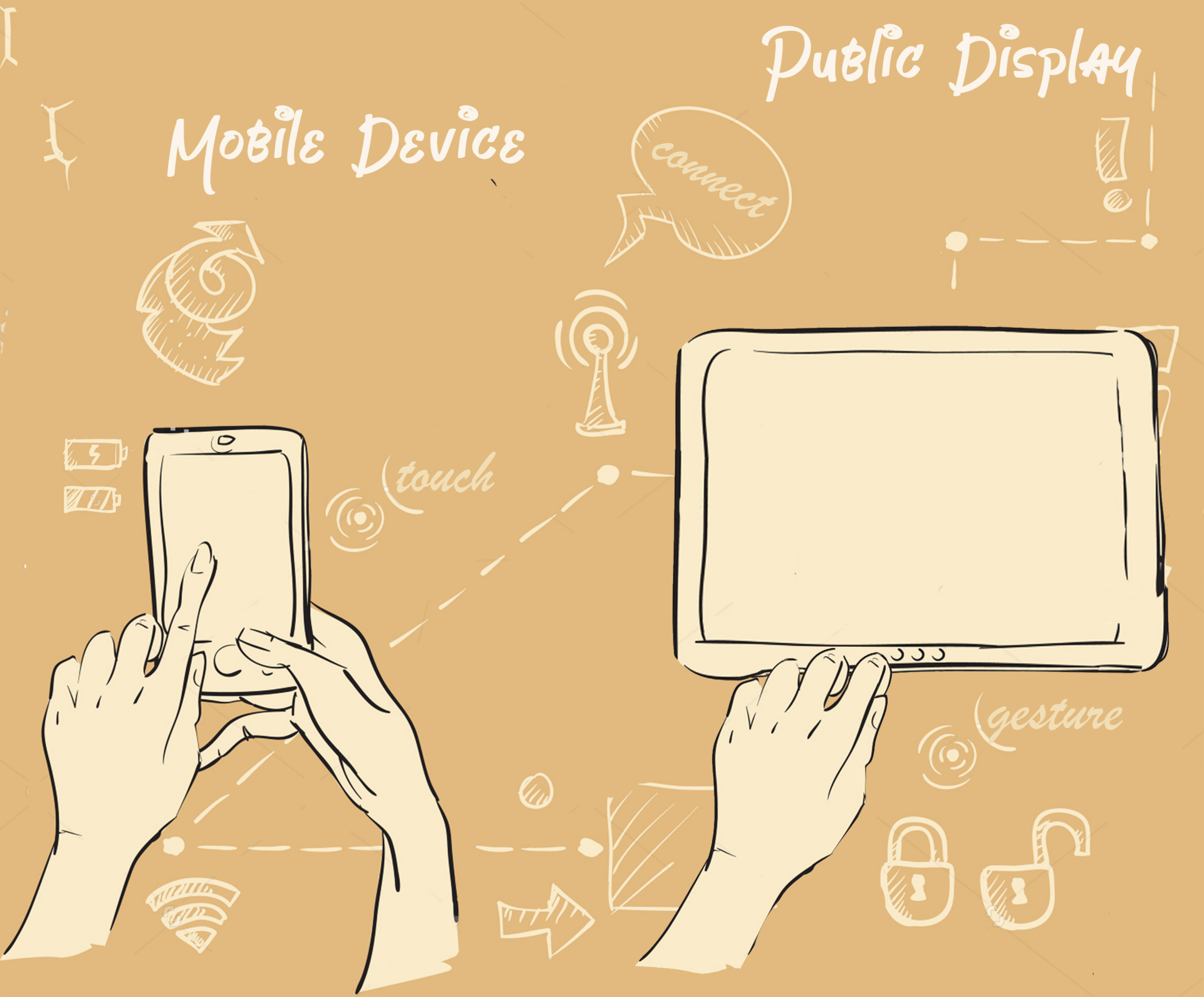
UNDERSTANDING THE USE OF INTERACTION TECHNIQUES

CLARITY, COMFORT, CONTEXT, AND CONTROL



IVAN S. PENCHEV
ELIAS RINGHAUGE

AALBORG UNIVERSITY
STUDENT REPORT

UNDERSTANDING THE USE OF INTERACTION TECHNIQUES
CLARITY, COMFORT, CONTEXT, AND CONTROL

SUMMARY:

A major challenge when creating large interactive displays in a public setting is creating and defining the ways for users to interact with them. This thesis presents the outcomes of user evaluation sessions designed to test a series of different cross-device natural user interaction techniques for people to interact with large displays in public setting. It is an initial step towards the entire perspective, the essence of being able to establish a natural means for immersive interactions with public displays. This thesis proposes four cross-device natural user interaction techniques for transferring data from a mobile device to a public displays in a realistic task scenario.

This thesis reports a quantitative and qualitative analysis on user responses to the proposed interaction techniques, leading to more in-depth understanding of the nature of cross-device interaction with a natural modality between people and large public display. Our contribution to HCI is 4 aspects, Clarity, Comfort, Context, and Control, that should be considered when designing and developing interaction techniques in order to support information transfer for interactive public displays.

Authors:

Ivan Penchev

Elias Ringhauge

Supervisor:

Assoc. Prof. Jeni Paay

Group:

IS106f16

Project period:

02.02.2016 – 06.06.2016

Number of pages:

66

AALBORG UNIVERSITY

Selma Lagerlöfsvej 300
Aalborg, Denmark

www.aau.dk

Phone 99 40 72 28

"Interaction design isn't only about fixing problems; it's also about facilitating interactions between people in richer, deeper, better ways – that is, finding new ways to better connect human beings to one another, and by doing so, make the world a better place to live"

Dan Saffer, *Designing for Interaction*

To our parents:

Juliana Yordanova
Svilen Penchev

Dorte Ringhauge
Lise Jensen
Kaj Jensen

For your love and support!

To our dearest:

Denica Dineva

Lotte Sørensen

For aiding us in our study by
proofreading, transcribing videos, and caring!

TABLE OF CONTENTS

2016
AALBORG UNIVERSITY

Preface

This report documents the work of a Master's thesis in Human Computer Interaction, Department of Computer Science at Aalborg University. It is a continuation of a 9th semester project. This report covers the work done between 02/02/2016 to 05/06/2016.

Our study work resulted in this report. The report consists of three parts and is structured as follows: First an overall introduction will explain the background and motivation for this project and its stages. Then the main content is enclosed in a ACM CHI SIG formatted research paper followed by a combined discussion of the two projects. Lastly, the 9th semester paper can be found in the appendix alongside pictures, participant information, and statistical data collected from the questionnaires.

The theme of our project is cross-device interaction techniques between mobile devices and public displays. More specifically understanding the user's responses when using different interaction techniques in a public setting.

A CD containing the source code of the developed systems is supplied together with this report. The CD also contains a digital copy of this report as well as appendices with e.g. interview transcriptions. Additionally, it contains all logged numbers used for the quantitative analysis, enclosed as database exports.

Finally, we would like to thank Jeni Paay, Associate Professor at Aalborg University, for providing ongoing feedback and guidance throughout the project as well as for her excellent support during this past year's ups and downs.

Introduction

As public displays become ubiquitous in the form of projectors, TVs, and kiosk, novel types of interaction applications emerge. In particular, we foresee applications that support user-generated content to rapidly gain importance, since they provide a benefit for the user, for example in the form of digital bulletin boards, discussion platform that foster public engagement, and applications that allow for self-expression. However, at the same time, the novelty of such applications infer several challenges: first, they need to provide suitable means for the passerby to contribute content to the application and second, mechanisms need to be employed that provide sufficient control for the display owner with regard to content moderation;

The foundation of our motivation for our Master's thesis can be found in our previous semester's work. Which began with us being interested in Natural User Interfaces(NUI) in public spaces, with regards to the interaction between the users and the system, as well as with each other. For example, a project that sparked our interest was one that was being conducted at Aalborg University. They were creating a large touch wall where multiple users could come up and create music together by touching the wall in different places. After a few iterations of our research idea we decided that instead of pursuing artistic installation type of work, we would focus on interaction side.

This report presents a study conducted on our 10th semester, at Aalborg University, as a continuation of our 9th semester project. The focus of our Master's thesis is on suitable means for passerby to contribute content to a public display. More specifically we examine cross-device natural user interaction as such potential means, this is done by investigating the user's responses to the different interaction techniques, both quantitatively and qualitatively.

Since cross-device natural user interaction is a relatively new research field, this gives us the opportunity to provide foundational research for use with public displays in applications such as bulletin boards and interactive advertising boards.

RESEARCH PAPER

Clarity, Comfort, Context, and Control. Understanding the use of Interaction Techniques for Cross-device Natural User Interaction in a Public Setting

Abstract

A major challenge when creating large interactive displays in a public setting is creating and defining the ways for users to interact with them. This paper presents the outcomes of user evaluation sessions designed to test a series of different cross-device natural user interaction techniques for people to interact with large displays in public setting. It is an initial step towards the entire perspective, the essence of being able to establish a natural means for immersive interactions with public displays. This paper proposes four cross-device natural user interaction techniques for transferring data from a mobile device to a public display in a realistic task scenario.

This paper reports a quantitative and qualitative analysis on proposed interaction techniques, leading to more in-depth understanding of the nature of cross-device interaction with a natural modality between people and large public display. Our contribution to HCI is 4 aspects, Clarity, Comfort, Context, and Control, that should be considered when designing and developing interaction techniques in order to support information transfer for interactive public displays.

2016

AALBORG UNIVERSITY

Clarity, Comfort, Context, and Control. Understanding the use of Interaction Techniques for Cross-device Natural User Interaction in a Public Setting

Ivan S. Penchev
Aalborg University
ipench14@student.aau.dk

Elias Ringhauge
Aalborg University
eringh10@student.aau.dk

ABSTRACT

A major challenge when creating large interactive displays in a public setting is creating and defining the ways for users to interact with them. This paper presents the outcomes of user evaluation sessions designed to test a series of different cross-device natural user interaction techniques for people to interact with large displays in public setting. It is an initial step towards the entire perspective, the essence of being able to establish a natural means for immersive interactions with public displays. This paper proposes four cross-device natural user interaction techniques for transferring data from a mobile device to a public display in a realistic task scenario.

This paper reports a quantitative and qualitative analysis on proposed interaction techniques, leading to more in-depth understanding of the nature of cross-device interaction with a natural modality between people and large public display. Our contribution to HCI is 4 aspects, Clarity, Comfort, Context, and Control, that should be considered when designing and developing interaction techniques in order to support information transfer for interactive public displays.

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous; See http://acm.org/about/class/1998/ for the full list of ACM classifiers. This section is required.

Author Keywords

Interaction Techniques; Cross-Device Interaction; Natural User Interaction; Kinect; Mid-air Gestures; Public Space; Public Displays; Large Displays

INTRODUCTION

Historically, most software applications are found on devices that are used in a personal context like a laptop, a personal computer, a mobile phone and so forth. However, as our, technical capabilities grew, so did our dreams, and we started imagining new places and uses for the current and future technology, in order to make our lives easier, more productive, more entertaining or even just different.

An example of this can be seen in the replacement of the traditional signs with digital displays. This is due to rapid progression in the technology for displays and projections, which led to lowering of the marginal price, thus *"displays have moved out of research laboratories into public spaces "* [19]. Now we have displays that are deployed in malls, shop windows and even in urban environments.We can see the transformation from traditional to digital in inner city areas, airports, train stations, and stadiums. With this transformation we can enable new forms of multimedia presentation, new user experiences, and interactivity. However, it is an open question whether we use this opportunity, as the vast majority of digital displays remain non interactive.

This trend of broadening the domain of multimedia to move beyond the private into the public space is promising, as it not only gives the possibility for technology to become more relevant for our lives, but it also presents the possibility of more immersive experiences. Presented with this opportunity, research has been done in the area of digital displays for use in a public context, marked with the term public displays. Some of the researchers gravitated around a specific system for a specific task, while others have worked on an abstract level. For example, Bringhul et al.[35] identified and tried to tackle the problem about how to entice people to participate in public displays, and Muller et al.[27] analysed the public design space contributing to the understanding of mental and interactional models. Cheung et al. [10] go further to analyse the barriers you face when interacting with public displays with a mobile device. What all these authors have in common, beside analysing the challenges for public displays, is that they agree that a key component of a public display is the interaction, marked with the term public interactive display.

There are different ways to interact with a public display. One way a person can do so is via a second device (cross-device interaction) or via gaze, touch or some other natural modality (natural user interactions); however, recent research in HCI has focused on combining these two fields in what could be called cross-device natural user interaction. An example can be seen within the work of researchers using spatial information [25, 26], others using touch [38], or combined touch with air gestures [8].

Based on the work of Lauridsen et al. [36] who empirically compared four cross-device natural user interaction techniques in a laboratory setting, we set out to investigate how interaction techniques tested in a laboratory, actually perform in a public situation. We ask the question: *"How can we understand people's responses to pushing information from a mobile device to a public display?"*.

In this paper we review the research that has been done within interaction with large public displays. This has been done to map the current understanding and practices, thereby helping future researchers within this field to gain a perspective on state of the art as well as identifying possible untapped opportunities for future work. This paper is structured in the following way: firstly, we look at related work in the domain of public displays to gain an insight of what has been done in terms of research. We then describe the methods used for our study, including our research approach, followed by a description of the study participants and the data analysed. We then give a description of the findings gathered from the user study sessions, including a list of areas of interest. In the concluding part of the paper, we discuss findings, future work and limitations of the study.

RELATED WORK

The current section will present related work in relation to large public displays, their domain in the public space, and ways of interacting with emphasis on cross-device natural user interaction.

Large Public Displays

One of earliest large public displays was Hole-in-Space, it was a window type, which enabled people from New York to see, hear and speak with head-to-toe, life sized images of people from the opposite side of the US, in Los Angeles [23]. Though the start was given as an artistic experiment, various research groups later scientifically explored the concept of interactions between people located at remote places over public displays with audio-video capabilities [13, 14, 4].

In the early nineties shared interactive display surfaces was an emerging research area, with early examples of shared drawing surfaces [5, 20]. Combining this with the new paradigm called "ubiquitous computing", Weiser showed examples of how display devices of different sizes could be embedded into a working environment to solve different tasks. He contemplated that these displays could be networked and exist in the periphery of the user, leaving him with the choice of interaction [45]. These two ideas resulted in richer examples of situated displays [41, 39], which gave us a better understanding of the role that public displays have in conversation and also their influence on group dynamics.

In the late nineties we witnessed the emergence of ambient display systems. Systems that could project water ripples on the ceiling of a room to denote different activities [21], or which used air bubbles rising vertically from a water tube to render small black and white images [18]. Also wearable displays, for example a LED display on a T-shirt became increasingly popular as an area of research, leading to displays being designed and studied in a more compact form[7, 33, 11].

In the beginning of the new millennium, technological progress of display technologies gave us the possibilities to have more affordable and less cumbersome displays and projection techniques. This research expanded to identify the social consequences of design and placement of displays. An example is Kaharalios et al.'s work [22], which even though it involved connecting two remote locations over audio-video link, had a different concept in mind. They blended the videos and converted them to make users look like graffiti, which was projected on the wall. This prompted users to move closer to the display and acted as a social catalyst to motivate users to engage in " conversation " with each other. This led to a new focus of research encouraging social interactions using situated displays. Greenberg et al.[15] combined multiple personal desktops and a large semi-public display to improve awareness among colleagues connected electronically. It consisted of a desktop client that allowed them to post multimedia content on a real time collaborative surface. A collage with randomly placed content appeared both on a large public display in a common area and on personal workstations. This acted as a start for social interaction between two people, for example, clicking on the live video stream would start a chat with that person.

However, despite this research, public displays used to support community and social activities revealed a major problem; the lack of willingness by users to participate[35].

The Public Space as a challenge

To address the participation problem, various researchers have presented different models on audience behaviour and interaction with public displays.

Streitz et al.[30] presented the three-zone model which defined three zones of interaction; ambient zone, notification zone and cell interaction zone based on the distance between the user and the display. This model lacks support for multiple users and assumes that a user in the cell interaction zone intends to interact with the system. Brignull and Rogers [35] presented an interaction model based on how people become aware of the existence of a display installation. They identified three activity spaces around a display installation: space of peripheral awareness, space of focal awareness, and space of direct interaction. They further identified that the transition zones between these spaces represent a key barrier to interacting with the display.

Researchers became very interested in using mobile devices to interact with public displays, when mobile devices became increasingly prevalent in the late 2000's early 2010. The advantages put forward for mobile phones were three fold. Firstly, the sensors in mobile devices made it possible to detect people around a display. Secondly, physical or touch buttons, microphones, inertial sensors and cameras made it possible to make it an interaction device. Lastly, mobile phones acted as a storage medium to transfer information from and to public displays.

Interacting with Public Displays

One of the fundamental aspects of public displays that make them useful is when they include interactivity. This aspect encourages participation and enables users to explore the content. One way of doing that is include interaction with other devices, such as mobile phones, making it cross-device interaction.

Cross-device studies with public displays are related to using mobile phones as interaction devices for controlling public displays, either with or without feedback on the mobile devices. An example of the former is the multiplayer Breakout game by Cheung et al. [10] that is played on a large public display. Multiple players can join the game by scanning the QR code located in front of the display using their mobile phones. This web-based client application allowed each player to control their paddle by tilting their phone. The mobile device provides feedback regarding the game and connection status (error, connected, disconnected). Vepsalainen et al. [43] presented a similar work wherein mobile phones were used as a gamepad for a controlling the game running on a large display, with feedback given on the mobile phones. In contrast, no feedback was provided on the users' mobile phones in Scroll, Tilt or Move, where Boring et al.[6] presented various techniques for controlling the pointer on a large display.

People use different modalities (speech, eye gaze, touch, gestures, facial expressions, body postures and others) while interacting with each other. Researchers have extended and explored the use of similar modalities to interact with public displays.

Speech can be recognized using microphone arrays near the displays to issue digital commands. For example, Ward provided air travel information based on the spontaneous speech commands given to the system [44]. However, using speech alone as an input method is error-prone especially when the content on the display is dynamic [32], therefore the majority of studies used a combination of speech and gestures, as this combination was natural and efficient to cope with the visual complexity of the display [2].

The term "gesture" is quite loosely defined in the field of HCI and it depends on the context of interaction. In our context, we refer to hand gestures and gestures using stylus-like devices or other external devices (such as the Wii or Kinect controller). The success of Kinect controller started a widespread interest in gesture interfaces within the research community. Users can walk into the vicinity and start interacting even before they realize it, and explore the system gradually [29]. While gestural interaction is desirable for interaction with public displays, they pose some challenges as well. A major challenge with gesture recognition is how to differentiate naturally occurring gestures, and gestures intended to interact with the system [46]. Moreover, the use of gestures could be embarrassing or disruptive to the user in a public environment [34].

