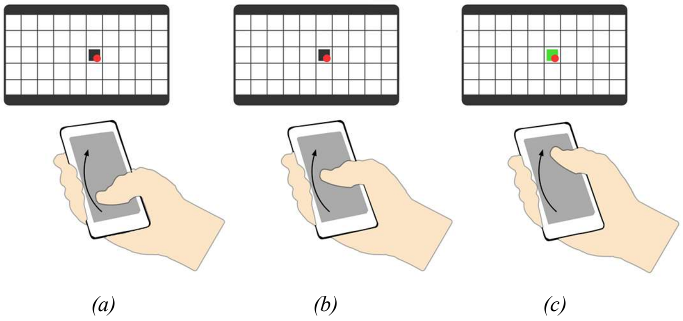

Figure 2. The Swipe technique.

3.1.3 Throw

The *Throw* technique (Figure 3) is a combination of a technique for pointing from Scheible et al. [21], that is, using a hand as a cursor in mid-air, and the throw technique described by Walter et al. [25] used for sharing information to large public displays. The inclusion of this type of natural user interaction technique is inspired by the call for more investigation into the use of natural gestures in the area of cross-device interaction by Schmidt et al. [22].

This technique was included based on its natural feel and playful design. The technique has a natural bodily gesture that mimics the real world scenario of throwing something like a ball. *Throw* is two-handed, and is the most complex of our techniques. It takes a little longer to execute because of the number of steps required. *Throw* is performed by:

1) pointing at an area on the large display with one hand (fig. 3a),
2) holding the phone in the other hand and tapping on the smartphone to select the piece of information the user wants to transfer (fig. 3b),
3) making a swinging motion with the phone towards the large display to finalize the transfer of data (fig. 3c).

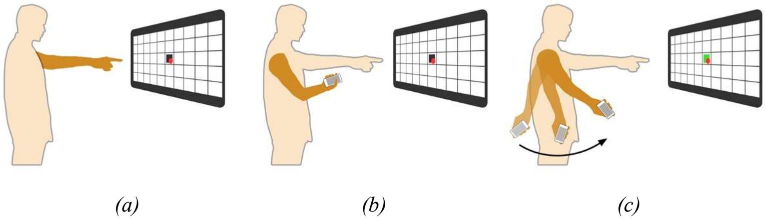

Figure 3. The Throw technique.

3.1.4 Tilt

The *Tilt* technique (Figure 4) is used in a collaboration system by Lucero et al. [12] to transfer an object from a large display to the user's smartphone. Our *Tilt* is a copy of Lucero et al.'s technique, but in reverse. Boring et al. [3] also use a tilt technique when directing a pointer on a large display using a phone. We chose this technique because it is one-handed, has a relatively low complexity, and like *Swipe*, is relatively fast to use. *Tilt* is performed by:

1) pointing at an area on the large display with the phone in an out stretched hand (fig. 4a),
2) selecting a target on the phone screen (fig. 4b),
3) tilting the phone forward (fig. 4c) to complete the data transfer.

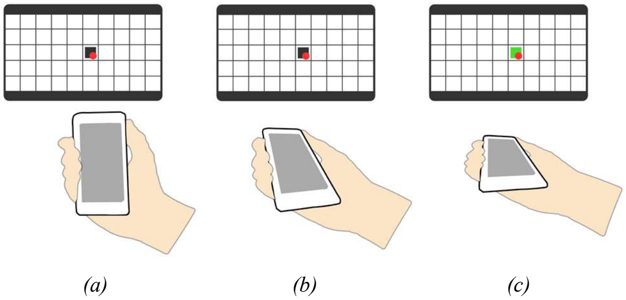

Figure 4. The Tilt technique

3.2 Experimental Design

The experiment was conducted as a within-subject design, with the four different *interaction techniques* and two *target sizes* as independent variables. This decision ensured that all 4 interaction techniques (*Pinch, Swipe, Throw* and *Tilt*) were performed in the experiment by all participants, thus maximising the number of data collected for all techniques. It also ensured that the possible different participants' characteristics would be similar in all experimental conditions. We also used two target sizes in our experiment to investigate the influence of target size on the results of the different techniques.

3.2.1 Participants

In total, 53 people took part in our experiment, which was conducted in a usability lab. Each filled in a short demographic survey. The participants where between 20-45 years old (M: 24.4, SD: 4.3) and were between 1.63 and 1.95 meters tall (M: 1.82, SD: 7.8). 88.7% of users were right handed, 90.6% were male, and 96.2% of them were smartphone users. Of those who owned smartphones, they had owned them for 2-15 years (M:5, SD: 2.1). Participants were recruited through a combination of social networking and posters around campus.

3.2.2 Experimental Setup

The purpose of our experiment was to investigate the performance of four representative interaction techniques that require the transfer of a selected piece of information to a specific area on a large screen. In order to make this transfer process more controlled and rigid, we took a series of actions in relation to the scenario our participants experienced. The number of pieces of information that the users could chose to transfer was limited to two, and in order to make the recognition process faster they were indicated as two shapes (circle and square, Fig. 6b). This decision allowed us to minimize the selection process time, a parameter that could had an impact on the performance of each technique, while still requiring the user to make an active selection before enacting a technique. This decision actively included the mobile phone in the interaction and simulated the real world situation of selecting data and moving it to a screen. The two shapes on the phone changed positions randomly, so users would have to check the phone for each target in order to select it. Users chose with which hand they held the phone and with which they pointed. They could also swap hands at any time during the experiment.