Touch, another natural modality we can use for interaction is accurate and provides a natural tactile feedback for the interaction. Although the rapid growth of touchscreen devices has increased the affordances for touch, not all displays currently deployed in public places support touch. Therefore, informing the user of which surfaces are touchscreens, and which are not, is crucial for its use. *MirrorTouch* supported both mid-air gestures and touch to increase the usage of the display[28]. However, using touch is not always possible because of varying display locations and/or sizes.

As different interactions have different pros and cons an obvious question is: *would a combination of interaction types create a synergy that would maximize the pluses and minimizes the minuses for information transfer to a public display?*

Cross-device natural user interactions for data transfer in Public Displays

Determining the recipient and sender devices for an information transfer in a public setting presents more challenges in contrast to controlled environments like offices, homes and workstations. One of the earlier attempts in this domain used a grid of QR codes to identify the desired content [3]. The authors developed two complementary interaction techniques using phone cam: Sweep and Point+Shoot. In the sweep technique, the phone camera acts like an optical mouse. Successive images are compared to determine the direction and displacement of the phone. In Point+Shoot technique, the user aims the phone camera (point) at the desired content on the display and presses the joystick on their phone (shoot) to retrieve that content to their personal device via QR code.

A different idea, presented by Hardy and Rukzio [16] allowed the user to touch the desired part of the public display with their mobile phone in order to perform interactions. This interaction technique was implemented using a grid of NFC tags, which would help the system dynamically identify the part of the application UI that the user was interacting with at that point. Since NFC tags have an id associated with them, it is trivial to identify the sender and recipient in the event of any information transfer between the public display and the personal device.

Researchers have also explored the use of other sensing technologies to detect similar touch events. For example, PhoneTouch relied on accelerometer data generated whilst touching the display surface to generate touch events when the phone touches the display surface [37]. Users could share content with the display surface, simply by selecting the content on their phone, followed by touching the display surface with their phone. However, the major drawback of this approach becomes visible when the size of display increases beyond the physical reach of a person.

To overcome this challenge Bragdon et al.[8] propose using a system that combines touch + air gesture hybrid interactions. They aim to design, implement and test a system that allows a group of users to interact using air gestures and touch gestures. The purpose is to increase control, support democratic access, and share items across multiple personal devices such as smartphones and laptops where the *"primary design goal is fluid, democratic sharing of content on a common display."* [8]. This method enables access, control and sharing of information through several different devices such as multi-touch screen, mobile touch devices, and Microsoft Kinect sensors.

Alt et al. [1] created an application that allows passersby (e.g., customers) to exchange content with a public display, similar to traditional public notice areas, such as bulletin boards. They present Dignifes, a digital public notice area built to investigate and compare possible interaction techniques. Based on a lab study they show that using direct touch at the display as well as using the mobile phone as a complementing interaction technology are most suitable. However, their findings indicate that while users had a similar performance using mobile phones and displays, they actually prefer mobile phones to create content in a privacy-preserving way / on-the-go.

The relish of possible further extension of the currently used public displays with the new interaction techniques is fascinating. This motivates a study conducting a field evaluation.

EXPERIMENT DESIGN

In this section we establish all the information needed for replication of our experiment. We begin by describing the building blocks and then proceed with the methods used to achieve our data analysis.

Interaction techniques

In our research we use four cross-device natural user interaction techniques - *Grab*, *Swipe*, *Throw*, *Tilt*. They have been chosen, after an extensive review, from the existing literature based on predetermine criteria. The most clear differential between techniques is the number of hands required to perform each of them. Two hands - *Grab* and *Throw* versus one hand - *Throw* and *Tilt*.

A full and meticulous description of the selection criteria, each technique as well as graphical representation of the technique's motion can be found in our previous work [36]. However, for continuity purposes we will shortly describe each technique. *Grab* is a complex two handed interaction technique, which is used in order to simulate the feeling of picking up an object from the phone and placing it at a desired location on a public display. *Swipe* is more simplistic technique which is one handed, the user points at the desired location with the phone stretching their arm, and then swipes their finger on the phone screen to transmit the information. *Throw* is performed as follows: The user points at a desired location on a public display with one hand, the other hand is used for the phone to transfer information. The user must then select the data and perform a swinging motion towards the public display to submit it. *Tilt* is performed as follows: The user is pointing at a desired location with one hand which also holds the phone, if they wish to transfer the data they must tilt the phone forward.

Choosing public space location

For our experiment we chose AAU cafeteria, located in department of Computer Science, the map of which is presented in Figure 1

Using Fischer's model [12] , which reflects on technology interventions by analysing their spatial configuration in relation to the structuring of interaction, we narrowed three candidate-spots for our experiment setup in the public space, those spots are indicated in Figure 1, with A, B, C notations.

Even though the process included Fisher's model, and we graded all spots by the model's criteria, for the final selection we had placed an emphasis on activities relating to small spaces, which best suited our current case.

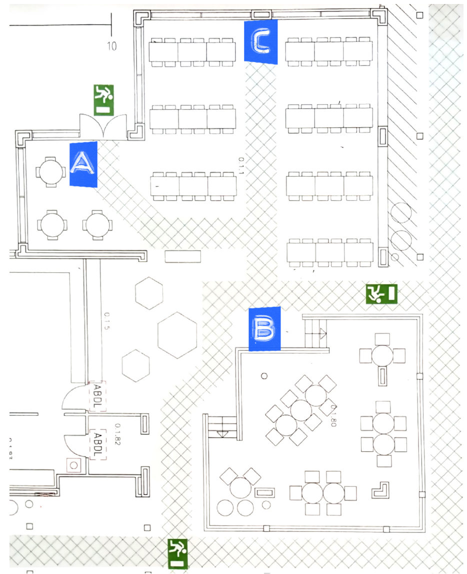


Figure 1. A map of the cafeteria we selected for our experiment

Results achieved by Spots A, B, C

	Spot A	Spot B	Spot C
Display Space	Small	Medium	Large
Interaction Space	Small	Medium	Small
Potential Interaction Spaces	Medium	Medium	Medium
Gap Spaces	Medium	Medium	Medium
Social Interaction Spaces	Small	Large	Medium
Comfort Spaces	Small	Small	Large
Activation Spaces	Small	Medium	Large

Table 1. Evaluation of possible public place spots using Fischer's Model

Table 1 shows the complete evaluation based on Fisher's Model. Our focus includes: Display Space is the area where our setup can be seen, it is something we wish to maximize. The Interaction space is the area where a potential interaction could occur, due to our system being single user, we would like to minimize this. The Comfort space is where psychological ease can be found, like walls, pillars and so on, we aim to maximize it. Finally Activation space is where a person can see the setup, triggering curiosity about it, but interaction is not possible we would like increased size of it. One could make an argument that all of the criteria are important, for example potential interaction space would be beneficial for converting an onlooker into a participant, we considered this; however, it is important to tailor the method/model for our instance.

Based on out evaluation of available spaces, Spot A is dropped due to smaller scale of the required zoning we would like to have, more precisely Displays space, Interaction space, Comfort space, and Activation space. Spot B and C are similar and even though Spot B offers improved lighting conditions, ultimately our selection finishes with Spot C, due to the fact that Spot C could be seen also from the second floor of the canteen, which Spot B is directly underneath and remains invisible from the upper floor. The second floor as a factor is beneficial because it increases the Comfort space.

Hardware and Software

Our aim was to study cross-device natural user interaction techniques in public settings and their use for a realistic task. To do this we extend on the system we built in our previous work [36]. That system hardware utilization is a Microsoft Kinect, a 65' inch Panasonic television with a 1920×1080 resolution and a Samsung Galaxy SII in order to create the experimental setup. Our current system makes some hardware changes. First we switch the mobile device to LG Optimus G E975, also we switch to Kinect for Windows v2,. The latter change is made due to the better capabilities of joint detection, this in return increases the robustness of the system, which we need for our field study. The server side of the software is written in C#, while the mobile part is in Java. Most of the participants interacted at a distance of 2.35 m. from the setup, this was chosen as it is an optimal operating distance for the Kinect, but it was not fixed. The height of the Kinect in regards to the floor was chosen in order to get the optimal position for a person who is 185 cm tall.

An overview of the experiment setup can be seen in Figure 2, and overview of the software interface can be seen in Figure 3.

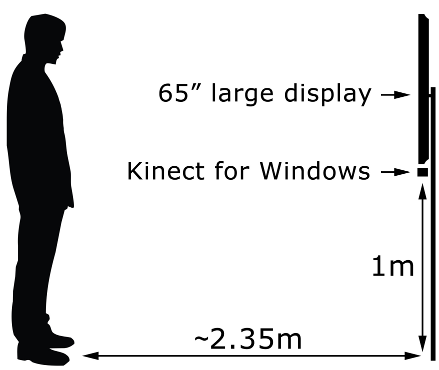
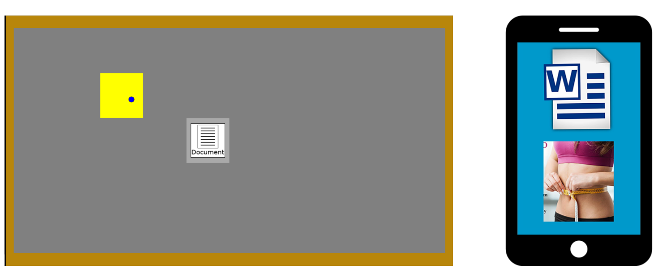


Figure 2. An overview of experiment setup



(a) (b)

Figure 3. An overview of software interface (a) The large display screen of the application, with the blue dot in the highlighted square as the aiming point and the document object as the target. (b) The phone screen showing a document and image object.

Test conditions

The experiment was conducted as a within-subject research, with the four different interaction techniques and two different send objects - image and document.

The within-subject research was chosen because we wanted to minimize the number of subjects needed in order to get significant results. We chose two objects so the users would need to orient themselves with the phone, but not spending too much time searching for the correct object. This is done so the users will not be performing the interaction techniques without paying any attention to the phone at all.

Tasks and Experimental procedure

The task given to a participant consists of pushing objects i.e., images and documents, from the phone to the large display, using four cross-device natural user interaction techniques.

For each technique there are 9 targets and 3 practice targets at the beginning of each technique test. Each participant must complete four technique tests in order to complete the whole experiment. The order in which the techniques were presented was randomized to minimize the learning effect of each technique. We also collected quantitative data from the system such as: did the person hit the target, how long did the user took to perform a technique and how far way from the target (in pixels) each user was. However, we ended up not using it, since our initial research idea shifted to a more human-centred approach with a focus on the people's perceptions of different interaction techniques.

As we aim to gather data about the persons perception of the techniques as well as their performance anxiety of using them in public settings, we concluded that observers need to be present during each test. For this reason we decided to recruit groups of minimum 3 people, in which all the group members would rotate and become participant, then observer. The procedure starts with a group member becoming a participant. He would then use all 4 techniques to transfer content to the large screen. After each technique, he would have to fill in a questionnaire for that technique. During this time the observers must observe and share their opinions of the performance of the techniques in their own questionnaire. After a participant is finished, he returns to the group pool to become an onlooker, and a new participant is selected from the pool. After all the people from the pool have had the experience of being a participant and onlooker they fill in a questionnaire about the whole experience.

	Participant	Observer
Participant Questionnaire (12 Rating questions)	4	-
Observer Questionnaire (4 Open questions)	-	1
Experience Questionnaire (5 Open & 1 rating question)	1	-

Table 2. Types of questionnaires, the questions they include and number of times a person must fill in each of them as well as during which phase (Participant or Observer)

There are 3 different questionnaires, where one is filled four times once for each technique, and the total number of questions are 13 Likert scale and 9 open ones. Table 2 shows how the questions are divided between questionnaires, as well as during which phase, participant or observer, each person should fill them in.

For example a group consisting of 3 members would produce 12 x Participants questionnaires (3 members x (4 questionnaires for the different techniques x (12 questions))), 3 x Observers questionnaires (3 members x 1 questionnaires x (4 questions)), and 3 x Experience questionnaire (3 members x 1 questionnaires x (5 questions)) resulting in 18 data sets each of which contains 22 data points. Table 3 exemplifies this.

Process for a group of three members

Time	Current	Pool
t_1	$P1_{PQ1,2,3,4}$	$O2_{OQ1}, O3_{OQ1}$
t_2	$P2_{PQ1,2,3,4}$	$O1_{OQ1}, O3_{OQ1}$
t_3	$P3_{PQ1,2,3,4}$	$O1_{OQ1}, O2_{OQ1}$
t_4		$O1_{EQ1}, O2_{EQ1}, O3_{EQ1}$

Table 3. Number of questionnaires filled by participants at different times. Parcitipant questionnaire (PQ), Observer questionnaire (OQ), Experience questionnaire (EQ)

Participants and Data Analysis

In total, 24 people took part in our experiment. The participants where between 21-55 years old (M: 26.5, SD: 7.3) and were between 1.63 and 1.95 meters tall (M: 1.77, SD: 0.09). 95.83% of users were right handed, 79.17% were male, and 95.83% of them were smartphone users. Of those who owned smartphones, they had owned one for 2-12 years (M: 6.29, SD: 2.7). They were recruited through a mixture of our social network and recruitment posters around the campus.

For the analysis we recorded each session on audio and video to be reviewed and transcribed. In the data analysis we applied a selection of coding techniques from Grounded theory [40, 24]. Open Coding was used during analysis to mark significant and important points from the participants. By applying an iterative thematic analysis approach [9], using an affinity diagram, lists of themes were discovered. In some cases, findings were considered of minor or no relevance and therefore excluded. Each participant session was analysed independently by each author to achieve concordance and inter-rater reliability, as well as to get different perspectives and assure all the important statements were covered.

Questionnaires

Evaluation of interactive public display systems is a challenge due to it's novelty and lack of standardized and well validated questionnaires [1]. Therefore the questions in the participant questionnaire were adopted from Turunen et al. [42], Hassenzahl [17] and 3 questions that we considered to be useful based on the experience from Lauridsen et al. [36]. The questions relate to affinity, amazement, clearness, ease of use, efficiency, entertaining, effort, naturalness, pleasantness, speed, reliability, robustness, and usefulness. These questions were selected in order to retrieve the participant's evaluation of various traits for the interaction techniques and ensure that all the participants gave their opinion so we could compare them.

The goal of providing the onlookers with an observer questionnaire was to get a different point of view on the interaction techniques that might reveal issues that the participant isn't aware of. We based the questions on what Lauridsen et. al. [36] in their literature review identified as being the 3 most important aspects that prevents people from using a public system: accessibility, learnability and social inhibition. Besides keeping the observers occupied with observing the participant we also use this opportunity to get their demographics.

In the Experience questionnaire we ask the participant which technique they preferred the most, dislike the most and if they felt pressure or embarrassment during the test. These questions are open so the participant may provide reasoning, which can contain details that could be important for understanding the strength and weaknesses of the technique in the public space. At the end of the questionnaire we allow for the participant to provide any additional comments they want to.

FINDINGS

This section presents the quantitative and qualitative data we gathered during the experiment.

Quantitative data

Our source for this data is the Participant questionnaire, that each participant filled in, after the performance of each technique. Before the start of analysis of the data we computed Cronbach's alpha, which is used as a (lowerbound) estimate of the reliability of the data, the alpha coefficient for the four items is .939, suggesting that the items have relatively high internal consistency. Table 4 presents the first and second ranked techniques that accumulated the most positive score of the participants in relation to the questions we asked.

After using descriptive statistics to construct Table 4, we used Friedman's ANOVA (also known as Friendman's test), to allow us comparing the repeated ordinal data that we collected.

Table 5 displays the result after computation of the statistical measurement. An interesting discovery is the fact that there is no statistical differences between the different interaction techniques for three questions, questions 4, 5, and 11. That can be translated as there was not statistically significant difference in between the level of confidence participants had during performance of each technique ($p-value = 0.07$), as well as the level of clearness each techniques had($p-value = 0.06$), and also the effort each technique required($p-value = 0.18$).

Based on our previous work [36], we had two specific questions in order to measure perceived success rate (Question 12) and time (Question 7) for each technique. We wanted to see if there would be a discrepancy in the results with previous findings from Lauridsen et al. [36].

In terms of success rate in placing the information at the right location, our previous quantitative results rank the techniques as follows: *Swipe*, *Throw*,*Grab*, and *Tilt*. We found out exactly the same from the perceived success rate from the participants in this study. In terms of time taken to use the techniques previous results rank the techniques *Swipe*, *Tilt*, *Throw*, and *Grab*. When asked the same question the participants reported that they felt *Throw* was faster than *Tilt*; however, the first and last positions remain the same, rearranging the techniques to *Swipe*, *Throw*, *Tilt*, and *Grab*, as perceived for time taken by the participants.

Results from participant's questionnaires

Question	First pick	Runner up
1. I (wouldn't / **would**) recommend this interaction technique to others	Swipe	Grab
2. This form of interaction technique is (not useful / **useful**) for putting content to a public display	Swipe	Grab
3. The interaction technique feels (unnatural/**natural**)	Swipe	Grab
4. I feel (doubtful/**confident**) about performing the interaction technique	Swipe	Grab
5. This interaction technique is (confusing/**clear**)	Swipe	Tilt
6. This form of interaction technique is (unpleasant/**pleasant**)	Swipe	Grab
7. This form of interaction technique is (slow/**fast**)	Swipe	Throw
8. This form of interaction is (ordinary/ **novel**)	Grab	Throw
9. Performing the interaction is (boring/**fun**)	Grab	Throw
10. Performing this interaction is (difficult / **easy**)	Swipe	Throw
11. The interaction required (too much/ **too little**) efforts	Swipe	Grab
12. The content (didn't went / **went**) where I placed it [wanted it] on the screen.	Swipe	Throw

Table 4. Current table displays the first and second technique as ranked by the participants for each question. The bolded option is the positive score.

Results from Friendman's test

Question	χ^2	p$-$*value*
1	19.04	0.00
2	12.78	0.01
3	15.00	0.00
4	6.94	0.07
5	7.31	0.06
6	19.28	0.00
7	8.48	0.04
8	14.35	0.00
9	8.70	0.03
10	15.19	0.00
11	4.90	0.18
12	7.75	0.04

Table 5. Computation of Friendman's test with included chi-square and p-values.

We also computed a correlation matrix to see the relation of results obtained for different questions. We found that ordinary/novel (question 8) did not correlate strongly with the other questions and therefore we will not include it in our comparison. From this analysis it can be seen that people's responses to techniques include the following relations:

- Recommended if useful,
- Useful if easy to use,
- Natural if pleasant
- Pleasant if natural
- Fast if easy to use
- Fun if natural or easy to use
- Easy to use if reliable
- Reliable if easy to use

These relations are represented in Figure 4.