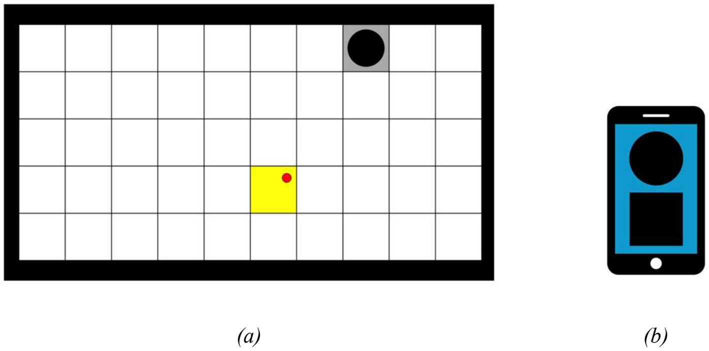
(a) *(b)*

Figure 5. The large display and the mobile phone screen.

The participants also needed to know, a) which of the two shapes to select in order to transfer to the screen, and b) where to place the selected shape on the large screen. Both needs were accomplished by using a target on the large screen that indicated both the shape as well as the placement position (grey shape, Figure 6a). Furthermore, a cursor was included in order to highlight where the participants were pointing (yellow shape, Figure 6a). The participants were presented with one target at a time. For each target and every technique, if the cursor was within the target's grid when the technique was performed, the system regarded it as a hit, and the target turned green. If the cursor was outside the target grid square, this was logged as a miss, and the target turned red.

Since we wanted to see if the target size has an effect on the interaction techniques we chose two include two different target sizes in our setup (small and large). We made sure that each participant experienced an equal number of small and large targets, and we randomised the order they were presented to them.

3.2.3 Tasks

After entering the usability lab, each participant was given a short introduction to the tasks they had to perform. We explained that they would experience four different interaction techniques that they would use to transfer data from the mobile device to the large display.

We then provided the participants with the smartphone, asked them to stand on a marked cross on the floor and commenced the test. The reason we used the marked cross was to experimentally control for the *distance* each participant stood from the screen. The mark was set at exactly 2.35m from the screen, based on the optimal operating distance for the Kinect.

The system then randomly chose one of the four techniques and played a short explanatory video of how to perform it on a second screen located beside the target display (See Figure 7a). We provided video instructions to make sure that all participants experienced the same narrative. Even though the four techniques were very different from each other, we chose to experimentally control for the *learning effect* and the *carry over effect* by mixing the sequence order the participants experienced the four techniques. In the end, the *Swipe* started 22.7% of tests, *Pinch* started 26.4% of all tests, *Throw* started 24.5%, and *Tilt* started 26.4% of tests.

(a) *(b)*

Figure 7. The experiment in progress.

Each participant was given three practice attempts per technique, in order to get familiar with it. After the practice phase, participants would go through the test, comprised of 18 attempts. In the end, each participant experienced three practice rounds (for each technique) which were not included in the results, and (9 small + 9 large targets) x (4 techniques) = 72 targets, as recorded data. The test per participant took on average 15 minutes. Repeat attempts at a target were not permitted and the system proceeded to the next target regardless of whether the participant hit the target, or missed. Since we deemed important that the distance the participants would have to cover on the large screen between the attempts would have an effect on the results, we chose to experimentally control for it by making sure that they would all experience the same distribution of distances for each technique. In order to successfully do so a calibration target would appear after the practice phase, and then all participants would experience the same distribution of distances in a random order.

3.2.4 System Setup

The experimental setup included a 65-inch Panasonic screen with 1920×1080 resolution mounted directly above a Microsoft Kinect (see Figure 5). The Kinect was mounted exactly 1m above floor level, based on an approximation of participants' average height. Each participant stood on a mark exactly 2.35m from the screen. From this point they used a Samsung Galaxy SII smartphone as their mobile device to complete the tasks.

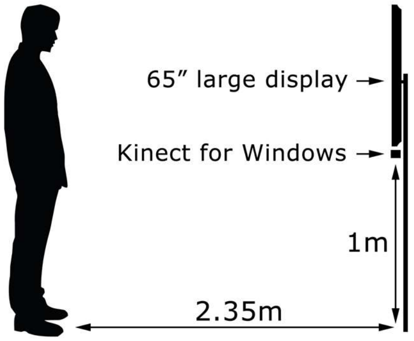

Figure 6. The System Setup.

3.3 Data Collection

A simple logging program was developed, which collected interaction data for each user. The collected data for all users included: a) a time stamp, b) whether they hit a target or not, c) whether they selected the correct shape on the mobile phone or not, d) the exact position of the target on the screen, e) the size and shape of the target (large-small, circle-square), and f) the exact position of the cursor on the screen when each attempt was performed.

From the collected dataset, the following measures were calculated and used in the analysis:

- *Effectiveness*: the number of successful attempts per user, per technique – a successful attempt, besides hitting the target, also required selecting the correct shape on the phone.

- *Efficiency*: the time each user spent to hit each of the targets - the time was initialized as soon as the target appeared on the screen,
- *Error size*: the distance between the cursor from the target in pixels, when a user missed a target.

Besides automatic logging of user's interactions, we also video recorded the experiment, both as a backup if the automatic data logging failed, but also to identify any technical problems with the gesture recognition software to be refined for any follow-up studies and to document participants' reactions when they experienced the four techniques.

4 Experimental Results

We analysed our data in respect to *effectiveness* (based on successful attempts), *efficiency* (based on time per target), and *error size* (based on distance to target). As each of the four interaction techniques were used 18 times per participant, in the end, each technique was performed 954 times.

The first step in the process was to clean the dataset up. We focused on each technique and removed the scores that belonged to outliers. Outliers were defined by applying the outlier-labelling rule and using 2.2 as a multiplier [9]. The majority of these scores came from participants that experienced glitches while interacting with the system. After removing outliers, the initial 4x954= 3816 attempts were narrowed down to 3564.

In the following subsections we present our experimental results for *effectiveness*, *efficiency*, and *error size*. All the statistical analyses were performed using SPSS.

4.1 Effectiveness

In this study, we define *effectiveness* as the number of successful attempts per user, per technique. The average successful attempts per user for each technique, for small and large targets respectively, are presented in Tables 1 and 2. Each user could have a maximum of 18 successful attempts (9 for the small targets and 9 for the large ones). In the resulting dataset we performed one-way ANOVA since our dependent variable (effectiveness) was an interval.

	Pinch	**Swipe**	**Throw**	**Tilt**
N=9	5.87	8.11	7.45	5.25

Table 1. Effectiveness: average successful attempts per user, per technique for small targets. N= maximum possible value.

For the small targets we identified significant differences among the techniques (F(3, 208)=34.713, p<.001). A pairwise comparison showed that all the techniques were significantly different from each other, except *Pinch* and *Tilt*, p=.054. Consequently, *Pinch* and *Tilt* had the lowest average successful attempts (5.87 and 5.25 out of 9, respectively, Figure 8), while *Swipe* the highest (8.11 out of 9, Figure 8).

	Pinch	Swipe	Throw	Tilt
N=9	7.02	8.64	8.4	7.02

Table 2. Effectiveness: average successful attempts per user, per technique for large targets. N= maximum possible value.

The same process was then followed for the large targets. Again we identified significant differences among the techniques (F(3, 208)=29.211, p<.001). A pairwise comparison showed that all the techniques were significantly different from each other, except *Swipe* and *Throw*, p=.284, and *Pinch* and *Tilt*, p=1.0. As was the case for the small targets, *Pinch* and *Tilt* performed worst (7.02 out of 9, Figure 8), while *Swipe* (8.64 out of 9, Figure 8) was the best followed by *Throw* (8.4 out of 9, Figure 8).

We then extended our analysis and we performed a two-way repeated measures ANOVA in order to examine the combined effect of technique and target size on effectiveness. The effect of each technique on effectiveness was significant (F(3, 416)= 62.264, p<.001) and so was the effect of target size (F(2, 416)= 62.285, p<.001). Furthermore, their interaction was also significant (F(3, 416)= 3.471, p=.016), showing that a combined effect of technique and target size on effectiveness exists.

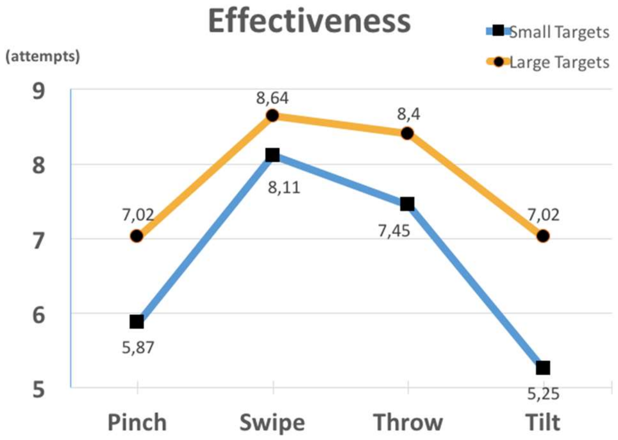

Figure 8. Effectiveness: Average successful attempts per user, per technique, for small and large targets.

4.2 Efficiency

By efficiency, we define the time it takes to complete an action as required by each technique. The number of attempts, the mean time to hit a target in seconds and the standard deviations (in parentheses) for each technique and for both target sizes, are shown in Tables 3 and 4, and combined in Figure 9.

To study the effect of each technique on efficiency for the different target sizes, we performed one-way repeated measures ANOVAs with a Greenhouse-Geisser correction, one for the small targets and one for the large ones. The reason to opt for a Greenhouse-Geisser correction was the fact that the repeated measures ANOVA assumption for sphericity was violated. The corrections provided us with new values for the degrees of freedom, while reducing the Type I error rate.