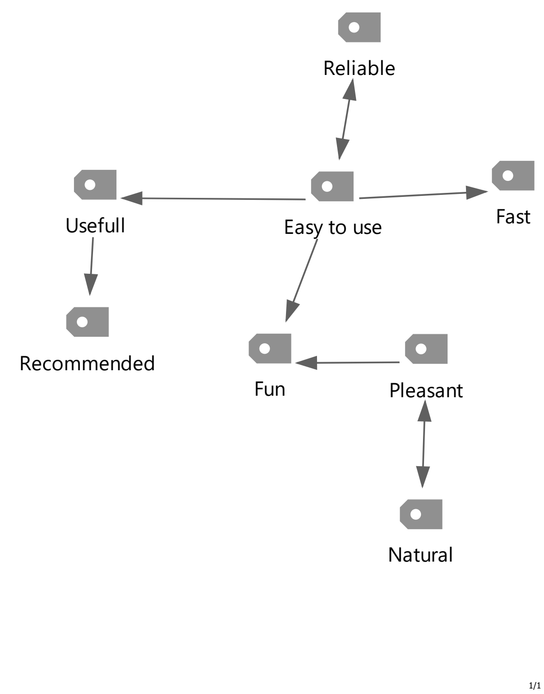
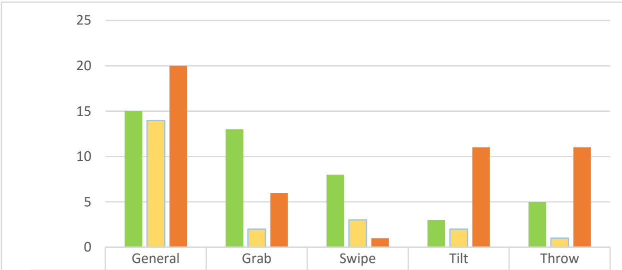


Figure 4. Correlated questions: A question is pointed to by an arrow from the question that is its highest correlation.

From these relations we can see that *Easy to use* has a strong correlation to other questions, meaning that a technique that is considered to be *Easy to use* more likely will be considered *Fast*, *Fun* and *Useful*. *Easy to use* and *Reliable* affect each other, this means that a technique that is not hitting where the users want it to hit will not be considered easy to use. Likewise we have the same relation between *Natural* and *Pleasant*. Slightly different for *Fun* is that there are two arrows pointing to it from two unrelated questions, this is due to the difference between the correlation from *Pleasant* to *Easy to use* is only 0.005. Meaning that we can expect *Fun* to mean that the technique is both easy and pleasant to use.

This helped us in understanding how to group the codes during the thematic analysis.



	General	Grab	Swipe	Tilt	Throw
Positive	15	13	8	3	5
Neutral	14	2	3	2	1
Negative	20	6	1	11	11

Figure 5. Codes divided into positive, neutral and negative groups.

Qualitative data

In this section we analyse the qualitative data gathered from video and audio recordings during the experiment. Both sources have been used in the process of transcribing the responses that occurred during the experiment. After an initial coding we divided the codes into a *positive*, *neutral* and *negative* group, each of the group containing both general and *techniques* specific codes. The distribution of the codes in these 3 groups can be seen in Figure 5. We identified common grounds for the *techniques* in each group, which we used to create sub-categories for the various techniques: *Grab*, *Swipe*, *Tilt* and *Throw*. Likewise we identified common ground for the general codes that resulted in dividing the general codes into a *people* category and an *other* category. The category *people* contains commonalities with a human factor; while any remaining codes would be collected in the *other* category as these are codes that did not fit well into the *people* and *techniques* categories, such as system codes and general comments.

After our thematic analysis we have 115 codes that have been grouped into 38 code containing sub-categories. By finding overlapping categories in the transcripts we are able to identify categories that can be related to each other, which further helped us to define a set of challenges related to cross-device natural user interactions. Those challenges are: Clarity, Comfort, Context and Control. The quotes in this section are labelled with Q for the interviewer and P for the participant.

Clarity: The challenge of ease of use for any interaction technique is complex, related to the person's expectations and how well those expectations are fulfilled. By examining the frequencies of co-occurrences between the techniques' characteristics and the participants' statements, we found out that two techniques stood out, those were *Swipe* and *Grab*. For example we see a participant describing how the swipe technique being fast, familiar, and natural makes it easier to use:

> (P 18) *Yeah, and I am also quite faster since it is a lot more natural. I am familiar with the gesture and I also don't have to think how to do it.*

In contrast to that we have *Grab*, a two-handed interaction that is more complex than *Swipe*. *Grab* is also considered to be easy to use; however, for a different reason.

> (P 12) *I think it's pretty positive that it feels like one motion; that I pick up the image and I place it again.*

> (P 13) *I actually found it to be both fast and accurate. It was like grabbing it and then bam.*

> (P 17) *I think this one feel quite natural it's almost the same way as you would have do it in real life.*

> (P 21): *Yes, because here you don't have to focus on two things at the same time You have to focus one time down here, then you hold onto it and throw it there.*

Comfort: As we performed our experiment in a public setting this naturally led to it being a challenge, which we examined with the participants.

A key issue in a public space is to get comfortable, this is mostly achieved by immersion in what you are doing. We were surprised that most participants did not experience any discomfort. The participants described their experience about that.

> (P 13) *Yeah I do(feel comfortable), I can't really see them (the people). [...] I have social anxiety.*

> (P 24) *I have actually not thought about me <looks around> being in public, because I'm very concentrated on this.*

> (P 8) *Yes I don't have any problems I feel confident. I still focus like 99 percent of my concentration to trying to what I am doing.*

While our participants were comfortable in the public space, the *Throw* technique stood out as being unacceptable due to its movement being exaggerated, larger than what the participants were comfortable with.

> (P16) *It would more be like if there were seventeen people standing behind me and I swing it like this...*
> (Q) *Bump into them...*
> (P 16) *Yeah. Or toss my phone at them.*

> (P 21): *Yeah if you are in London underground then there will be a lot of people behind you. There it won't always be good to make a swing with the arm. There will the other (tilt) perhaps be a little better.*

The *Tilt* was not considered unacceptable; however, it was uncomfortable to some of our participants due to pain.

> (P 13) *It's straining [his hand] a bit.*
> (Q) *You don't feel physically comfortable with this gesture?*
> (P 13) *No.*

> (P 2) *I think that overall maybe the tilt could create some cramps.*

Context: We found that the context is important for the evaluation of the techniques. A technique that is not fitting for transferring files from a mobile device to a public display might be considered more fitting in a different context. We also discovered that the concern for accuracy would depend on the context; since accuracy becomes a characteristic of the techniques it is a question about whether accuracy is useful or not in the context.

One of the participants did not find it important to aim for the targets in the experiment:

> (P 2) *Yeah, Well, I just take ehh like… they are suggestions.*

The phenomenon of not hitting or hitting around the target, during the experiment, formed a tendency which was recurring from time to time. We were curious about the cause behind it.

> (P 21): *Maybe I won't be precise but I would say hey it's ok because it's up there.*

> (P 13) *Depending on the context. If it was in the group rooms or some room where you need to make a lot of posts right next to each other, then precision would be key, but in a place where I just need to show this document, or this picture and it just like a presentation, they need to see this right now then precision wouldn't be key.*

> (Q) *Because the image overlaps, do you feel like it was close enough to the target even though you missed the target?*
> (P 16) *Yeah, I guess so.*

Some of the participants suggested alternative uses for the *Throw* and *Tilt* techniques. The *Throw* had too large movement to be good in confined spaces and the *Tilt* motion was painful in consecutive use.

Throw:

> (P 7) *[...]I would say this one is really good like for games for somewhere outside a bar maybe. Yeah I would say its super useful for games.*

> (P 23) *It might be, if I was in a sports stadium doing something..competing in a sports thing, its a lot more… <throws as if throwing a ball>*
> (Q) *Or perhaps if you had in a game where I…*
> (P 23) *Yeah, it would have to fit the context.*

Tilt:

> (P 8) *Yes, i think if for instance you want to print an image really fast and you don't have a printer at home you can just go and do this to the local printer.*

<u>Control:</u> During the experiment we saw some people had a lack of control with the use of one-handed techniques *Swipe* and *Tilt* due to visibility, fatigue and grip of the phone.

> (P15) *[...] It is just the target thing, because if I have to stretch my hand it's quite hard to see what I am swiping.*

> (P 9) Participant states that it is uncomfortable having to stretch fully her hand (with the phone) in order to point.

> (Q) *So put the phone up and aim with it, now put it down and aim with your hand, what do you like more?*
> (P19) *Well I think in the long run without the phone. Also less weight so don't get tired.*

The grip issue caused some of the participants to prefer performing the one-handed techniques with two hands.

> (Q) *So do you feel like you would rather want to use two hands for swiping rather than one hand?*
> (P 23) *Yes. Otherwise, I think I'm gonna drop it.*

> (P 14) *I felt really good. The only problem I had it is difficult to have a hand motion that is possible to both swipe the top off and the bottom off accurately. This is the reason why I decided to use my second hand.*

Additionally many of the participants reported the two-handed gestures to be advantageous due to increased coordination.

> (P 13) *Yeah I feel like it is more easy this way. Basically I use one hand for pointing and the other is like the controlling one.*

> (P 14) *There are definite advantages to using two handed techniques because you can handle both aiming with one hand and managing stuff with the other hand.*

> (P 21) *[...]I was able to do Throw very precisely because you use 2 hand coordination.*

Some of the participants were more concerned; however, they were concerned in a control related matter i.e. for a specific technique, which was the Tilt.

> (P 12) *Okay. I feel it's hard to both keep my hand steady and make this hard motion at the same time. I don't feel very confident that the image will land where I'm trying to throw it, because I feel the bop - the pointer moves when I do this gesture (tilt).*

> (P 21) *Tilt was quite less accurate as you would do a tilt and that would move the pointer, so i always hit the area around (the target).*

> (P 18) *Not completely accurately, it doesn't always register the tilt. And it's pretty annoying that changes (the position of the pointer).*

DISCUSSION

In the qualitative findings we identified 4 design considerations based on the participants statements - Clarity, Comfort, Context, and Control. We are able to see that the perceived usefulness does not correspond directly with the statistical findings from the qualitative data in this paper as well as Lauridsen et al. [36] This becomes apparent though the analysis of these design considerations:

Clarity: This aspect is based on the input of the users from the qualitative study involving multiple characteristics around each technique, from perceived speed to perceived intuitiveness. In the quantitative study we directly asked the question is it easy to use or not, we did so in order to have a baseline on which to base our in-depth analysis.

Based on that benchmark question the easiest to use techniques are *Swipe* and *Throw*, while *Grab* and *Tilt* are perceived as less easy. However, once we went through the video logs we noticed that actually *Grab* has more "points" in the ease of use department, we asked why this discrepancy? The result of our query was that the naive approach did not properly reflect the feelings of the participants.

Further inquiry depicted that situation: intuitiveness, speed, and enjoyment are factors also to be taken into consideration. An example could be seen when a participant states that *Grab* " could " have been good if a two-finger release was enough, and even though the nature of the statement is negatively said the participant still points to the technique as a most liked. Another participant who also indicated *Grab* as the most liked technique states that it felt seamless and " after " getting used to it, was the easiest to run for the longest time period.

It is clear that cross-device natural user interaction techniques being faced with a single binary variable - " easy / not easy " to use is not a viable option, instead a balance should be considered in order to create an optimal solution. The techniques could, and should be observed in a situated performance e.g. in our case a public space.

Comfort: In our experiment many of our participants were comfortable and forgot about their surrounding from being immersed with the techniques and the system. We see that multiple of the participants stated they were concentrating on what they were doing and we believe this is the cause that made them forget about their surroundings. It also means that the participants did not find any of the techniques so unacceptable or embarrassing that they would focus on that rather than the task.

By leading the participants to think about the space or a different space with less room to move about, we were able to identify that excessive movements was a concern due to personal space and safety. However, these concerns were also dependent on the context, that is that the context requires certain conditions to be met for a technique to be unacceptable.

A discomfort that is not related to the context is pain and we discovered that most participants felt pain from performing the *Tilt* technique over time. Since discomfort should be avoided we would suggest to use the *Tilt* with a slow motion or minimize the amount of repeated motions.

Context: It is important to keep in mind what context the technique is being used in, as contexts sets differing conditions and social etiquette. We see this clearly by how the participants suggested the *Throw* to be used for playing games in a context that is more suitable for fun and large motions.

Our system could be considered to be used for sharing files in a work space, presenting material during a meeting, being a digital bulletin board in a public space or a game in a public space. This allowed the user to imagine different context and therefore better relate to how the techniques could be used. However, the system resembled and was presented as a digital bulletin board. This was more fitting for the selected public area, but could lead to participants being concerned about etiquette other than if the system resembled a game.

We also noticed that some were less concerned about the accuracy of the techniques, where one stated that this was dependent on the purpose of the task. Therefore it is possible that the techniques we evaluated in our context, a university canteen during lunchtime, would be ranked differently in another context.

Control: It is important to be in control of the actions you perform to ensure reliable use of the system, which in turn is important for user satisfaction and engagement [31].

The one-handed techniques *Swipe* and *Tilt* can lead to perceived lack of control and therefore require more focus from a user, distracting them from focusing on the task. In many cases, the participants reported the two-handed techniques to be advantageous due to separating the aiming and the interaction, and thereby providing a better coordination. The *Swipe*, which requires a touch motion to be done with the hand holding the phone, causes a weaker grip on the phone and thus some participants feared they might drop the phone. This ultimately led to some participant using two-handed *Swipe* to continue with the experiment, refusing to use only one hand. Being able to use the technique at all becomes important for the engagement of the system, a participant stated that if there had not been a tutorial system then *"If I wasn't shown how to do it [...] then I will just go away and say it's just not working for me (interaction with the system)"* and therefore it is important to ensure that the techniques are reliable and easy to perform.

LIMITATIONS

This study is limited due to the amount of participants involved, which could have been higher. Some of the participants also were involved with a similar study conducted by different researchers.

The implementation of *Grab* was made by telling the Kinect to read if a palm is open or not. However, most people grabbed the object from the phone with two fingers and tried to release it like that, for some it worked, for some it did not. This was related to the size of the hand and if the surface of an two-finger opened hand is large enough to be read by the Kinect.

Even at an optimal distance from the Kinect and the use of a calibration system to account for the participant's height, some participants complained about how high they had to raise their hand to reach the top of the screen on the large display. Some people complained about the size of the phone used for the experiment, stating that it was difficult to reach the top object on the mobile device when using the *Swipe* technique. It is possible that by removing these limitations that we would have sen a different outcome in our study.

CONCLUSION

In this paper we investigated people's responses to four different push techniques to find out how we can understand them in relation to pushing information from a mobile device to a public display. We established a field study that contained questionnaires, open questions and unstructured interview that was recorded. This resulted in quantitative and qualitative data. We used the quantitative data to identify relations between characteristics of a technique and to show the statistical evaluation of the techniques for the perceived usability and utility. By analysing the qualitative data we found four aspects of interest that should be considered when designing interaction technique, these were Clarity, Comfort, Context and Control.

The qualitative data allowed us to derive that the perceived characteristics of a technique can be conditional by the public setting and context that the data is gathered in; therefore it isn't enough to count 'points' in the quantitative data and assume that these points determine what technique is best. Even the tilt which both statistically and as perceived by the participants is the worst technique of the 4, has been suggested an alternative context where the technique could succeed.

This research contributes with a framework of four design considerations: Clarity, Comfort, Context and Control to aid designers of cross-device natural user interaction and public displays.

FUTURE WORK

The relations as well as the areas of interest require further evaluation in future research through a longitudinal implementation of a system that can be evaluated in the wild by subjective as well as objective measures. By doing so, we will be able to refine the relations as well as the design considerations.

Also, the study limits itself to focus only on mobile devices, such as smartphones. It would be interesting to implement the interaction techniques through other technologies, such as tablets or smarthwatches for further understanding on how they can be applied and how does that affect the relations as well as the areas of interest.

Finally it would be interesting to investigate more closely gender differences, focus on different types of content, and explore effects on local communities.

REFERENCES

1. Florian Alt, Alireza Sahami Shirazi, Thomas Kubitza, and Albrecht Schmidt. 2013. Interaction Techniques for Creating and Exchanging Content with Public Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 1709–1718. DOI: http://dx.doi.org/10.1145/2470654.2466226

2. Antonella Angeli, Laurent Romary, and Frederic Wolff. 1999. *Modeling and Using Context: Second International and Interdisciplinary Conference,CONTEXT'99 Trento, Italy, September 9–11, 1999 Proceedings*. Springer Berlin Heidelberg, Berlin, Heidelberg, Chapter Ecological Interfaces: Extending the Pointing Paradigm by Visual Context, 91–104. DOI: http://dx.doi.org/10.1007/3-540-48315-2_8

3. Rafael Ballagas, Michael Rohs, and Jennifer G. Sheridan. 2005. Sweep and Point and Shoot: Phonecam-based Interactions for Large Public Displays. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems (CHI EA '05)*. ACM, New York, NY, USA, 1200–1203. DOI: http://dx.doi.org/10.1145/1056808.1056876

4. Sara A. Bly, Steve R. Harrison, and Susan Irwin. 1993. Media Spaces: Bringing People Together in a Video, Audio, and Computing Environment. *Commun. ACM* 36, 1 (Jan. 1993), 28–46. DOI: http://dx.doi.org/10.1145/151233.151235

5. S. A. Bly and S. L. Minneman. 1990. Commune: A Shared Drawing Surface. In *Proceedings of the ACM SIGOIS and IEEE CS TC-OA Conference on Office Information Systems (COCS '90)*. ACM, New York, NY, USA, 184–192. DOI: http://dx.doi.org/10.1145/91474.91514

6. Sebastian Boring, Marko Jurmu, and Andreas Butz. 2009. Scroll, Tilt or Move It: Using Mobile Phones to Continuously Control Pointers on Large Public Displays. In *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7 (OZCHI '09)*. ACM, New York, NY, USA, 161–168. DOI: http://dx.doi.org/10.1145/1738826.1738853

7. Richard Borovoy, Fred Martin, Sunil Vemuri, Mitchel Resnick, Brian Silverman, and Chris Hancock. 1998. Meme Tags and Community Mirrors: Moving from Conferences to Collaboration. In *Proceedings of the 1998 ACM Conference on Computer Supported Cooperative Work (CSCW '98)*. ACM, New York, NY, USA, 159–168. DOI:http://dx.doi.org/10.1145/289444.289490

8. Andrew Bragdon, Rob DeLine, Ken Hinckley, and Meredith Ringel Morris. 2011. Code Space: Touch + Air Gesture Hybrid Interactions for Supporting Developer Meetings. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces (ITS '11)*. ACM, New York, NY, USA, 212–221. DOI: http://dx.doi.org/10.1145/2076354.2076393

9. Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. DOI: http://dx.doi.org/10.1191/1478088706qp063oa

10. Victor Cheung, Diane Watson, Jo Vermeulen, Mark Hancock, and Stacey Scott. 2014. Overcoming Interaction Barriers in Large Public Displays Using Personal Devices. In *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces (ITS '14)*. ACM, New York, NY, USA, 375–380. DOI:http://dx.doi.org/10.1145/2669485.2669549

11. Jennica Falk and Staffan Bjork. 1999. The BubbleBadge: A Wearable Public Display. In *CHI '99 Extended Abstracts on Human Factors in Computing Systems (CHI EA '99)*. ACM, New York, NY, USA, 318–319. DOI: http://dx.doi.org/10.1145/632716.632909

12. Patrick Tobias Fischer and Eva Hornecker. 2012. Urban HCI: Spatial Aspects in the Design of Shared Encounters for Media Facades. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 307–316. DOI: http://dx.doi.org/10.1145/2207676.2207719

13. Robert S. Fish, Robert E. Kraut, and Barbara L. Chalfonte. 1990. The VideoWindow System in Informal Communication. In *Proceedings of the 1990 ACM*

Conference on Computer-supported Cooperative Work (CSCW '90). ACM, New York, NY, USA, 1–11. DOI: http://dx.doi.org/10.1145/99332.99335

14. William Gaver, Thomas Moran, Allan MacLean, Lennart Lövstrand, Paul Dourish, Kathleen Carter, and William Buxton. 1992. Realizing a Video Environment: EuroPARC's RAVE System. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '92)*. ACM, New York, NY, USA, 27–35. DOI:http://dx.doi.org/10.1145/142750.142754
15. Saul Greenberg and Michael Rounding. 2001. The Notification Collage: Posting Information to Public and Personal Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '01)*. ACM, New York, NY, USA, 514–521. DOI: http://dx.doi.org/10.1145/365024.365339
16. Robert Hardy and Enrico Rukzio. 2008. Touch & Interact: Touch-based Interaction of Mobile Phones with Displays. In *Proceedings of the 10th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '08)*. ACM, New York, NY, USA, 245–254. DOI: http://dx.doi.org/10.1145/1409240.1409267
17. Marc Hassenzahl. 2005. *Funology: From Usability to Enjoyment*. Springer Netherlands, Dordrecht, Chapter The Thing and I: Understanding the Relationship Between User and Product, 31–42. DOI: http://dx.doi.org/10.1007/1-4020-2967-5_4
18. Jeremy M. Heiner, Scott E. Hudson, and Kenichiro Tanaka. 1999. The Information Percolator: Ambient Information Display in a Decorative Object. In *Proceedings of the 12th Annual ACM Symposium on User Interface Software and Technology (UIST '99)*. ACM, New York, NY, USA, 141–148. DOI: http://dx.doi.org/10.1145/320719.322595
19. Uta Hinrichs, Sheelagh Carpendale, Nina Valkanova, Kai Kuikkaniemi, Giulio Jacucci, and Andrew Vande Moere. 2013. Interactive Public Displays. *IEEE Comput. Graph. Appl.* 33, 2 (March 2013), 25–27. DOI: http://dx.doi.org/10.1109/MCG.2013.28
20. Hiroshi Ishii and Minoru Kobayashi. 1992. ClearBoard: A Seamless Medium for Shared Drawing and Conversation with Eye Contact. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '92)*. ACM, New York, NY, USA, 525–532. DOI:http://dx.doi.org/10.1145/142750.142977
21. Hiroshi Ishii, Craig Wisneski, Scott Brave, Andrew Dahley, Matt Gorbet, Brygg Ullmer, and Paul Yarin. 1998. ambientROOM: Integrating Ambient Media with Architectural Space. In *CHI 98 Cconference Summary on Human Factors in Computing Systems (CHI '98)*. ACM, New York, NY, USA, 173–174. DOI: http://dx.doi.org/10.1145/286498.286652
22. Karrie Karahalios and Judith Donath. 2004. Telemurals: Linking Remote Spaces with Social Catalysts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '04)*. ACM, New York, NY, USA, 615–622. DOI: http://dx.doi.org/10.1145/985692.985770
23. Sherrie Rabinowitz Kit Galloway. 1980. Hole in Space. (November 1980). http://www.ecafe.com/getty/HIS/
24. J. Lazar, J.H. Feng, and H. Hochheiser. 2010. *Research Methods in Human-Computer Interaction*. John Wiley & Sons. https://books.google.dk/books?id=H_r6prUFpc4C
25. Nicolai Marquardt. 2011. Proxemic Interactions in Ubiquitous Computing Ecologies. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems (CHI EA '11)*. ACM, New York, NY, USA, 1033–1036. DOI: http://dx.doi.org/10.1145/1979742.1979691
26. Nicolai Marquardt, Ken Hinckley, and Saul Greenberg. 2012. Cross-device Interaction via Micro-mobility and F-formations. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology (UIST '12)*. ACM, New York, NY, USA, 13–22. DOI: http://dx.doi.org/10.1145/2380116.2380121
27. Jörg Müller, Florian Alt, Daniel Michelis, and Albrecht Schmidt. 2010. Requirements and Design Space for Interactive Public Displays. In *Proceedings of the 18th ACM International Conference on Multimedia (MM '10)*. ACM, New York, NY, USA, 1285–1294. DOI: http://dx.doi.org/10.1145/1873951.1874203
28. Jorg Muller, Gilles Bailly, Thor Bossuyt, and Niklas Hillgren. 2014. MirrorTouch: Combining Touch and Mid-air Gestures for Public Displays. In *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services (MobileHCI '14)*. ACM, New York, NY, USA, 319–328. DOI:http://dx.doi.org/10.1145/2628363.2628379
29. Jorg Muller, Robert Walter, Gilles Bailly, Michael Nischt, and Florian Alt. 2012. Looking Glass: A Field Study on Noticing Interactivity of a Shop Window. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 297–306. DOI: http://dx.doi.org/10.1145/2207676.2207718
30. Richard Stenzel Norbert Streitz, Thorsten Prante and Daniel van Alphen. 2003. Situated Interaction with Ambient Information: Facilitating Awareness and Communication in Ubiquitous Work Environments. In *Human-Computer Interaction*.
31. Heather L. O'Brien and Elaine G. Toms. 2008. What is user engagement? A conceptual framework for defining user engagement with technology. *JASIST* 59, 6 (2008), 938–955. DOI:http://dx.doi.org/10.1002/asi.20801
32. Sharon Oviatt and Philip Cohen. 2000. Perceptual User Interfaces: Multimodal Interfaces That Process What Comes Naturally. *Commun. ACM* 43, 3 (March 2000), 45–53. DOI:http://dx.doi.org/10.1145/330534.330538

33. Bradley J. Rhodes. 1997. The Wearable Remembrance Agent: A System for Augmented Memory. In *Proceedings of the 1st IEEE International Symposium on Wearable Computers (ISWC '97)*. IEEE Computer Society, Washington, DC, USA, 123–. http://dl.acm.org/citation.cfm?id=851036.856441

34. Julie Rico and Stephen Brewster. 2010. Usable Gestures for Mobile Interfaces: Evaluating Social Acceptability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 887–896. DOI: http://dx.doi.org/10.1145/1753326.1753458

35. Harry Brignulland Yvone Rogers. 2003. Enticing People to Interact with Large Public Displays in Public Space. In *Proc*. INTERACT 3,, INTERACT, 17–24.

36. Bjarke Martin Lauridsen Ivan Svilenov Penchev Elias Ringhauge Eric Vignola Ruder. 2016. *Cross-device interaction between Mobile Phones and Large Displays*. Technical Report. Aalborg, Denmark. http://projekter.aau.dk/projekter/da/studentthesis/crossdevice-interaction-between-mobile-phones-and-large-displays(33529544-532c-4148-aff7-ac7b929ee3df).html.

37. Dominik Schmidt, Julian Seifert, Enrico Rukzio, and Hans Gellersen. 2012. A Cross-device Interaction Style for Mobiles and Surfaces. In *Proceedings of the Designing Interactive Systems Conference (DIS '12)*. ACM, New York, NY, USA, 318–327. DOI: http://dx.doi.org/10.1145/2317956.2318005

38. Julian Seifert, Adalberto Simeone, Dominik Schmidt, Paul Holleis, Christian Reinartz, Matthias Wagner, Hans Gellersen, and Enrico Rukzio. 2012. MobiSurf: Improving Co-located Collaboration Through Integrating Mobile Devices and Interactive Surfaces. In *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces (ITS '12)*. ACM, New York, NY, USA, 51–60. DOI: http://dx.doi.org/10.1145/2396636.2396644

39. Dave Snowdon and Antonietta Grasso. 2002. Diffusing Information in Organizational Settings: Learning from Experience. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '02)*. ACM, New York, NY, USA, 331–338. DOI: http://dx.doi.org/10.1145/503376.503435

40. A. Strauss and J.M. Corbin. 1998. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. SAGE Publications. https://books.google.dk/books?id=wTwYUnHYsmMC

41. Norbert A. Streitz. 1999. i-LAND: An Interactive Landscape for Creativity and Innovation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '99)*. ACM, New York, NY, USA, 120–127. DOI: http://dx.doi.org/10.1145/302979.303010

42. Markku Turunen, Jaakko Hakulinen, Aleksi Melto, Tomi Heimonen, Tuuli Laivo, and Juho Hella. 2009. SUXES - user experience evaluation method for spoken and multimodal interaction. In *Annual Conference of the International Speech Communication Association*. 2567–2570.

43. Jouni Vepsalainen, Antonella Di Rienzo, Matti Nelimarkka, Jouni A. Ojala, Petri Savolainen, Kai Kuikkaniemi, Sasu Tarkoma, and Giulio Jacucci. 2015. Personal Device As a Controller for Interactive Surfaces: Usability and Utility of Different Connection Methods. In *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces (ITS '15)*. ACM, New York, NY, USA, 201–204. DOI: http://dx.doi.org/10.1145/2817721.2817745

44. W. Ward. 1991. Understanding Spontaneous Speech: The Phoenix System. In *Proceedings of the Acoustics, Speech, and Signal Processing, 1991. ICASSP-91., 1991 International Conference (ICASSP '91)*. IEEE Computer Society, Washington, DC, USA, 365–367. http://dl.acm.org/citation.cfm?id=1170742.1170864

45. M. Weiser. 1991. The computer for the 21st century. *Scientific American* 265, 3 (Sept. 1991), 94–104. DOI: http://dx.doi.org/10.1038/scientificamerican0991-94

46. Alan Wexelblat. 1998. Research Challenges in Gesture: Open Issues and Unsolved Problems. In *Proceedings of the International Gesture Workshop on Gesture and Sign Language in Human-Computer Interaction*. Springer-Verlag, London, UK, UK, 1–11. http://dl.acm.org/citation.cfm?id=647590.728557

Concluding Remarks

This section will summarise this report, in which the approach and outcome are covered. This in terms will lead to a clear and meticulous description of our contributions within the research of HCI and interaction techniques. Finally, we will take a look at future work that needs to be conducted in order to accomplish a more complete understanding of what can be considered helpful within cross-device natural user interaction for mobile devices and public displays.

1 Project summary

This project has been an extension of our 9th semester project that took place in laboratory settings. In that project we explored and compared, empirically, four cross-device natural user interaction techniques based on previous studies and theory, to help people interact with large displays. We learned about the challenges of creating successful, efficient and accurate cross-device interaction techniques that feel natural for the users to perform, in a way that they are capable of walking up to the display and performing the technique with as little guidance as possible. The multitude of findings and observations that were gathered from that project, justified continued work within this field for the 10th semester with a shift in the approach we take.

In this study we decided to increases the factors that a user performing an interaction technique is exposed to, by doing it in a public setting, compared to laboratory setting. To achieve our goal we developed a system, consisting of a client and server, which uses four different cross-device natural user interaction techniques for a realistic task - transferring files and images from a mobile device to a public display. We then conducted a series of test sessions with groups, where each group member was an observer, and a user of the system. After our experiment period ended we carefully gathered and examined our data. Our approach resulted in four design considerations which should be accounted for when one design cross-device natural user interaction techniques.

2 Research contribution

We discovered four design considerations that we consider important when designing systems to support cross-device natural user interactions. These aspects were: Clarity, Comfort, Context, and Control. This framework can be used to analyse and understand existing public systems and design new ones. We found that while quantitative analysis of the interaction techniques can provide measurement about usability and be used for finding the relation between important aspects, it does not provide details that can be crucial for engaging with a public system. We suggest to perform qualitative measurement in a context that is similar or identical to the context that the system has to be used in, as the context can play a factor in how the techniques are perceived. The qualitative measurement is useful for understanding what is socially acceptable in a specific context and may contain information that can lead to an overall improvement of the system. Bringing the qualitative and quantitative measurements together draws a picture of which techniques would be most fitting for the context the examination has been conducted in and thereby aiding the developers in selecting a technique that will be engaging for the user.

3 Future work

In this Master's thesis we have defined relations as well as design considerations. Future research is required to further evaluated them, through the use of a longitudinal implementation of a system that can be evaluated in the wild by subjective as well as objective measures.

Also, the study limits itself to focus only on mobile devices, such as smartphones. In our paper we suggest that it would be interesting to implement the interaction techniques through other technologies, such as tablets or smarthwatches. This could also be extended to more exotic examples such as mind reading wearable like Thync or Muse. This will aid for further understanding on how the technology can be applied and how does that affect the relations as well as design consideration.

An examination that would be interesting to investigate more closely is gender differences, as well as focus on different types of content, and explore effects on local communities.

Finally we suggest that the findings gathered in the study can be used as the basis of alternative designs within this domain. Further work, which includes these findings, is needed in order to design and achieve technology that can fully support people within this context.

APPENDICES

Contents

2016
AALBORG UNIVERSITY

Appendix A: Previous study

Comparison of Push Techniques for Cross-Device Interaction Between Mobile Devices and Large Displays

Bjarke M. Lauridsen
Aalborg University
blauri10@student.aau.dk

Ivan S. Penchev
Aalborg University
ipench14@student.aau.dk

Elias Ringhauge
Aalborg University
eringh10@student.aau.dk

Eric V. Ruder
Aalborg University
eruder10@student.aau.dk

ABSTRACT
In recent years, research into cross-device interaction techniques has increased. Much of this research focuses on interaction between mobile devices and large displays. We contribute to this body of knowledge with an empirical comparison of four different push techniques - *Pinch*, *Swipe*, *Throw*, and *Tilt* for interaction between mobile devices and large displays. We report on success rate, efficiency and accuracy. We also present the ease of use of techniques as perceived by users. We show that *Swipe* was the most effective in terms of success rate, efficiency and accuracy. Furthermore, *Swipe* gathered the highest score, in regards to ease of use, by users. Participants also reported that *Pinch* was the most fun to use.

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous; See http://acm.org/about/class/1998/ for the full list of ACM classifiers. This section is required.

Author Keywords
Interaction Techniques; Cross-Device Interaction; Natural User Interaction; Kinect; Mid-air Gestures; Smartphones; Large Displays; Data Transfer

INTRODUCTION
The evolution of ways people interact with the digital world is noteworthy considering the short life-span of computing. How we use our devices, which devices we use, and the context in which we use them has been continually under transformation. From portable personal computers originally considered mostly for specialized field applications such as accountancy, military use, or for sales representatives, which addressed mobility of a person's workspace, to modern hand-held devices which presents their users with such degrees of freedom that ultimately workspaces are becoming more ubiquitous. In non-work context people are now connected mostly everywhere, which aids us in a search for information or in communicating, changing the way we interact. This expansion has not only increased mobile computing due to greater convenience, but also made it widespread.[6]

As numerous divergent devices are being adopted in different domains and contexts, understanding cross-device interaction is currently becoming more important and relevant; after all, people take their hand-held devices into situations where other technologies are active. This ubiquitous presence of devices means that they can be used to enhance everyday situations in all kind of places. Imagine if a public display could morph from a one-way broadcast device that merely shows visual content to a two-way interaction device that provides a more engaging and immersive experience. This emergence of cross-device communication opportunities prompts a need to understand how different interaction techniques perform in use, i.e. in terms of how easily, quickly and accurately, or in terms of how enjoyable or satisfying it is to interact in this way.

Research in the area of cross-device interaction is increasing with the changing trends. Earliest examples are in the late 90's, within ubiquitous computing, with Rekimoto's work. He argued for what he called multi computer user interface and that interaction techniques must overcome the boundaries among devices in multi-device settings[16].

Recent HCI research has focused on how to include natural modality more in cross-device interaction, contributing to what should be know as cross-device natural user interaction. Some researchers used spatial information [11, 12], others used touch [19], or combined touch with air gestures [4]. But we still have limited understanding of how to design cross-device natural user interaction techniques and we lack empirical studies of this.

Inspired by the opportunities presented by such challenges, this paper reports on a empirical study between four different cross-device natural user interaction techniques for data transfer between mobile devices and large displays. We discovered that out of the four techniques we developed and implemented, *Swipe* was the most effective technique. We also show that even though *Pinch* is not as effective as the others, users described it as a fun technique.

RELATED WORK

Public displays are an inherently visual medium with graphics and animations that are increasingly used to visualize data. However the general case is that presented data can only be viewed and not interacted with. The current trend is a change from non-interactive information broadcasters to active medium information exchange. Combining this with the recent trend of increased number of hand-held device per capita an opportunity for new cross-device applications in public space arises, providing new opportunities for exploring cross-device interaction (CDI), in this context.

In a historic perspective, one of the earliest working cross-device applications is by Myers et al. [13]. One of the applications realized within their *Pebbles project* is SlideShow Commander that utilized Personal Devices Assistants (PDA) to control a PowerPoint presentation running on other computer or laptop. It was possible not just moving between slides, but also scribbling and writing on the PDA slides, while annotations are shown on the presentation for the audience. However the idea of cross-device has deeper roots in ubiquitous computing. Rekimoto's work on *pick-and-drop technique* is one of the earliest examples for exploring a technique that spawns between multiple devices. The technique " *allows a user to pick up an object on a display and drop it on another display as if he/she were manipulating a physical object.*" [15]. These two early work examples, even though in different fields, are related and would provide a foundation for additional research.