	Pinch N=448	**Swipe** N=449	**Throw** N=439	**Tilt** N=456
N=1741	7.94 (3.92)	5.67 (2.16)	6.32 (2.25)	6.03 (3.02)

Table 3. Efficiency: means and (standard deviations) of time spent for each technique per target for small targets.

For the small targets (Table 3), we identified significant differences among the techniques (F(2.333, 111.506)=119.799, p<.001). A pairwise comparison showed that all techniques were significantly different from each other (for all cases p<.001, except *Swipe* and *Tilt*, p=.043).

	Pinch N=447	**Swipe** N=421	**Throw** N=454	**Tilt** N=450
N=1772	6.84 (3.22)	4.22 (1.14)	5.54 (1.89)	4.81 (1.97)

Table 4. Efficiency: means and (standard deviations) of time spent for each technique per target for large targets.

For the large targets, we also identified significant differences (F(2.001, 952.686)=210.596, p<.001). A pairwise comparison showed the same result as the small targets. All techniques were statistically different from each other (for all cases p<.001) in relation to efficiency.

We also extended our analysis and we performed a two-way repeated measures ANOVA in order to examine the combined effect of technique and target size on efficiency. The effect of each technique on efficiency was significant (F(2.230, 2123.045)=309.362, p<.001) and so was the effect of target size (F(2.230, 2123.045)=57.386, p<.001). On the contrary, their interaction was not significant (F(2.230, 2123.045)=1.462, p=.23).

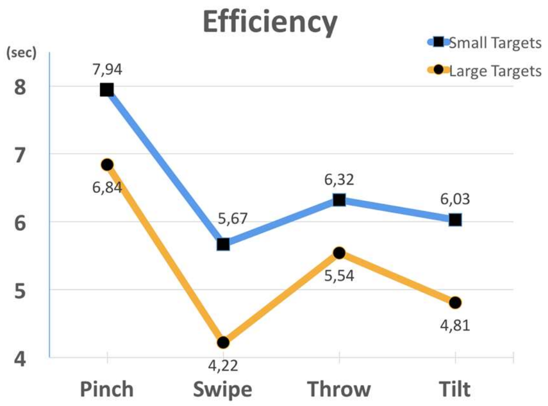

Figure 9. Efficiency: average time spent per target for each technique for small and large targets.

4.3 Error Size

We defined error size as the distance in pixels the cursor had from a target every time a participant missed the target. Means and standard deviations for the distance from target in pixels can be seen in Tables 5 and 6 and combined in Figure 10.

Pinch N=169	**Swipe** N=44	**Throw** N=80	**Tilt** N=197
218.37 (241.29)	51.66 (143.89)	111.06 (210.48)	182.25 (236.57)

Table 5. Error Size: means and (standard deviations) in pixels per technique for small targets.

As a first step we performed two one-way ANOVA's, one for each target size. For the small targets our result show that there is a significant effect of technique on error size (F(3, 483) = 8.457, p<.0001). The pairwise comparisons showed that all techniques were significantly different from each other (p < .001), except *Pinch* and *Tilt* (p = .793), and *Pinch* and *Swipe* (p =.987).

Pinch N=105	**Swipe** N=16	**Throw** N=31	**Tilt** N=104
342.57 (261.99)	68.343 (207.41)	190.75 (264.34)	270.07 (278.96)

Table 6. Error Size: means and (standard deviations) in pixels per technique for large targets

For the large targets, the result was the same as before, F(3, 256) = 6.494, p<.0001). Pairwise comparisons showed that all pairs were significantly different from each other (p<.05) except *Pinch* and *Tilt* (p=.302), *Swipe* and *Throw* (p=.821), and *Throw* and *Tilt* (p=.882)

We continued our analysis and performed a two-way ANOVA in order to examine the combined effect of technique and target size on error size. The effect of each technique on error size was significant (F(3, 735)=18.903, p=.019) and so was the effect of target size (F(1, 735)=12.028, p=.009). On the contrary, their interaction was not significant (F(3, 735)=.774, p=.509).

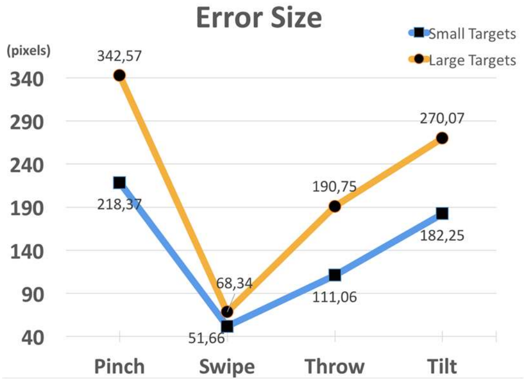

Figure 10. Error Size: average distance in pixels per technique for small and large targets.

5 Discussion

Our results, based on empirical data collected through the reported experiment, make an important contribution to the research area of cross-device interaction techniques by giving new insights in respect to knowing more about how these four different techniques compare in terms of effectiveness, efficiency and error size.

5.1 Effectiveness

The technique with the highest average successful attempts was *Swipe*, both for large and small targets. In contrast, the least successful technique was *Tilt*, especially for the small targets. We believe that *Swipe* was the most successful for two reasons.

Firstly, the users are already experienced with the technique as they are used to swiping during their everyday interactions with technology, and in particular with mobile phones. Secondly, because it was easier for the participants to keep the phone reasonably still while performing *Swipe*. For example, when a mobile phone is being used as a pointing device, then the user needs to keep pointing onto the selected area on the screen, while enacting the data transfer gesture. In this situation, *Tilt* proved to be particularly challenging because when users tilted the phone forward, often this coincided with an unintentional phone movement, causing the cursor to move away from users' intended position on the target shape.

Throw also had relatively high average successful attempts, and we believe that this can be explained by the fact that it required a very natural and easy gesture from the users. The metaphor of throwing something toward the screen matched the gesture required to transfer data from the mobile to the large display. Also, with *Throw*, the pointing hand can be easily held quite still, while the second one is being used to perform the throwing movement. Finally, *Pinch* also performed slightly better than *Tilt* (especially for the small targets). This was most likely due to the fact that even though the pinching hand also acted as the pointer, releasing the data (completing the transfer) required the minimal movement of simply opening the hand while keeping it in place.

5.2 Efficiency

All the techniques were statistically different from each other, both for large and small targets. This result was validated when we examined the combined effect of technique and target size on efficiency. What is interesting though, is that both had a significant effect on efficiency, but there was no significant effect of their interaction. Thus, the larger the target the faster a user is and this effect is similar for all techniques.

As with effectiveness, the best performed technique was *Swipe*. The users were very fast in swiping and hitting the correct shape onto the target on the display. They were also significantly faster when they interacted with large targets. The slowest technique was *Pinch*. We surmise that the reasons for having this result were twofold. Firstly, *Pinch* required a relatively complex gesture from the users to capture and release the data, that is, they had to pinch the shape on the phone, lift their hand up, point it on the screen, and then let go (see Figure 1). Thus this complexity might explain why it took them more time to hit a target. Secondly, often users would spend a considerable amount of time pinching the correct shape on the mobile phone screen once they had identified it. This did not happen for the *Swipe* technique as the users would simply touch on the correct shape on the mobile phone screen and swipe their finger forward in a single move. *Throw* and *Tilt* had significant differences to each other in relation to efficiency, but if we look closer at the averages (Figure 9) their statistical differences are not that meaningful, especially for the small targets (0.29 sec).

5.3 Error Size

Error size is defined as the distance in pixels between the cursor and the target when users failed to hit it. With respect to error size, the results were slightly different than for efficiency.

For small targets, whenever there was a failure, *Pinch* and *Tilt*, and *Pinch* and *Swipe* did not have as statistically significant differences as the rest of the pairs. For large targets, *Pinch* and *Tilt*, *Swipe* and *Throw*, and *Throw* and *Tilt* did not have statistically significant differences. As with efficiency, the effect of the interaction technique and target size on error size was also not significant. Interestingly, our results show that when the target size increases then the size of the error somehow becomes identical for all techniques.

If we observe the actual error size in number of pixels (Figure 10), then it is clear that the technique that produced the smallest error in pixels was *Swipe*. When users failed, then they were relatively close to their target. The technique that produced the largest error was *Pinch*, signifying that when users failed, the cursor was positioned in a quite large distance from the target when the transfer was performed. What is interesting though is that the actual size of the error in pixels was larger when users interacted with large targets. We treat this finding as somehow unexpected and suspect that a possible explanation for this result is that users were more careful to perform the techniques when the targets were small.

5.4 User Feedback

To balance the quantitative data, we would like to briefly discuss some qualitative data that was also collected during the experiment. At the end of the attempts with each interaction technique, we asked participants to complete a questionnaire about their experience, to get additional understanding on how users perceived them. The questionnaire had 6 items, taken from the USE questionnaire [13], focusing on *ease of use* and *ease of learning*:

- Three items for *ease of use*: “This technique is easy to use”, “I can use this technique successfully every time”, and “Using this technique is effortless”, and
- Three items for *ease of learning*: “It is easy to learn to use this technique”, “I quickly became skilful with this technique”, and “I learned to use it quickly”.

Users scored their answers to each item on a 7-point Likert scale. We also used the video of the sessions to transcribe comments made during the experiment, and made notes on their actions, to get a better overall impression of user responses.

The results (Figure 11) showed that *Swipe* was perceived as the easiest to use (5.59) as well as the easiest to learn (6.38). *Pinch* collected the lowest scores both on *ease of use*, as well as *ease of learning*. From these results we can see that, in line with our quantitative findings, users considered *Swipe* as the most useful technique. *Throw* and *Tilt* were rated close to each other on ease of use, while *Pinch* trailed quite significantly behind. However, in the experiment *Throw* outperformed *Tilt* considerably, both in regards to successful attempts and time taken, indicating that perceived ease of use does not guarantee greater effectiveness or efficiency. Similarly, *Pinch* was perceived by users as the most difficult to use, even though it had an average successful attempt rate similar to *Tilt*.