An example of such research is that of Boring et al. [2] who not only build a cross-device application but explored the implications of different techniques on it. Boring et al. explored the transfer of data from a large public display onto a mobile device. They created a method of transferring data from a large screen by using the camera on the mobile device. The user would take a picture of whatever content they were interested in, after which the application would query the content server with the picture taken from the user. Through visual analyses, the content server would determine what content the user was interested in and would return that content to them. They show that there is a need for enabling data exchange between mobile devices and public displays. In another study, Boring et al. [3] investigated cross-device interaction between large displays and mobile phones. Investigating three different interaction techniques in order to continuously control a pointer on a large screen from a mobile device. *Move* and *Tilt*, two of the three interaction techniques, enabled faster selection time compared to the last one, *Scroll*, but at the cost of higher error rates. They showed that different interaction techniques have certain strengths and weaknesses, and depending on the context and use, certain techniques are more effective.

Boring et al.'s idea [3] of how to control a public display is only one side of cross-device interaction, a different idea from Nielsen et al. [14] uses collaboration surface made from multiple devices to investigate the use of multiple device together, by allowing a number of devices being put next to each other and " pinched " together to form a larger collaborative workspace. In order to expand on the idea of common workspace, a movement from use of multiple devices to build one large display was needed. Schmidt et al. [18] proposed a cross-device interaction style for mobiles and surfaces where one can use multiple phones to interact with a digital surface.The researchers point out 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"*. In order to overcome this they propose the use of mobiles as tangible input on the surface in a stylus like fashion.

A combination of the ideas is presented by Skov et al. [20] illustrating six different cross-device interaction techniques for the case of card playing. A player can see their own cards on their phone and use three different techniques for playing a card from the hand-held to the tablet, which is placed on a table. In the other direction, i.e. when drawing a card, the player also has three techniques to choose from. The study aims to quantitatively evaluate each of the techniques and shows that there is a difference in time and number of errors between the techniques. They recorded two types of errors, namely, interaction errors and play errors. The number of interaction errors shows how difficult it is for a user to perform a given technique while play errors represent the errors related to the game and is recorded when the user plays a wrong card. The difference in interaction errors is apparent, especially between two of the techniques for playing a card.

The research above illustrates two different direction movement in cross-device interaction, what they have in common, is data transfer between devices. Hamilton and Wigdor's work [7] aggregates much of the works above and clearly articulates the data transfer. They create a prototype framework for cross-device applications by combining a number of interaction techniques for data transfer, chaining tasks, and managing interactions sessions. Data transfer is a challenge and as such there are different approaches for solution. Marquardt et al. [12] study cross-device interaction on tablets with a natural modality, by involving spatial information through proxemics.Based on the constructs of f-formation, micro-mobility, and co-present collaboration, they build their prototype with the idea of support for fluid and minimally disruptive interaction in document transfer. Bragdon et al.[4] propose using a combination of air gestures and touch.They aim to design, implement and test a system that allows, a group of users, to interact using air gestures and touch gestures. The purpose is to increase control, support democratic access, and share items across multiple personal devices such as smartphones and laptops where the primary design goal is fluid, democratic sharing of content on a common display.

INTERACTION TECHNIQUES

In this section we illustrate and describe techniques that can be used to push information from mobile devices to large displays with the purpose of empirically comparing them to each other. We want to find intuitive techniques that allow a user to walk up and use a large display. The techniques are characterized by allowing the users to interact with a large screen in a natural way using one or both hands and their mobile phone.

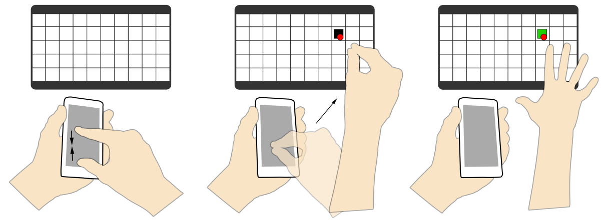
All the techniques were found in the literature and chosen based on a set of criteria outlined in Table 1. We are also interested in examining the effect of each technique on targets of two different sizes, large and small, relative to the size of the screen because the need for precision of an application is not always the same. Sometimes an applications needs to allow the user to place some data in really specific locations, where sometimes a general approximation is enough. We chose four techniques named *Pinch*, *Swipe*, *Throw* and *Tilt* because they fulfill different aspects of the criteria requirements in Table 1 and allow us to compare them to each other in an experimental setup with simple tasks.

Criteria	Description
Number of hands	There must be both one-handed and two-handed techniques.
Previously used	To avoid designing and testing a set of novel techniques, we had the criterion that all techniques must have been used by others before we would use them.
Complexity	The techniques must differ in their complexity and therefore we included techniques with different amount of steps.
Natural feel	There must be a natural feel to the techniques in some way.
Time	The time it takes to perform the different techniques must be different.

Table 1. This table describes the set of criteria.

In the following we explain why these four techniques were chosen and how they should be performed.

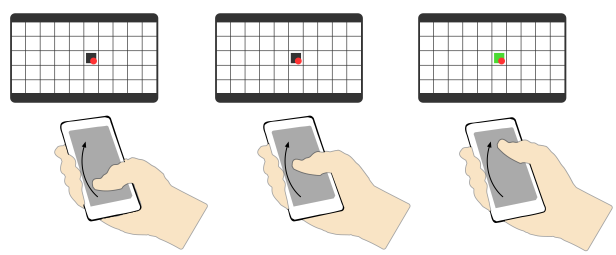
The *Pinch* technique (Figure 1) is used in [8] by Ikematsu et al., as part of a drag-and-drop method for moving data objects between devices. Chen et al. uses a pinching gesture in [5] for cross-device interaction between a smartphone and a smartwatch to control volume. Benko and Wilson [1] used the *Pinch* technique for interacting with omnidirectional visualizations in a dome. This technique is again a combination of the pointing technique used by Scheible et al. [17] and the aforementioned pinching techniques. The reason for including this technique was to imitate the natural action of picking up a real object e.g. piece of paper, and then moving it to another location. With *Pinch* we get a two handed technique which requires the user to perform a series of steps and are thus considered more complex and time consuming than the one handed *Swipe* and *Tilt*. The *Pinch* technique is performed by *1*) holding the phone in one hand and making a pinch gesture on the phone with the other hand (fig. 1a), subsequently closing the hand; *2*) pointing at a target on the large display with the closed hand (fig. 1b); and *3*) opening the hand to complete the technique (fig. 1c).



(a) (b) (c)

Figure 1. (a) First step of the *Pinch* technique is to make a pinch gesture on the phone using the index finger and thumb. (b) Second step is to move and use the pinched hand as a pointer. (c) Third step is to release by opening the hand.

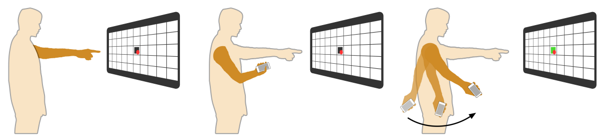
The *Swipe* technique (Figure 2) is used by Bragdon et al. [4] in Code Space, a system using the Kinect and smartphones to support developer meetings. Bragdon et al. describe the technique as: *"cross-device interaction with touch and air pointing"* and the swipe motion is described as *"flicking up on the touch screen"*. This technique was chosen because of its simplistic design and the low level of complexity. Only one hand is needed and the amount of effort and time required to execute this technique is minimal compared to other techniques. *Swipe* is copied exactly as described in [4]. The *Swipe* technique is performed by *1*) pointing at a target on the large display with the phone in a stretched arm, and *2*) making a forwards swipe motion with the thumb on the phone's screen (fig. 2).



(a) (b) (c)

Figure 2. The *Swipe* technique is performed by using the thumb to swipe from (a) to (c).

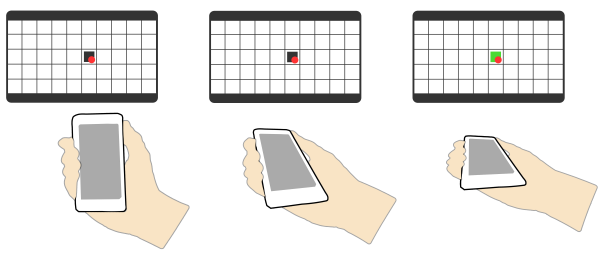
The *Throw* technique (Figure 3) is a combination of a technique for pointing [17] i.e. using a hand as a cursor in mid-air, and a throw technique described by Walter et al. [21] used in a system for sharing information on large public displays. We chose to include this technique based on its natural and playful design. The technique mimics the real world scenario of throwing something like a ball somewhere or to someone. *Throw* is two handed, more complex than aforementioned techniques, and takes a little longer to execute because of the increased number of steps. The *Throw* technique is performed by *1*) pointing at a target on the large display with one hand (fig. 3a); *2*) holding the phone in the other hand and selecting data (fig. 3b); and *3*) making a swinging motion towards the large display (fig. 3c).



(a) (b) (c)

Figure 3. (a) First step of the *Throw* technique is to point at the screen. (b) Second step is to select data. (c) Third and final step is to do a swinging motion with the phone towards the screen.

The *Tilt* technique (Figure 4) is used in a collaborative application by Lucero et al. [9] to transfer an object from a large display to the user's smartphone. Boring et al. [3] use the tilt technique when moving a pointer on a display using a phone and though not the same application, the execution of the technique is the same. We chose this technique because it is one handed, relatively low complexity, and much like the *Swipe* it is generally easy to use. When the direction is reversed, *Tilt* is an exact copy of the way Lucero et al. describe the technique. The *Tilt* technique is performed by *1*) pointing at a target on the large display with the phone in a stretched arm, and *2*) making a forwards tilt with the phone (fig. 4).



(a) (b) (c)

Figure 4. The *Tilt* technique is performed by doing a forward tilt motion with the phone from (a) to (c).

As mentioned, the techniques have been used in previously published papers where they were part of applications or systems, facilitating interaction between devices such as phone to display interaction and vice versa. The described techniques are different in the way they are performed but also the number of hands that are required to make them work. We chose two techniques that are one-handed (*Swipe* and *Tilt*) and two techniques that are two-handed (*Throw* and *Pinch*). Also, to vary the complexity we chose two techniques that require more steps (*Throw* and *Pinch*), and two techniques that require less steps for the user to complete the technique (*Swipe* and *Tilt*).

All the techniques make use of some combination of pointing in mid-air and touch gestures on the smartphone screen. The mid-air pointing interaction is achieved by using the Microsoft Kinect for Windows which uses a depth camera making it possible to track a user's hand in mid-air. As for the touch gestures, smartphones have an accelerometer and a touchscreen, making it possible to detect motion input and detect contact between e.g. a finger and the screen. These technologies are, in combination with each other, used to recognize the four techniques described in this section.

EXPERIMENT

The four cross-device interaction techniques mentioned above were implemented and then evaluated in a lab study in order to judge their performance compared to each other. We are interested in knowing whether or not the different techniques with different target sizes have an effect on the efficiency, accuracy, and ease of use of pushing information to a large display. Therefore, we developed an application that would allow us to run experiments and test the effect of the different techniques and target sizes. We utilized a Microsoft Kinect, a 65' inch Panasonic television with a 1920×1080 resolution and a Samsung Galaxy SII to create the experimental application. An overview of the experiment setup can be seen in Figure 5.

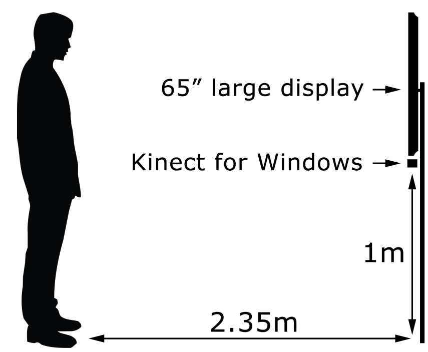


Figure 5. An overview of the entire setup of our experiment.

Implementation

The 4 techniques (*Swipe*, *Tilt*, *Throw*, and *Pinch*) were implemented in order to push data onto the large display. They were implemented in a simple and short target practice application, where the goal was to "hit" the target on the display with the shown technique.

A grid system was implemented in the test application, where each cell of the grid is a possible target. As mentioned, we were interested in measuring the effect of different target sizes on the different techniques, therefore the grid is implemented in two different sizes. The grid system can have large cells, where the grid is 5×10 cells and each cell is 61 pixels or 7.3 cm wide, or it could also have small cells, where the grid is 10×20 cells and each cell is 122 pixels, or 14.6 cm wide. The target is located in one of those cells, and scales accordingly to the size of the cell(See Figure 6a).

A red dot works as the pointer in the screen; it is the location that would be hit when the user performed the given technique. The yellow highlighted cell is the cell in which the pointer is currently located inside of. This is extra feedback for the user so that he knows exactly were he will hit once we performs the technique.

The developed mobile application was simple. It showed two shapes, a circle and a square, which the user could choose to push to the display. The display would tell which shape was the correct one by having that shape as the target in one of the grid cells. We chose two shapes so that it did not become a search problem with users spending too much time searching for the correct shape. We wanted the user to spend some time orienting him or her self with the phone and not just simply performing the gesture without paying any attention to the phone at all. The phone screen can be seen in Figure 6b.

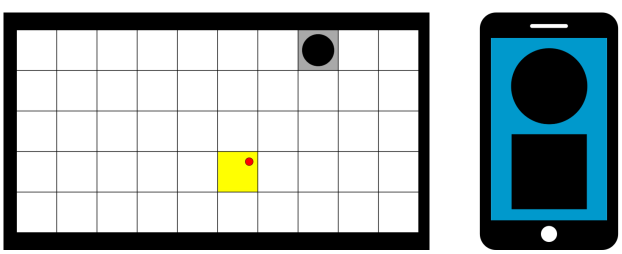
Users would control the pointer on the large display with their hands. Which ever hand was closest to the screen would determine the position of the pointer on the large screen. This meant that users could switch hands whenever they pleased at any point during the test.

The *Pinch* technique was implemented with the help of the Kinect and the touch screen on the mobile phone. This technique started by having the user pinch the shape on the screen of the mobile phone and close his or her hand around it, as if to grab it. The Kinect would then look for a opening of the hand motion, on the pointer hand, and take that as the target point.

The *Swipe* technique was implemented with the touch screen of the phone. Here, we detected when a significant swipe happened on the screen, and then use the pointer location to place the shape that was swiped up onto the screen.

The *Throw* technique was also implemented with the help of the Kinect and the accelerometer on the mobile phone. The Kinect looked at the user to recognize when a user moved the mobile phone from 10 centimeters behind the hip to 10 centimeters in front of the hip. At the same time, the phone detects when a significant change in the accelerometer happened, so to not simply detect an unintentional wave of the arm. The Kinect would then use the position of the other hand to see where on the screen the user intended to perform the *Throw* technique towards.

The *Tilt* technique was implemented mostly with the accelerometer of the phone, by checking for a significant change in the z and y axis of the accelerometer, as if tilting the phone forward.



(a) (b)

Figure 6. (a) The large display screen of the application, with the red dot in the highlighted square as the starting point and the black circle as the target. (b) The phone screen showing the two shapes.

Experimental Design

The experiment was conducted as a within-subject research, with the four different interaction techniques and two target sizes as independent variables.

The within-subject research was chosen because we wanted to minimize the amount of subjects needed in order to get a significant result. We also believed that the learning effect would not be as pronounced since the four techniques are very different from each other. We chose to investigate techniques because we were interested in learning about the way people interact with large displays, and the two target sizes to investigate precision. Sometimes users need to be as precise as possible, and sometimes they just need to be able to interact with a large display.

For the dependent variables, different measures of completion time and hit success were used, as well as a short questionnaire to get the user's satisfaction with regard to the given interaction technique. Which technique started the test was randomized in order to mitigate the learning effect on the entire set of tests. In the end, the *Pinch* gesture started 26.4% of all tests, *Swipe* started 22.7% of all tests, *Throw* started 24.5% of all tests, and *Tilt* started 26.4% of all tests. All of this was automatically logged, and every test session was also video recorded in order to be able to go through them in case we wanted to go into detail in one of the test sessions.

A simple logging mechanism was developed, which created a unique file for each user and outputted all attempts into that file. In the end, the result was a list of 53 files, one for each test participant, where each file would have a list of attempts and target size switches. Each attempt would have a time stamp, whether the user hit the target or not, whether he selected the correct shape, were the target was, and where the participant hit. These where the following measures that we were able to deduce from the log files that were generated:

Total Time: This was the time each user spent completing the test for a given interaction technique. This was measured from the time each user had hit his first target after a practice period of three tries until he had hit his last target. There were a total of 18 targets, plus the first target used for calibration.

Time per target: This was the time each user spent hitting each of the targets. This was measured as the time since each user last hit a target until he hit the next one.

Hit success: Whether or not each user hit the given target. Current pointer and target position (in both cell and pixels coordinates) were also recorded in order to give a precise measure of accuracy for each attempt in terms of *distance to target*.

Ease of use: Each user was given a questionnaire after having gone through each interaction technique. There were 6 questions, all taken from the USE questionnaire [10]. These were asked to get an understanding of how useful and easy to use each technique was. The 6 questions were the following:

- It is easy to use
- Using it is effortless
- It is easy to learn to use
- I can use it successfully every time
- I quickly became skillful with it
- I learned how to use it quickly

Users were able to rate their answers to each question on a 7 point Likert scale. We also wrote down any comments made during the experiment and combined them with the questionnaire responses to get a better understanding of the user's response to each of the techniques.

Participants

In total, 53 people took part in our experiment, which was conducted in a usability lab. 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 one for 2-15 years (M:5, SD:2.1). They were recruited through a mixture of our social network and recruitment posters around the campus.

The Experiment

Each test subject was taken into the usability lab and given a short introduction to what we were doing and why. We then explained how the system worked and what they had to do. We would hand them a phone, ask them to stand on a marked cross, so that the distance to the screen would always be the same, and start the test.

The application chose at random one of the four techniques and displayed a short explanatory movie of how to perform the technique on a screen right beside the main application display(See Figure 7a).

(a) (b)

Figure 7. (a) The main screen on the left with the tutorial video screen on the right. (b) The experiment in progress, as seen from the right.

The user would then have three practice attempts, in order to get familiar with the technique. Nothing was logged during the practice phase. A shape would appear, either a square or a circle, at one of the cells in the grid. The user would have to choose the correct shape on the phone and perform the technique with that shape selected. The shapes on the phone (fig. 6b) would randomly change positions, so that the user would have to check the phone after every technique. The target would also randomly change size from small to large or vice-versa. In reality, the target sequences where hard coded by us in such a way that there was an equal distribution of small and large targets. We also made sure that there was an equal distribution of distances between each target. We classified them as short jumps, medium jumps, and long jumps. A short jump was 2 large cells(4 small cells), a medium jump was 4 large cells(8 small cells), and a large jump was 6 large cells(12 small cells). After a practice phase of 3 practice targets, a calibration start target would be shown. This is so that we could calculate the distance between all other targets correctly. The user would then go through the rest of the test (18 targets), going through a total of 22 targets.

That means that our experiment had the following list of conditions:

- Technique (4)
- Target size (2)
- Target jump distance (3)
- Repetition (3)

This means that each user had a total of $4 \times 2 \times 3 \times 3 = 72$ targets.