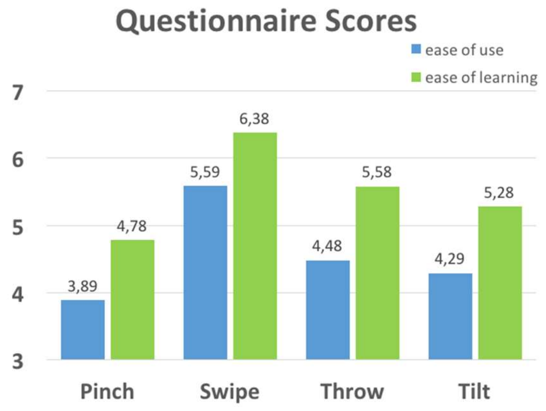

Figure 11. Average scores per technique for ease of use and ease of learning.

From the videos as well as the notes the researchers took while performing the experiment we collected some interesting findings in relation to the learnability of each technique, the distance the users were placed from the screen (2.35m), and the mobile device used in the experiment. In terms of learnability, some mentioned that the *Pinch* technique was hard to learn, and a high number of participants had to be repeatedly told how to perform both the *Pinch* and the *Throw* techniques. This was surprising, as these two were the techniques with the most natural bodily gestures both having analogies with the actions of pinching and throwing real world objects. An explanation of this could be that the system set-up was quite rigid in its recognition of these gestures, while the participants tended to revert to an intuitive behaviour for pinching and throwing, rather than the precise detailed movements given in the tutorial video. This lack of recognition of a gesture led to participant frustration with a technique, when they thought they were performing the technique correctly and nothing was happening on the display. This frustration was commented on for the *Pinch* technique, but also for the *Tilt* technique, which relied on a distinct movement of the mobile phone to register an accelerometer change. Not surprisingly, in the questionnaire, both *Tilt* and *Pinch* were given very low scores on successful and skilful ratings.

The fact the participants had to be stationary at 2.35m from the screen for consistency purposes, also played a role in shaping their experiences. Some users mentioned having trouble reaching all areas of the screen, and almost all users showed signs of trouble, for example, by standing on their toes or stretching their arms as far as possible whenever they had to select targets that were placed at the edges of the screen. One user got so frustrated that she asked for a chair to stand on. Some users also mentioned a problem of the mobile phone obscuring their view of the target on the large display whenever the targets were too high and they had to stretch their arms up. Since they were not allowed to freely position themselves towards the large screen, we expected that such instances could occur for some of our participants. We expect such problems not to exist in real world interactions with large screens as usually users are allowed to freely move around, but nevertheless researchers that are planning similar experiments need to minimize as much as possible such instances as they can have an effect on the way interaction techniques are experienced. Furthermore, such findings also indicate that relative rather than absolute pointing techniques could be more suitable for cross-device interaction between a handheld mobile device and a large display, and should be explored more within HCI research.

In regards to interacting with the mobile device our results are somehow contradictory. Some users complained that the screen of the mobile device was too small when performing a *Pinch*, making it hard to precisely select the correct shape. This may have influenced their low scoring on ease of use, and yet *Pinch* was not the worst in the successful placing of targets on the screen. Other users complained that the mobile device screen was too large when performing a *Swipe*, since it was hard to reach the correct shape with their thumbs while still maintaining precision with the pointer. Some users solved this by using two hands for the gesture. This increased both swipe-ability and pointer precision for them. These observations indicate that the size of the mobile phone may influence the way each interaction technique is experienced, especially in rigid experimental conditions such as ours. Again, we expect such issues to appear less in real world conditions, where the users will be interacting with their own mobile phone, which they are already familiar with. Nevertheless, the size of a mobile phone should be taken into consideration in similar experiments as ours since it is known that it may affect effectiveness, efficiency and perceived usability [20].

Finally, there is also the fun aspect to take into consideration. Many of the users mentioned having fun while performing the *Pinch* technique. They compared it to casting a spell or causing explosions on the screen. Users mentioned that this technique was especially interesting, and enjoyable to use, and yet according to the questionnaire, *Pinch w*as the hardest technique for users to use and to learn.

5.5 Implications for Design

From this research we can learn something about the comparative strengths of these four different cross-device interaction techniques.

In an application where speed is important when sending data from a mobile device to a large display, then *Swipe* is the fastest, and should be considered the interaction technique of choice.

Where accuracy is important, and placement of data from a smartphone to a shared screen needs to be precise, then *Swipe* has the highest success rates in terms of hitting the right place on the display. This holds true irrespective of the size of the target space being aimed at.

Another consideration in favour of *Swipe* is that users perceive it overall as the easiest to use, easiest to learn, most effortless to use, quickest to learn, and the most successful and skilful of the techniques.

An interesting finding for designers to note is that despite poor quantitative performance statistics and low qualitative usefulness ratings and comments, people seemed to have a fascination and fondness for the *Pinch* technique. It should therefore be considered in situations where effectiveness, efficiency and accuracy are not so important, for example for playing games.

Target size matters for all of the studied techniques. This means that in transferring data from a mobile device to a large display, the larger the target – meaning the less precise people have to be in their placement of the pointer – the more effective and efficient they are. This is not a surprising result, however, we also found that when they miss, they actually miss by a greater distance when they are aiming at a bigger target. This was unexpected, and probably indicates that smaller targets encourage users to try more to be precise in the interaction, while with the large ones they believe they are more competent than they actually are in reality. Of course, we believe that it is relevant to investigate more the effect of target size by including more targets in future experiments. This will allow us to have a better understanding on the target size effect.

A final consideration, based on informal observations made during the experiments, is that the laboratory set up and experimental design influences the way that people enact the techniques. During the experiment, we noticed that users would spend relatively little time using their pointing device (mobile phone or hand) to place the cursor in the general vicinity of the target and would spend most of their time, in each attempt, trying to place the cursor exactly on top of the actual target. This indicates that without the need for precision, as imposed by the rigid experimental set up, and given a more realistic task of transferring images or text from personal phones to a shared public display, the comparative outcomes might be different. However, this will have to be investigated through further research.

5.6 Limitations

There are some limitations in our experiment that are mostly related to our implementation of the four techniques. Registering correct gestures with the Kinect was the most problematic. Firstly, when users had to use a hand gesture to complete a data transfer, often the system would register interim hand movements as the completion movement, before they had actually finished performing the interaction. Secondly, the Kinect often had problems determining different arm joints and hands, especially when they moved behind each other or too close to each other. Both issues caused target errors and/or additional time to complete the tasks. Such incidents were noted down during the experiment and were removed from the dataset in order to not affect the results, but the ability of the Kinect technology to recognize different gestures was frustrating for users, and caused more problems with some gestures than others. This may have affected our comparative results, as some techniques could be registered more clearly than others. Furthermore, our collected data are highly dependent on Kinect. Due to our rigid experimental setup, we believe that our findings can be generalised to our types of gesture tracking technologies, but we believe that future research work should look at alternative, and perhaps more robust technologies for recognizing these gestures.

Furthermore, another limitation of our study has to do with the duration of the experiment. Even though our participants took on average about 15 minutes to complete the experiment, they were asked to constantly interact with the screen for those 15 minutes, with the exception of small breaks in-between the interaction techniques. This resulted to some cases where a few participants got tired towards the end of the experiment as they reported that the number of movements we required them to perform was excessive for their standards. Even though we believe our sample size is adequate enough to have valid, generalizable results, we still believe that we should have taken this problem under consideration. Therefore, we strongly recommend to future researchers to purposefully introduce longer breaks in order to allow the participants to recover from any possible muscle fatigue.

6 Conclusion and Future Work

We have presented a study on cross-device interaction techniques focusing on moving data from a handheld mobile device to a large screen display. Specifically, we have compared the use of four different techniques (*Pinch*, *Swipe*, *Throw* and *Tilt*) in conjunction with two different target sizes, to investigate their respective effectiveness, efficiency and error size.

Our findings show that *Swipe* performed best on all measures. *Swipe* was the most effective technique, having the highest number of successful attempts, for both small and large targets. At the same time *Swipe* was also the most efficient technique, being the fastest one to use, for both small and large targets. The *Swipe* technique was also the most accurate one when looking at the size of the errors encountered. Again, this was the case for both small and large targets.

From this we conclude that to design a cross-device interaction technique used to send data from a handheld mobile device to a large display, a swiping technique, like the one presented here, should be a first consideration. In terms of effectiveness, the *Tilt* technique performed the worst, and especially with small targets. In terms of efficiency and error size, the *Pinch* technique was the slowest and also the most imprecise, but at the same time it is reported as the most enjoyable one from many participants. We also found that target size mattered considerably for all techniques, confirming previous research, but surprisingly when people missed a target, they missed larger targets to a greater degree than smaller ones.

Our experiment has investigated four specific cross-device techniques, and the specific measures of effectiveness, efficiency and error size, in a laboratory setting. Future research should expand this with additional techniques and measures, such as usefulness for concrete tasks, and perceived user experience. We would also like to see comparative studies carried out in real-world settings in order to see if the size of the mobile device and the distance from the screen has a significant effect. Finally, we recommend to other researchers that will carry out similar experiments as ours not only to consider the fatigue effect onto their experiment, but to thoroughly research it, as it can be extremely relevant for specific types of tasks. For example, it would be interesting to know which of our four interactions techniques is less tiring for the users, when they are required to repeatedly interact with a large screen.

Overall, our findings can be used to inform the design of applications with cross-device interaction through knowledge about their relative strengths and weaknesses in terms of effectiveness, efficiency and error size.

Acknowledgements

We would like to thank all our participants for their valuable contribution to our study.

References

1 Benko, H. and Wilson, A.D., Multi-point interactions with immersive omnidirectional visualizations in a dome. in ITS '10: ACM International Conference on Interactive Tabletops and Surfaces, (2010), 19-28.

2 Boring, S., Altendorfer, M., Broll, G., Hilliges, O. and Butz, A., Shoot & copy: phonecam-based information transfer from public displays onto mobile phones. in Mobility '07: 4th international conference on mobile technology, applications, and systems and the 1st international symposium on Computer human interaction in mobile technology, (2007), 24-31.

3 Boring, S., Jurmu, M., and Butz, A., Scroll, tilt or move it: using mobile phones to continuously control pointers on large public displays. in OZCHI '09: 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group, (2009), 161-168.

4 Bragdon, A., DeLine, R., Hinckley, R. and Morris, M.R., Code space: touch + air gesture hybrid interactions for supporting developer meetings. in ITS '11: ACM International Conference on Interactive Tabletops and Surfaces, (2011), 212-221.