After going through every target for one technique, the user would then be asked to fill in the short questionnaire regarding the technique just tried in terms of how natural it felt based on ease of use measures.

This entire process would be repeated four times in total, once for each technique. After that, we presented them with a short demographics questionnaire, in order to better understand the user. We asked them about their age, height, if they were left or right handed, if they had a smartphone, for how long, if they had any experience with a Kinect, Wii, Playstation Move, or any other similar air gesture based technologies, and how often they used them. Finally, we thanked them for their time. The entire test took on average 15 minutes.

RESULTS

We will now present the results that we achieved through out our experiment and also how they were achieved. We will first present our findings in respect to success rate(based on hit success), then in respect to efficiency(based on time per target), and finally in respect to accuracy(based on distance to target). Finally, we will look at the questionnaires and show the significant findings there in terms of ease of use. We will discuss these results later, in the Discussion section. Each of the four interaction techniques, *Pinch*, *Swipe*, *Throw* and *Tilt* were completed 18 times per participant. For each target size, each technique was performed 477 times.

Success Rate

In this section we will present the results related to the success of hitting a target. We will perform an analysis of the hit success rate and the effect of each technique with respect to target sizes.

The success rate's mean and standard deviation, M(St.D.) for each technique for small(S) and large(L) target sizes can be seen in Table 2 and in Figure 8 the results are shown as a graph.

Hit Success Means

	Pinch (n = 477)	Swipe (n = 477)	Throw (n = 477)	Tilt (n = 477)
S	0.65 (0.48)	0.91 (0.29)	0.83 (0.37)	0.58 (0.49)
L	0.78 (0.41)	0.97 (0.18)	0.94 (0.25)	0.78 (0.41)

Table 2. Mean hit and standard deviation for each technique per target for small(S) and large(L) targets.

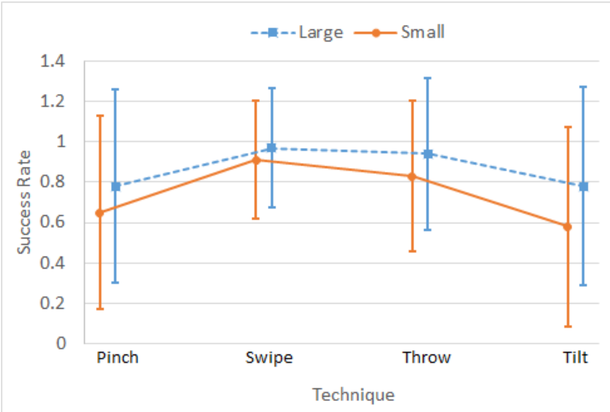


Figure 8. Mean and standard deviation for each technique in regards to hit success rate per target.

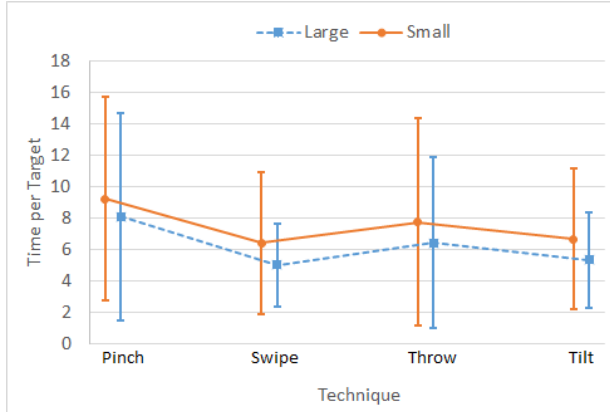


Figure 9. Mean and standard deviation for each technique in regards to time per target.

In order to see the effect of each technique on the hit ratio per target for the different target sizes, we performed two different one-way ANOVA's, where we split the data between the two different target sizes. We then performed a post-hoc pairwise LSD test to see where the significant difference were.

For the the small target, we got ($p < 0.001$), ($F(2.722, 1295.674) = 62.754$), (Greenhouse-Geisser correction: 0.907). The pairwise test showed that all techniques were significantly different. *Pinch* and *Swipe* had ($p < 0.001$), *Pinch* and *Throw* ($p < 0.001$), *Pinch* and *Tilt* ($p = 0.031$), *Swipe* and *Throw* ($p = 0.001$), *Swipe* and *Tilt* ($p < 0.001$) and finally, *Throw* and *Tilt* had ($p < 0.001$).

For the large target, we got ($p < 0.001$), ($F(2.472, 1176.749) = 42.773$), (Greenhouse-Geisser correction: 0.824). The pairwise test showed that all of the techniques, with the exception of *Pinch* and *Tilt*, were statistically different from each other. The results were as following: *Pinch* and *Swipe* ($p < 0.001$), *Pinch* and *Throw* ($p < 0.001$), *Pinch* and *Tilt* ($p = 1.000$), *Swipe* and *Throw* ($p = 0.025$), *Swipe* and *Tilt* ($p < 0.001$), and finally *Throw* and *Tilt* ($p < 0.001$).

Efficiency

In this section we present the efficiency results which defines the amount of time spent performing a technique. We perform an analysis of the efficiency and the effect of each technique with respect to target sizes.

Table 3 shows the mean time per target in seconds for each of the techniques as well as their standard deviation for both target sizes.

Time per Target Means

	Pinch (n = 477)	Swipe (n = 477)	Throw (n = 477)	Tilt (n = 477)
S	9.23 (6.48)	6.41 (4.49)	7.73 (6.60)	6.67 (4.49)
L	8.09 (6.60)	5.01 (2.66)	6.42 (5.43)	5.33 (3.04)

Table 3. Mean time and standard deviation for each technique per target for small(S) and large(L) targets.

To get the effect of each technique on the time for the different target sizes, we performed two one-way ANOVA's, one for the small target and another for the large target. A post-hoc pairwise LSD test to see where the significant difference were.

For the the small target, we got ($p < 0.001$), ($F(2.740, 1304.290) = 26.523$), (Greenhouse-Geisser correction: 0.913). The pairwise test showed that all techniques were significantly different except *Swipe* and *Tilt*. *Pinch* and *Swipe* had ($p < 0.001$), *Pinch* and *Throw* ($p < 0.001$), *Pinch* and *Tilt* ($p < 0.001$), *Swipe* and *Throw* ($p < 0.001$), *Swipe* and *Tilt* ($p = 0.354$) and finally, *Throw* and *Tilt* had ($p = 0.004$).

For the large target, we got ($p < 0.001$), ($F(2.221, 1057.144) = 44.539$), (Greenhouse-Geisser correction: 0.740). The pairwise test showed that all of the techniques, with the exception of *Swipe* and *Tilt*, were statistically different from each other. The results were as following: *Pinch* and *Swipe* ($p < 0.001$), *Pinch* and *Throw* ($p < 0.001$), *Pinch* and *Tilt* ($p < 0.001$), *Swipe* and *Throw* ($p < 0.001$), *Swipe* and *Tilt* ($p = 0.077$), and finally *Throw* and *Tilt* ($p < 0.001$).

Accuracy

Finally, we will perform an analysis of the accuracy and the effect of each technique with respect to target sizes. Here, we took three different measures of accuracy; the distance between where the user hit and the target cell as well as taking the x and y axis independently. These were all measured in pixels. This was because there were signs that certain techniques might miss in a specific direction, and we wanted to see if the data supported that. An overview of the distance mean and standard deviation in pixels can be seen in Table 4 and Figure 10.

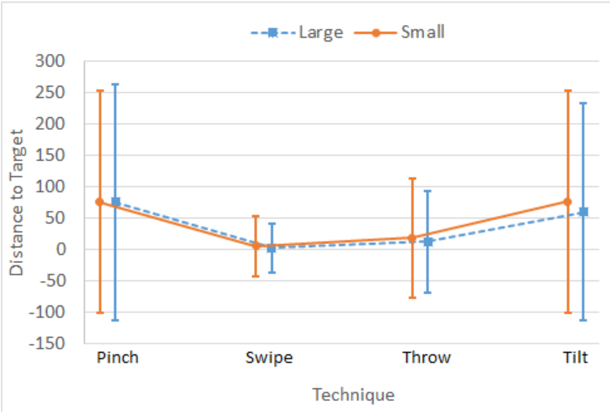


Figure 10. Mean and standard deviation for each technique in regards to distance.

Distance Means

	Pinch (n = 477)	Swipe (n = 477)	Throw (n = 477)	Tilt (n = 477)
S	75.97 (176.15)	5.40 (47.81)	18.60 (95.29)	76.51 (177.45)
L	75.41 (187.58)	2.29 (38.82)	12.37 (81.36)	59.88 (172.22)

Table 4. Mean and standard deviation for the distance for each technique per target for small(S) and large(L) targets.

We performed two one way ANOVA's, one for each target size, on the distance data.

For the the small target, we got ($p < 0.001$), ($F(2.341, 1114.249) = 37.504$), (Greenhouse-Geisser correction: 0.780). The pairwise test showed that all techniques were significantly different except *Pinch* and *Tilt*. *Pinch* and *Swipe* had ($p < 0.001$), *Pinch* and *Throw* ($p < 0.001$), *Pinch* and *Tilt* ($p = 0.961$), *Swipe* and *Throw* ($p = 0.008$), *Swipe* and *Tilt* ($p < 0.001$) and finally, *Throw* and *Tilt* had ($p < 0.001$).

For the large target, we got ($p < 0.001$), ($F(2.176, 1036.004) = 33.315$), (Greenhouse-Geisser correction: 0.725). The pairwise test showed that all of the techniques, with the exception of *Pinch* and *Tilt*, were statistically different from each other. The results were as following: *Pinch* and *Swipe* had ($p < 0.001$), *Pinch* and *Throw* ($p < 0.001$), *Pinch* and *Tilt* ($p = 0.171$), *Swipe* and *Throw* ($p = 0.015$), *Swipe* and *Tilt* ($p < 0.001$) and finally, *Throw* and *Tilt* had ($p < 0.001$).

X and Y Distance Means

	Pinch (n = 477)	Swipe (n = 477)	Throw (n = 477)	Tilt (n = 477)
S XD.	54.33 (140.87)	3.78 (40.15)	10.32 (81.84)	49.23 (151.36)
L XD.	55.32 (159.88)	1.88 (31.74)	10.32 (74.01)	49.23 (157.35)
S YD.	42.90 (110.30)	2.00 (26.16)	8.00 (49.37)	39.3 (99.82)3
L YD.	37.41 (104.18)	1.14 (22.37)	4.33 (34.21)	22.86 (73.75)

Table 5. Mean and standard deviation for distance on the X-Axis(XD) and distance on the Y-Axis(YD) for each technique per target for small(S) and large(L) targets.

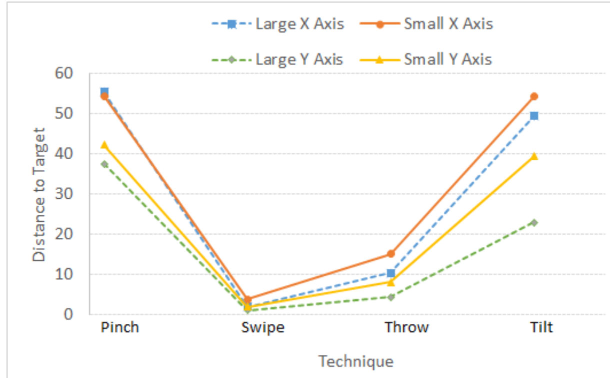


Figure 11. Mean and standard deviation for each technique in regards distance on the x-axis and y-axis.

To get a better understanding of the distance we decided to examine the distance in regards to the x-axis and in regards to the y-axis. The results can be seen in Table 5 and Figure 11

Results of Ease of Use Questionnaire

Our questionnaire was based on USE, which used Likert scale 4.2, when encoding the data we did it as continuous variable, as such *"strongly disagree"* got a value of 1, and *"strongly agree"* a value ot 7. After that the cumulative value per technique, based on the different questions, was calculated, the data was ploted and presented in Figure 12.

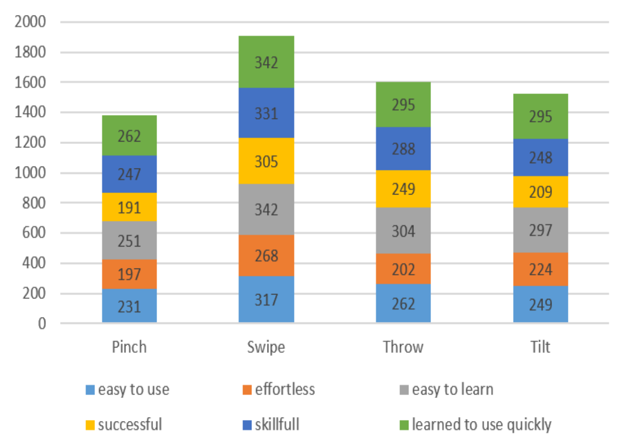


Figure 12. Cumulative values of survey questions per technique

A One-Way MANOVA was applied ($F(18, 574.66) = 5.118$, $p < 0.000$) which showed that there is statistical differences between each techniques. In order to specify where this differences lie we perform an post hoc test, we summarize the results in regards to each question from the survey, as follow: *1)* "It is easy to use" *Swipe* is statistically different, however there is no statistical difference between *Throw*, *Tilt*, and *Pinch*; *2)* "Using it is effortless" *Swipe* is statistically different, however there is no statistical difference between *Throw*, *Tilt*, and *Pinch*; *3)* "It is easy to learn to use" *Swipe* and *Pinch* are statistically different, however there is no statistical difference between *Tilt*, and *Throw*; *4)* "I can use it successfully every time" *Swipe* and *Throw* are statistically different, there is no statistical difference between *Tilt*, and *Pinch*; *5*) "I quickly became skillful with it" *Swipe* and *Throw* are statistically different, however there is no statistical difference between *Pinch*, and *Tilt*; *6*) "I learned how to use it quickly" *Swipe* is statistically different, however there is no statistical difference between *Throw*, *Tilt*, and *Pinch*. The results show that the ratings received by the techniques differ in the different aspects of the areas covered by the survey.

DISCUSSION

When discussing our results, it is important to redefine the terms that we use. When talking about success rate, we mean whether or not the technique hit the given target. When we mention efficiency, we are talking about the time it takes to successfully perform that technique. When talking about accuracy, we mean the distance the attempt was from the target(in pixels). It is also important to note that the standard deviation in some of our measures are quite high. This is primarily because the experimental system was not robust enough to get a perfect reading on the users intention to perform a technique. Sometimes, the system would misunderstand the gesture a user made, so that it either activated too early, or did not activate all. Activation means that the system interpreted the gestures the user performed as an attempt to hit the target. This is further discussed in Section 6.7

When looking at the results on the effect of the four techniques, *Pinch*, *Swipe*, *Throw* and *Tilt*, as well as the two target sizes, small and large, on the time per target, the results tell a rather interesting story.

Success Rate

If we look at the results in regards to the effect of the technique on the success rate of each attempt, it is interesting to note that the two techniques that were not significantly different from each other were *Tilt* and *Pinch*. These two techniques both used the hand that controlled the pointer to activate the technique. When tilting the phone forward, usually the hand would move together with the phone causing the pointer to displace itself from the users intended position. When releasing the *Pinch*, the Kinect would sometimes reevaluate the location of the hand joint, now that it could see the entire hand, which would also cause the pointer to displace itself from the intended position. *Pinch* and *Tilt* were also the techniques that had the largest amount of activation errors due to the implementation of the system. Sometimes, users would show large amount of their palms to the Kinect during a *Pinch*, even though their hand was closed, causing the Kinect to interpret that as an opening of the hand and activate the technique. *Tilt* would sometimes activate if the user moved the mobile around too quickly, especially when orienting the pointer up and down on the screen.

Swipe and *Throw* both had reasonably high success rates. *Throw* did not require the user to actually move the pointer hand while activating the technique. While *Swipe* did require the user to perform some movement on the hand that was used as a pointer, it was very little movement. This is also a technique all smartphone users are familiar with, since a lot of applications use some form swiping to activate some functionality.

Efficiency

There was a significant difference between all techniques, with the exception of *Swipe* and *Tilt*. These were the two one handed techniques that we chose. The range of movement needed in order to activate these two techniques was rather limited, the full motion could be achieved quite quickly and is quite similar for both of them. This is why they are not statistically different from each other. *Swipe* and *Tilt*, are on average, at least a second faster then the other two techniques. Their standard deviations are also smaller, which means that users were more consistent, with regards to how long it took to hit each target, with these two techniques.

Looking at the two other techniques, *Pinch* and *Throw*, their times also reflect the range of motion needed in order to activate each technique. *Pinch* requires the user to pinch the shape on their phone, lift their hand up, direct it on the screen, and then finally let go. This can be seen in its mean, where it takes almost 1.59 seconds longer to perform than the second longest technique, *Throw*. *Throw* also requires a considerable range of motion in order to activate: point with one arm, select the shape on the phone with the other arm, bring your arm back and then finally swing it forward. Both two handed techniques take significantly longer to perform than their one handed counter-parts.

We noticed that users would spent relatively little time getting into the general vicinity of the target, and would spend most of their time per attempt getting the pointer on top of the actual target. This was more pronounced for the small target, were users would perform smaller, more careful adjustments in order to not overshot the target, which can be seen when comparing the mean times of each technique for the two different target sizes.

Accuracy

If we look at the results regarding the distance from the target, it paints quite a clear image of which of the techniques are more accurate. *Swipe* is by far the most accurate of the four techniques. It is so precise, that it is actually more precise for the small target than all other techniques are for the large target. Figure 13 is an image that shows the location of each hit compared to the given target cell.

Figure 13 shows that *Swipe* and *Throw* have a large concentration of hits inside the target cell, where *Pinch* and *Tilt* have quite a spread of hits outside the target cell. This shows us that users are capable of hitting the target with *Swipe* and *Throw* more consistently and accurately than with the other two techniques by considerable amounts.

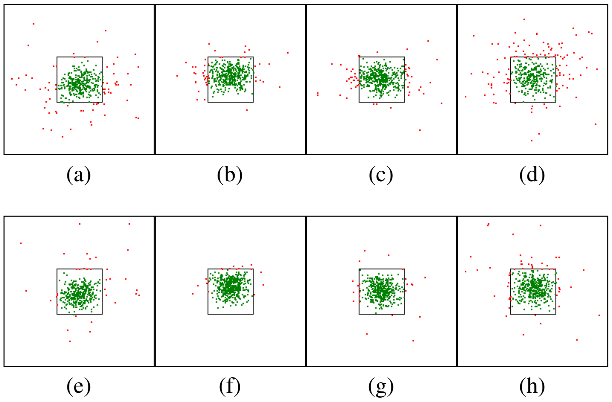


Figure 13. Hit location illustration for each technique. Green is a hit, red is a miss (a) Pinch (Small target) (b) Swipe (Small target) (c) Throw (Small target) (d) Tilt (Small target) (e) Pinch (Large target) (f) Swipe (Large target) (g) Throw (Large target) (h) Tilt (Large target)

Also, an interesting thing to notice is that there is a trend where the largest distance from the target is located on the x axis. This can be seen in Table 5, where the distance is broken into the two different axis. For all four techniques, the x axis has a larger mean distance from the target then the y axis. Further research should be done though if one were to make some conclusive statement in regards to the effect of the different axis on the accuracy of the techniques.

Target Size

If we compare the effect of the target size on the different techniques, it does not have that big of an impact on the success rate of each technique compared to each other. *Swipe* and *Throw* are still the techniques with the highest success rate. Our results also show that target size does not change the efficiency of each technique compared to each other. The same is true for the distance from the target cell. Our results point towards a tendency were the size of the target does not influence the measured effectiveness of each technique relative to each other. This of course would require further research, since a sample size of two target sizes is not enough.

Easy of Use

Looking at the questionnaire, we can look at each question and see there is a trend. If the user gave it a high score, then he strongly agreed with the given question. We can then say the accumulated scores of each technique for all questions show a tendency towards the user agreeing that the given technique was easy to use and hence more natural to use. The higher the score, the easier the user felt the given technique was to perform. This means that in general, users considered *Swipe* much easier to use then the other techniques. *Throw* and *Tilt* were considerably close to each other, while *Pinch* trailed quite significantly behind. It is interesting to note that *Throw* and *Tilt* were so close to each other in ease of use, even though *Throw* outperformed *Tilt* considerably, both in regards to time and hit success, as well as the consistency of the technique, shown by the standard deviation.

User Comments

There is also a qualitative aspect to take into account here, which is not reflected well in the surveys or the test results, but were recorded on video and notes during each experiment. For example, a large finding was that correct mapping between the direction in which the user is pointing and the pointer icon on the screen is critical to the success of the application. It is extremely important for the experience of the user to have as close to absolute mapping as possible. Eleven users mentioned having trouble reaching all areas of the screen, but almost all users showed sign of trouble, by for example standing on their toes or stretching their arms as far as possible. One user got so frustrated that she asked for a chair to stand on. With more or different sensors, placed on hands, fingers and phones, we could have had a more precise pointer by being able to determine the direction of the phone and the arm and not solely rely on the position of the users hands. This would most likely lead to more precise results, because the the mapping between the pointer and the users pointing direction would be much closer to a absolute mapping.

In regards to the mobile phone, four users complained that the screen was too small when performing a *Pinch*, making it hard to precisely select the correct shape. Four other users complained that the 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. Four users mentioned that it was hard to orient themselves with the phone while performing the *Throw* technique, having to break their flow to look down on the phone to select the correct shape. Three users mentioned the same problem with *Swipe*, whenever the targets were too high. This sometimes lead to the mobile phone covering up the target and making it hard for the user to orient themselves to the large display. This was an effect of the relative mapping though, since it was hard to see the screen when their arm was stretched far above their head in order to reach the high targets. The same error could have occurred with *Tilt*, since the user might also end up lifting the phone in front of their field of view, but none of the users mentioned it there.

There is also a the learning aspect of each technique. Six people actually mentioned that the *Pinch* technique was hard to learn, and a large portion of the participants had to be told more specifically how to perform it. The same held true for the *Throw* technique, a large number of the participants had to be told that they had to perform a slightly larger motion in order for the application to understand that a throw motion was being attempted. Very few of the participants had to have further instructions on how the *Swipe* technique worked, and few people needed further help with the *Tilt*. This is most likely a combination of the complexity of some of the techniques as well as the tutorial movies not being descriptive enough. This also lead to frustration, were users thought they were performing the technique correctly and nothing was happening. Four users mentioned being frustrated by the *Pinch* technique, while three users got frustrated with the *Tilt* technique

The fatigue effect is also something to take into consideration. 13 people mentioned being fatigued through out the test and 11 of them first mentioned it during the *Throw* technique. Some users commented that it was because one arm had nothing to do but be uplifted and point to the screen, while the other arm performed all the motion. One user mentioned it would not have been that noticeable if the pointing arm had some motion to perform. We have hopefully accounted for this by rotating the order in which the techniques are performed.

Finally, there is also the fun aspect to take into consideration. Nine users actually mentioned having fun while performing the *Pinch* technique. They compared it to casting a spell or causing explosions on the screen. Three other users mentioned that this technique was especially interesting. It is still worth remembering though that *Pinch* was, by large, the hardest technique for users to learn.

Limitations

There are of course some limitations to the system we developed. Firstly, the intention with the *Tilt* technique was that the users would point and tilt with the phone, but because of our implementation, it was possible for users to point with one hand and tilt the phone with the other. The same holds true for the *Swipe* technique, where users were able to point with one hand and swipe with the other.

The way the system detects an open hand is not very robust: sometimes, depending on the profile of the hand, it misreads the users intentions and believes the user opened his hand.

The system also has a very narrow definition of what throwing means. This is something that can be seen when users were told to "throw" the data from the phone to the screen. Some would perform a much larger tilt motion, others would perform a baseball-like pitching motion.

The Kinect also had some problems determining where the different arm joints were. If the elbow joint was directly behind the hand joint from the Kinects perspective, it would cause the pointer to move erratically since the Kinect was not absolutely sure were the hand joint was. Another problem occurred when the user put their two hands close to each other. The Kinect would have problems determining where the individual hand joints were located.

CONCLUSION

We conducted a study on cross-device interaction techniques, where our focus was on pushing data from a mobile device onto a large display. We compared four different techniques (*Pinch*, *Swipe*, *Throw* and *Tilt*) with two different target sizes. Our main concern was investigating the success rate, efficiency, accuracy and ease of use of the user while using the different interaction techniques.

Our findings show that *Swipe* was the technique with the highest success rate, and at the same time, the most efficient and accurate technique. This was also the technique the users felt easiest to use. We also found that *Pinch* was considered to be a fun and entertaining technique by many of our users. Finally, we also found that the mapping between the screen and the users pointing direction is critical to the applications success. Great care should be taken to achieve as close to an absolute mapping as possible.

FUTURE RESEARCH

Based on our experiment, in which we used two target sizes, we can not determine the difference in effect between targets of different sizes on a large display. As such we would suggest more research including a greater varity of target sizes.

Whereas our experiment only consider push techniques, we would strongly suggest looking at pull techniques for cross-device natural user interaction with large displays. With pull techniques we imagine research investigating the opposite direction i.e. pulling information from a large display to a mobile device. The techniques which may be preferred for pushing information to a large display might not be the best choice for pulling information.

Our experiment has exclusively been concerned with specific measures like success rate, efficiency and accuracy for each technique, and we could suggest for future research that other measures be included in experiment. This includes, but is not limited to, measures on user experience and which techniques users prefer to use for interaction with displays located in public places.

We focused only on the interaction between mobile phones and large displays. In the future though, the range of different personal devices will probably be much more widespread than today, and we suggest further research in this area. An example of this research is the interaction between large displays and devices such as smart watches, tables, and smart glasses.

As a final suggestion, a framework for cross-device natural user interactions might help, for example, developers and researchers with techniques, guidelines and designs for interacting with large displays in the future.

ACKNOWLEDGMENTS

We would like to thank Jeni Paay for the excellent guidance we received throughout the entire project. We would like to thank Dimitrios Raptis for helping us set up the experiment as well as understanding all the data we got out of it. We would also like to thanks Jesper Kjeldskov for making sure we had all the resources we needed. Finally, we would like to thank all the test participants in our experiment.

REFERENCES

1. Hrvoje Benko and Andrew D. Wilson. 2010. Multi-point Interactions with Immersive Omnidirectional Visualizations in a Dome. In *ACM International Conference on Interactive Tabletops and Surfaces (ITS '10)*. ACM, New York, NY, USA, 19–28. DOI: http://dx.doi.org/10.1145/1936652.1936657

2. Sebastian Boring, Manuela Altendorfer, Gregor Broll, Otmar Hilliges, and Andreas Butz. 2007. Shoot & Copy: Phonecam-based Information Transfer from Public Displays Onto Mobile Phones. In *Proceedings of the 4th International Conference on Mobile Technology, Applications, and Systems and the 1st International Symposium on Computer Human Interaction in Mobile Technology (Mobility '07)*. ACM, New York, NY, USA, 24–31. DOI:http://dx.doi.org/10.1145/1378063.1378068

3. Sebastian Boring, Marko Jurmu, and Andreas Butz. 2009. Scroll, Tilt or Move It: Using Mobile Phones to Continuously Control Pointers on Large Public Displays. In *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7 (OZCHI '09)*. ACM, New York, NY, USA, 161–168. DOI: http://dx.doi.org/10.1145/1738826.1738853

4. Andrew Bragdon, Rob DeLine, Ken Hinckley, and Meredith Ringel Morris. 2011. Code Space: Touch + Air Gesture Hybrid Interactions for Supporting Developer Meetings. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces (ITS '11)*. ACM, New York, NY, USA, 212–221. DOI: http://dx.doi.org/10.1145/2076354.2076393

5. Xiang 'Anthony' Chen, Tovi Grossman, Daniel J. Wigdor, and George Fitzmaurice. 2014. Duet: Exploring Joint Interactions on a Smart Phone and a Smart Watch. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 159–168. DOI: http://dx.doi.org/10.1145/2556288.2556955

6. Larry Francis. 1997. Mobile Computing: A Fact in Your Future. In *Proceedings of the 15th Annual International Conference on Computer Documentation (SIGDOC '97)*. ACM, New York, NY, USA, 63–67. DOI: http://dx.doi.org/10.1145/263367.263378

7. Peter Hamilton and Daniel J. Wigdor. 2014. Conductor: Enabling and Understanding Cross-device Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 2773–2782. DOI: http://dx.doi.org/10.1145/2556288.2557170

8. Kaori Ikematsu and Itiro Siio. 2015. Memory Stones: An Intuitive Information Transfer Technique Between Multi-touch Computers. In *Proceedings of the 16th International Workshop on Mobile Computing Systems and Applications (HotMobile '15)*. ACM, New York, NY, USA, 3–8. DOI: http://dx.doi.org/10.1145/2699343.2699352

9. Andrés Lucero, Jussi Holopainen, and Tero Jokela. 2012. MobiComics: Collaborative Use of Mobile Phones and Large Displays for Public Expression. In *Proceedings of the 14th International Conference on Human-computer Interaction with Mobile Devices and Services (MobileHCI '12)*. ACM, New York, NY, USA, 383–392. DOI:http://dx.doi.org/10.1145/2371574.2371634

10. AM Lund. 2001. Measuring Usability with the USE Questionnaire. STC Usability SIG Newsletter. (2001).

11. Nicolai Marquardt. 2011. Proxemic Interactions in Ubiquitous Computing Ecologies. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems (CHI EA '11)*. ACM, New York, NY, USA, 1033–1036. DOI: http://dx.doi.org/10.1145/1979742.1979691

12. Nicolai Marquardt, Ken Hinckley, and Saul Greenberg. 2012. Cross-device Interaction via Micro-mobility and F-formations. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology (UIST '12)*. ACM, New York, NY, USA, 13–22. DOI: http://dx.doi.org/10.1145/2380116.2380121

13. Brad A. Myers. 2001. Using Handhelds and PCs Together. *Commun. ACM* 44, 11 (Nov. 2001), 34–41. DOI:http://dx.doi.org/10.1145/384150.384159

14. Heidi Selmer Nielsen, Marius Pallisgaard Olsen, Mikael B. Skov, and Jesper Kjeldskov. 2014. JuxtaPinch: Exploring Multi-device Interaction in Collocated Photo Sharing. In *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services (MobileHCI '14)*. ACM, New York, NY, USA, 183–192. DOI: http://dx.doi.org/10.1145/2628363.2628369

15. Jun Rekimoto. 1997. Pick-and-drop: A Direct Manipulation Technique for Multiple Computer Environments. In *Proceedings of the 10th Annual ACM Symposium on User Interface Software and Technology (UIST '97)*. ACM, New York, NY, USA, 31–39. DOI: http://dx.doi.org/10.1145/263407.263505

16. Jun Rekimoto. 1998. Multiple-Computer User Interfaces: A Cooperative Environment Consisting of Multiple Digital Devices. In *Proceedings of the First International Workshop on Cooperative Buildings, Integrating Information, Organization, and Architecture (CoBuild '98)*. Springer-Verlag, London, UK, UK, 33–40. http://dl.acm.org/citation.cfm?id=645968.674728

17. Jürgen Scheible, Timo Ojala, and Paul Coulton. 2008. MobiToss: A Novel Gesture Based Interface for Creating and Sharing Mobile Multimedia Art on Large Public Displays. In *Proceedings of the 16th ACM International Conference on Multimedia (MM '08)*. ACM, New York, NY, USA, 957–960. DOI: http://dx.doi.org/10.1145/1459359.1459532

18. Dominik Schmidt, Julian Seifert, Enrico Rukzio, and Hans Gellersen. 2012. A Cross-device Interaction Style for Mobiles and Surfaces. In *Proceedings of the Designing Interactive Systems Conference (DIS '12)*. ACM, New York, NY, USA, 318–327. DOI: http://dx.doi.org/10.1145/2317956.2318005

19. Julian Seifert, Adalberto Simeone, Dominik Schmidt, Paul Holleis, Christian Reinartz, Matthias Wagner, Hans Gellersen, and Enrico Rukzio. 2012. MobiSurf: Improving Co-located Collaboration Through Integrating Mobile Devices and Interactive Surfaces. In *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces (ITS '12)*. ACM, New York, NY, USA, 51–60. DOI: http://dx.doi.org/10.1145/2396636.2396644

20. Mikael B. Skov, Jesper Kjeldskov, Jeni Paay, Heidi P. Jensen, and Marius P. Olsen. 2015. Investigating Cross-Device Interaction Techniques: A Case of Card Playing on Handhelds and Tablets. In *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction (OzCHI '15)*. ACM, New York, NY, USA, 446–454. DOI: http://dx.doi.org/10.1145/2838739.2838763

21. Robert Walter, Gilles Bailly, Nina Valkanova, and Jörg Müller. 2014. Cuenesics: Using Mid-air Gestures to Select Items on Interactive Public Displays. In *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services (MobileHCI '14)*. ACM, New York, NY, USA, 299–308. DOI: http://dx.doi.org/10.1145/2628363.2628368

Appendix B: Participant Questionnaire

Technique: ☐ Grab ☐ Swipe ☐ Throw ☐ Tilt   ID:

---

1. I _______ recommend this interaction technique to others.

would not ☐ ☐ ☐ ☐ ☐ ☐ ☐ would

2. This form of interaction technique is _________ for putting content to a public display.

not useful ☐ ☐ ☐ ☐ ☐ ☐ ☐ useful

3. The interaction technique feels __________.

unnatural ☐ ☐ ☐ ☐ ☐ ☐ ☐ natural

4. I feel ________ about performing the interaction technique.

doubtful ☐ ☐ ☐ ☐ ☐ ☐ ☐ confident

5. This interaction technique is ___________.

confusing ☐ ☐ ☐ ☐ ☐ ☐ ☐ clear

6. This form of interaction technique is ___________ .

unpleasant ☐ ☐ ☐ ☐ ☐ ☐ ☐ pleasant

7. This form of interaction technique is _________ .

slow ☐ ☐ ☐ ☐ ☐ ☐ ☐ fast

8. This form of interaction is _________.

ordinary ☐ ☐ ☐ ☐ ☐ ☐ ☐ novel

9. Performing the interaction is ___________.

boring ☐ ☐ ☐ ☐ ☐ ☐ ☐ fun

10. Performing the interaction is __________.

difficult ☐ ☐ ☐ ☐ ☐ ☐ ☐ easy

11. The interaction requires ___________ effort.

too much ☐ ☐ ☐ ☐ ☐ ☐ ☐ too little

12. The content _________ where I placed it (wanted it) on the screen.

didn't went ☐ ☐ ☐ ☐ ☐ ☐ ☐ went

Appendix B: Observer Questionnaire

ID:

Age: ________ Height: ________ Email: ________________

Gender: ☐ Male ☐ Female

Left or right handed: ☐ Left ☐ Right

Smartphone user: ☐ Yes ☐ No How long have you had one?: ___________ years

Have you ever played any motion sensor based games such Wii games, Kinect games, or PlayStation Move games?

☐ Yes ☐ No

By looking at the user of the system, do you see something embarrassing?
(Draw/Describe):

By looking at the system, does it seem easy to learn how to use it?

(If yes, describe why.):

By looking at placement of the system, does it look accessible?

By looking at the system, do you feel like you want to try it yourself?

I do not agree ☐ ☐ ☐ ☐ ☐ ☐ ☐ I agree

Appendix B: Experience Questionnaires

ID:

What technique did you like the most and why?

What technique did you dislike the most and why?
(Do you remember the name of it?)

Did you feel pressured by performing in this public area?
(if yes why?)

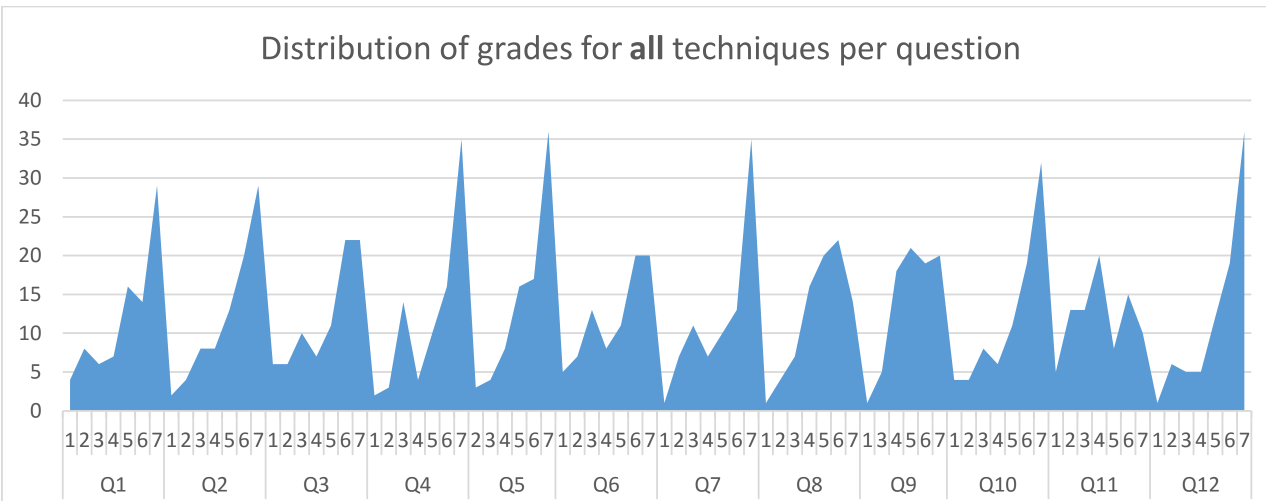
Did you feel embarrassed during these interactions, why?, when?