5 Chen, X., Grossman, T., Wigdor, D.J. and Fitzmaurice, G., Duet: exploring joint interactions on a smart phone and a smart watch. in CHI '14: SIGCHI Conference on Human Factors in Computing Systems, (2014), 159-168.

6 Dachselt, R. and Buchholz, R., Natural throw and tilt interaction between mobile phones and distant displays. in CHI '09 EA: CHI 2009 Extended Abstracts on Human Factors in Computing Systems, (2009), 3253-3258.

7 Hamilton, P. and Wigdor, D.J., Conductor: enabling and understanding cross-device interaction. in CHI '14: SIGCHI Conference on Human Factors in Computing Systems, (2014), 2773-2782.

8 Hincapié-Ramos, J.D., Guo, X., Moghadasian, P., Irani, P., Consumed endurance: a metric to quatify arm fatique of mid-air interactions. in CHI '14: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, (2014), 1063-1072.

9 Hoaglin, D.C., Iglewicz, B. and Tukey, J.W. Performance of some resistant rules for outlier labeling. Journal of the American Statistical Association, 81(396), 1986. 991-999.

10 Ikematsu, K. and Siio, I., Memory Stones: An Intuitive Information Transfer Technique between Multi-touch Computers. in HotMobile '15: 16th International Workshop on Mobile Computing Systems and Applications, (2015), 3-8.

11 Jokela, T., Ojala, J. and Olsson, T., A Diary Study on Combining Multiple Information Devices in Everyday Activities and Tasks. in CHI '15: 33rd Annual ACM Conference on Human Factors in Computing Systems, (2015), 3903-3912.

12 Lucero, A., Holopainen, J. and Jokela, T., MobiComics: collaborative use of mobile phones and large displays for public expression. in MobileHCI '12: 14th international conference on Human-computer interaction with mobile devices and services, (2012), 383-392.

13 Lund, A.M., Measuring Usability with the USE Questionnaire. STC Usability SIG Newsletter, (2001).

14 Marquardt, N., Hinckley, K. and Greenberg, S., Cross-device interaction via micro-mobility and f-formations. in UIST '12: 25th annual ACM symposium on User interface software and technology, (2012), 13-22.

15 Myers, B.A., Using handhelds and PCs together. Communications of the ACM, 44 (11), 2001. 34-41.

16 Nielsen, H.S., Olsen, M.P., Skov, M.B. and Kjeldskov, J., JuxtaPinch: exploring multi-device interaction in collocated photo sharing. in MobileHCI '14: 16th international conference on Human-computer interaction with mobile devices & services, (2014), 183-192.

17 Rekimoto, J., Pick-and-drop: a direct manipulation technique for multiple computer environments. in UIST '97: 10th annual ACM symposium on User interface software and technology, (1997), 31-39.

18 Rekimoto, J., Multiple-Computer User Interfaces: A cooperative environment consisting of multiple digital devices. in CoBuild '98: First International Workshop on Cooperative Buildings, Integrating Information, Organization, and Architecture, (1998), 33-40.

19 Rädle, R., Jetter, H., Schreiner, M., Lu, Z., Reiterer, H. and Rogers, Y., Spatially-aware or Spatially-agnostic?: Elicitation and Evaluation of User-Defined Cross-Device Interactions. in CHI '15: 33rd Annual ACM Conference on Human Factors in Computing Systems, (2015), 3913-3922.

20 Raptis, D., Tselios, N., Kjeldskov, J. and Skov. M.B., Does size matter?: investigating the impact of mobile phone screen size on users' perceived usability, effectiveness and efficiency. in MobileHCI '13: Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services, (2013), 127-136.

21 Scheible, J., Ojala, T. and Coulton, P., MobiToss: a novel gesture based interface for creating and sharing mobile multimedia art on large public displays. in MM '08: 16th ACM international conference on Multimedia, (2008), 957-960.

22 Schmidt, D., Chehimi, F., Rukzio, E. and Gellersen, H., PhoneTouch: a technique for direct phone interaction on surfaces. in UIST '10: 23rd annual ACM symposium on User interface software and technology, (2010), 13-16.

23 Schmidt, D., Seifert, J., Rukzio, E. and Gellersen, H., A cross-device interaction style for mobiles and surfaces. in DIS '12: Designing Interactive Systems Conference, (2012), 318-327.

24 Skov, M.B., Kjeldskov, J., Paay, J., Jensen, H.P. and Olsen, M.P., Investigating Cross-Device Interaction Techniques: A Case of Card Playing on Handhelds and Tablets. in OzCHI '15: Annual Meeting of the Australian Special Interest Group for Computer Human Interaction, (2015), 446-454.

25 Sørensen, H., Raptis, D., Kjeldskov, J. and Skov, M.B., The 4C framework: principles of interaction in digital ecosystems. in UbiComp '14: 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing, (2014), 87-97.

26 Walter, R., Bailly, G., Valkanova, N. and Müller, J., Cuenesics: Using Mid-air Gestures to Select Items on Interactive Public Displays. in MobileHCI '14: 16th international conference on Human-computer interaction with mobile devices & services, (2014), 299–308.