Other comments?

Appendix C: Summary of Quantitative data

Question	Technique	Mean	Cum. (max 147)	Std	Final
I (wouldn't / would) recommend this interaction technique to others	Grab	5.24	110	2.02	would
	Swipe	6.24	131	1.19	would
	Throw	4.86	102	1.67	would
	Tilt	4.29	90	1.93	wouldn't
This form of interaction technique is (not useful / useful) for putting content to a public displays	Grab	5.48	115	1.71	useful
	Swipe	6.14	129	1.21	useful
	Throw	5.10	107	1.54	useful
	Tilt	4.90	103	1.82	useful
The interaction technique feels (unnatural/natural)	Grab	5.29	111	1.83	natural
	Swipe	5.90	124	1.69	natural
	Throw	4.38	92	1.73	unnatural
	Tilt	4.29	90	1.91	unnatural
I feel (doubtful/confident) about performing the interaction technique	Grab	5.48	115	1.47	confident
	Swipe	6.14	129	1.08	confident
	Throw	5.19	109	2.01	confident
	Tilt	4.95	104	2.03	confident
This interaction technique is (confusing/clear)	Grab	5.43	114	1.62	clear
	Swipe	6.33	133	0.84	clear
	Throw	5.62	118	1.25	clear
	Tilt	5.67	119	1.55	clear
This form of interaction technique is (unpleasant/pleasant)	Grab	5.43	114	1.68	pleasant
	Swipe	5.67	119	1.58	pleasant
	Throw	4.24	89	1.66	unpleasant
	Tilt	3.95	83	2.01	unpleasant
This form of interaction technique is (slow/fast)	Grab	5.05	106	1.81	fast
	Swipe	6.00	126	1.51	fast
	Throw	5.19	109	1.92	fast
	Tilt	5.14	108	1.78	fast
This form of interaction is (ordinary/ novel)	Grab	5.86	123	1.21	novel
	Swipe	4.19	88	1.71	ordinary
	Throw	5.24	110	1.11	novel
	Tilt	4.90	103	1.11	ordinary
Performing the interaction is (boring/fun)	Grab	5.67	119	1.28	fun
	Swipe	5.29	111	1.12	fun
	Throw	5.57	117	1.33	fun
	Tilt	4.76	100	1.31	fun

Performing this interaction is (difficult / easy)	Grab	5.29	111	1.72	easy
	Swipe	6.14	129	1.39	easy
	Throw	5.62	118	1.50	easy
	Tilt	4.52	95	2.11	difficult?
The interaction required (too much/ too little) efforts	Grab	4.24	89	1.82	much
	Swipe	4.57	96	1.50	little
	Throw	3.86	81	1.78	much
	Tilt	4.00	84	1.90	much
The content (didn't went / went) where I placed it [wanted it] on the screen.	Grab	5.67	119	1.64	went
	Swipe	6.24	131	0.92	went
	Throw	5.90	124	1.48	went
	Tilt	4.76	100	1.92	went



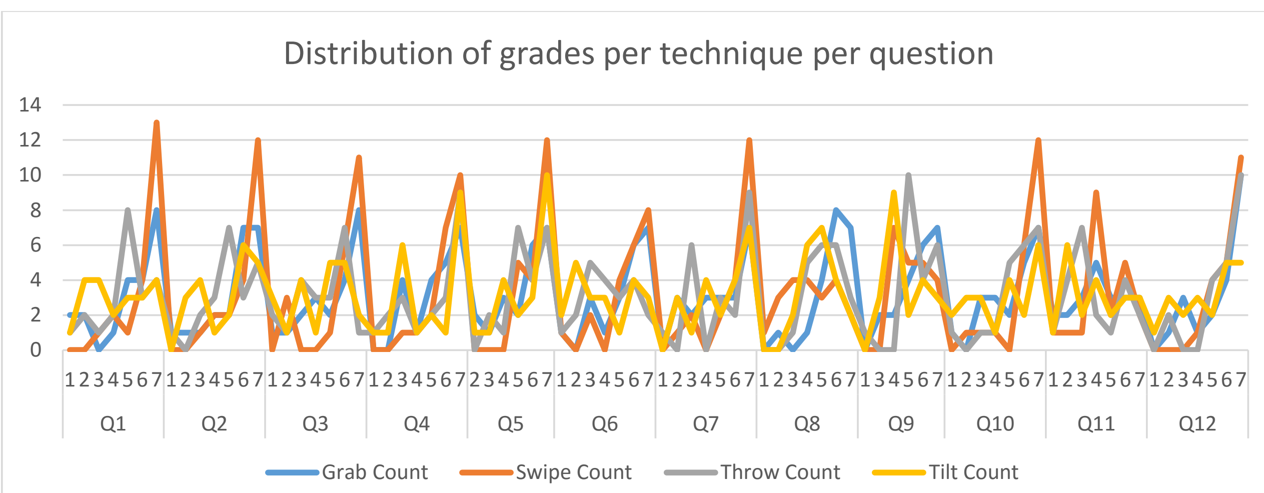


Table with values/grades per technique and in total

		Grab	Swipe	Throw	Tilt	Total
	Value	Count	Count	Count	Count	Count
Q1	1	2	0	1	1	4
	2	2	0	2	4	8
	3	0	1	1	4	6
	4	1	2	2	2	7
	5	4	1	8	3	16
	6	4	4	3	3	14
	7	8	13	4	4	29
Q2	1	1	0	1	0	2
	2	1	0	0	3	4
	3	1	1	2	4	8
	4	2	2	3	1	8
	5	2	2	7	2	13
	6	7	4	3	6	20
	7	7	12	5	5	29
Q3	1	1	0	2	3	6
	2	1	3	1	1	6
	3	2	0	4	4	10
	4	3	0	3	1	7
	5	2	1	3	5	11
	6	4	6	7	5	22
	7	8	11	1	2	22
Q4	1	0	0	1	1	2
	2	0	0	2	1	3
	3	4	1	3	6	14
	4	1	1	1	1	4
	5	4	2	2	2	10
	6	5	7	3	1	16
	7	7	10	9	9	35
Q5	2	2	0	0	1	3
	3	1	0	2	1	4
	4	3	0	1	4	8
	5	2	5	7	2	16
	6	6	4	4	3	17
	7	7	12	7	10	36
Q6	1	1	1	1	2	5
	2	0	0	2	5	7
	3	3	2	5	3	13
	4	1	0	4	3	8
	5	3	4	3	1	11
	6	6	6	4	4	20
	7	7	8	2	3	20
Q7	1	0	0	1	0	1
	2	3	1	0	3	7
	3	2	2	6	1	11
	4	3	0	0	4	7
	5	3	2	3	2	10
	6	3	4	2	4	13
	7	7	12	9	7	35
Q8	1	0	1	0	0	1
	2	1	3	0	0	4
	3	0	4	1	2	7
	4	1	4	5	6	16
	5	4	3	6	7	20
	6	8	4	6	4	22
	7	7	2	3	2	14
Q9	1	0	0	1	0	1
	3	2	0	0	3	5
	4	2	7	0	9	18
	5	4	5	10	2	21
	6	6	5	4	4	19
	7	7	4	6	3	20
Q10	1	1	0	1	2	4
	2	0	1	0	3	4
	3	3	1	1	3	8
	4	3	1	1	1	6
	5	2	0	5	4	11
	6	5	6	6	2	19
	7	7	12	7	6	32
Q11	1	2	1	1	1	5
	2	2	1	4	6	13
	3	3	1	7	2	13
	4	5	9	2	4	20
	5	3	2	1	2	8
	6	3	5	4	3	15
	7	3	2	2	3	10
Q12	1	0	0	0	1	1
	2	1	0	2	3	6
	3	3	0	0	2	5
	4	1	1	0	3	5
	5	2	4	4	2	12
	6	4	5	5	5	19
	7	10	11	10	5	36

	Correlation Matrix(a)												
		Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12
Correlation	Q1	1.00	0.84	0.75	0.66	0.60	0.74	0.68	0.10	0.62	0.78	0.68	0.60
	Q2	0.84	1.00	0.75	0.71	0.61	0.74	0.65	0.10	0.60	0.78	0.58	0.59
	Q3	0.75	0.75	1.00	0.66	0.54	0.82	0.56	-0.03	0.51	0.72	0.63	0.49
	Q4	0.66	0.71	0.66	1.00	0.71	0.64	0.61	0.19	0.47	0.76	0.57	0.59
	Q5	0.60	0.61	0.54	0.71	1.00	0.54	0.61	0.06	0.42	0.65	0.55	0.53
	Q6	0.74	0.74	0.82	0.64	0.54	1.00	0.58	0.16	0.65	0.72	0.68	0.59
	Q7	0.68	0.65	0.56	0.61	0.61	0.58	1.00	0.24	0.54	0.63	0.65	0.47
	Q8	0.10	0.10	-0.03	0.19	0.06	0.16	0.24	1.00	0.35	0.16	0.14	0.17
	Q9	0.62	0.60	0.51	0.47	0.42	0.65	0.54	0.35	1.00	0.62	0.59	0.47
	Q10	0.78	0.78	0.72	0.76	0.65	0.72	0.63	0.16	0.62	1.00	0.67	0.75
	Q11	0.68	0.58	0.63	0.57	0.55	0.68	0.65	0.14	0.59	0.67	1.00	0.47
	Q12	0.60	0.59	0.49	0.59	0.53	0.59	0.47	0.17	0.47	0.75	0.47	1.00
Sig. (1-tailed)	Q1		0	0	0	0	0	0	0.19	0	0	0	0
	Q2	0		0	0	0	0	0	0.19	0	0	0	0
	Q3	0	0		0	0	0	0	0.41	0	0	0	0
	Q4	0	0	0		0	0	0	0.04	0	0	0	0
	Q5	0	0	0	0		0	0	0.30	0	0	0	0
	Q6	0	0	0	0	0		0	0.07	0	0	0	0
	Q7	0	0	0	0	0	0		0.02	0	0	0	0
	Q8	0.19	0.19	0.41	0.04	0.30	0.07	0.02		0.00	0.07	0.10	0.06
	Q9	0	0	0	0	0	0	0	0.00		0	0	0
	Q10	0	0	0	0	0	0	0	0.07	0		0	0
	Q11	0	0	0	0	0	0	0	0.10	0	0		0
	Q12	0	0	0	0	0	0	0	0.06	0	0	0	

Reliability Statistics

Cronbach's Alpha	N of Items
.938	12

Appendix D: Demographic data

Age	Heght	Gender	Hand	Smartphone	SmartphoneTime	Motion_games
21	183	Male	Right	0	0	1
21	175	Male	Right	1	6	1
26	173	Male	Right	1	8	1
27	187	Male	Right	1	5	1
26	173	Male	Right	1	2	1
25	160	Female	Right	1	7	1
34	178	Male	Right	1	10	1
26	181	Male	Right	1	7	1
23	168	Female	Right	1	6	1
26	179	Male	Right	1	2	1
25	162	Female	Right	1	6	1
21	182	Male	Right	1	6	1
21	183	Male	Right	1	11	1
23	180	Male	Right	1	6	1
25	178	Male	Right	1	5	0
23	195	Male	Right	1	6	1
23	187	Male	Right	1	8	1
26	176	Male	Right	1	10	1
23	165	Male	Right	1	12	1
32	180	Male	Left	1	5	1
22	186	Male	Right	1	5	1
21	179	Male	Right	1	6	0
55	157	Female	Right	1	7	0
40	169	Female	Right	1	5	1

Age	Heght	Gender	Hand	Smartphone	SmartphoneTime	Motion_games
21	183	Male	Right	0	0	1
21	175	Male	Right	1	6	1
26	173	Male	Right	1	8	1
27	187	Male	Right	1	5	1
26	173	Male	Right	1	2	1
25	160	Female	Right	1	7	1
34	178	Male	Right	1	10	1
26	181	Male	Right	1	7	1
23	168	Female	Right	1	6	1
26	179	Male	Right	1	2	1
25	162	Female	Right	1	6	1
21	182	Male	Right	1	6	1
21	183	Male	Right	1	11	1
23	180	Male	Right	1	6	1
25	178	Male	Right	1	5	0
23	195	Male	Right	1	6	1
23	187	Male	Right	1	8	1
26	176	Male	Right	1	10	1
23	165	Male	Right	1	12	1
32	180	Male	Left	1	5	1
22	186	Male	Right	1	5	1
21	179	Male	Right	1	6	0
55	157	Female	Right	1	7	0
40	169	Female	Right	1	5	1

Appendix E: Example of Experiment Transcript

Day 4, 0. MTS
Begins at 00:00

(V) Participant fools around with the kinect, trying different hand motions.

Throw Gesture

02:23
(V) Participant watches the instructions.

02:47
(P) Should I move it up here? (his hand/pointing finger)
(Q) Yes, it's running now.
(V) Participant performs successful throw gestures.

03:00
(Q) I'll say you got that pretty quickly.
(P) <Inaudible>
(Q) Yeah.
(Q) Do you feel this technique is fairly intuitive?
(P) <Inaudible> seems a little weird, but I think it actually works pretty well using both hands; one for the gesture itself and the other for pointing. But I don't quite understand why I should use throw - why can't I just click and make it appear?
(Q) Yes, well...that is a fair question.
(P) It's true, it seems...oh, what's happening?
(Q) Oh, that's because the testing of this technique is done.
(P) Okay.
(Q) But you feel you'd like something easier then the throw gesture?
(P) Yes - I think it (throw) could be fun in a game, but if it's just about putting something on a display, when I need to select on the phone anyway, I've already performed an action and am also already pointing.
(Q) Yeah. So you'd rather just point and select?
(P) Yes, then it's there. I don't see the point of throwing. If I selected on the phone, pointed at the large display and then performed a throw, then it might make sense. But when I start out by pointing and then select on the phone, I've already decided where it (the shape/icon) should go. <the participant demonstrates a light forward push> This could perhaps be used instead.
(Q) I can reveal that you will experience that later. <smile>
(P) I hope I do. <smile>
(Q) So, you think throw is a fun gesture, but you'd like something faster. Do you feel it's accurate enough?

(P) Well, tracking isn’t that accurate, but I find it easy enough to elevate my arm while using the phone, but when looking down for a moment may misplace my selected position on the large display.
(Q) Yeah, okay.
(V) Participant briefly repeats himself.

05:10
(Q) I actually forgot to do this in english. <grins> Oh, well…
(P) Sorry about that.
(Q) No, it’s quite okay.

(V) Questionaire time!

Pinch Gesture

07:20
(Q) Does it distract you? (referring to the live recording)
(P) Maybe a little. <smiles awkwardly and looks into the camera - yes, it will devour you!>
(Q) Maybe a little distracting.
(P) Then it’s gone in a moment, I think...okay.
(Q) Yeah, okay.
(P) So I have to grab the image <performs the pinch technique>
(Q) Yeah.
(P) Then...it has to register and that takes some time. <opens his pinched fingers only at first, then his entire hand>
(Q) Yeah.
(P) Yeah. It didn’t work very well with just two fingers - I have to open my whole hand to register completely. That’s a bit annoying. But maybe that’s er…
(Q) Okay, that went like a false positive.
(P) Yeah, I think so. Um, so...I think...er. Since only, since I’m grabbing with two fingers and releasing with all of my hand, that feels a bit unnatural.
(Q) Okay. So you’d like a more precise detection of your fingers?
(P) Yes.
(Q) How do you feel about the gesture in usage generally? Is it fun or fast or?
(P) I think I feel it’s pretty natural that it’s one...there’s a lot of false positives right now. I think it’s pretty positive that it feels like one motion; that I pick up the image and I place it again. That was the problem with previous one (throw) was that it felt like two distinct things; first, I aim, then I chose and then I threw.
(Q) So you really feel all three sequences where you don’t really, you see two of the sequences being more merged together into one sequence than this one.
(P) Before, it was looks at the screen, look at this one (phone), look at the screen. Now it’s more like just one <makes a throwing motion toward the large display>... that’s better. The finger

tracking is more precise - if it can actually track that I open two fingers, that would be better. But that's probably something with the hardware. I guess.
(Q) There's a limitation there, yes. So, before we proceed I just want to ask you a question. Now, you've been through half of the test and you had onlookers and people are surrounding you - how do you feel about being in a public space during this type of testing and usage?
(P) I think if you were doing like it in a real world usage for for example for using like putting something on a digital board, I think it would probably feel pretty natural. It's not like I'd draw a lot of attention to myself just by pinching and releasing. If it was something like I had to go (jumps up and down and flails his arms above his head) ...then I'd feel more worried about it. But it doesn't draw a lot of attention.
(Q) Okay, so you feel rather comfortable right now?
(P) Yeah, I think so.
(Q) Okay, that's nice.
(V) Participant enthusiastically continues to demonstrate self-invented motions which in fact, but unbeknownst to him, is the swipe gesture. Q smiles broadly.

(V) Questionaire time!

SWIPE Gesture

11:18
(P) Okay, now it's coming. <smile> That one.
(V) Participant proceeds with testing the swipe technique. Seems to get it right away.
(Q) Do you have problem reaching the top of the screen?
(P) Yes. I do have a bit of a problem. It's hard to keep my hand that high. It could probably better work if I had the phone closer to my chest and point the phone instead of my whole arm. But then again, I'm not sure about the technicalities of that, but that would probably be...er.
(Q) That would be a different technique than what we have <inaudible>. So ehm…
(P) This is good. I think this is…er.
(Q) The swipe is much better.
(P) As long as there's not too many...if I have to focus on too many things on the phone like that 3 by 4 small icons and I have to choose one of those and swipe that one, then it would be confusing, but as long as I can find the icon, then look up at the screen and just swipe it - that's good.
(Q) Yeah, so the user interface of the phone for real life uses would be very important for you to be simplified in some way?
(P) Yeah, yeah. I would have to choose the file and then, like you'll have to fulfill the whole screen, so I could just point and swipe.
(Q) Yeah.
(P) If I had to choose and swipe in the same, like in the same movement, then that would be confusing.

(Q) Yeah.
(P) Also, for the demo here, you should allow me to continue swiping things onto the (screen).
(Q) Yeah, okay. After the progress…
(P) Yeah, after the progress is done. <smile>
(Q) <grinning> That's an interesting...eh, yeah.
(P) Because now it feels like nothing is happening. <swipes away, smiling> And that's because nothing *is* happening.
(Q) Yeah, exactly. Well, I already have the next questionaire ready for you, so…

(V) Questionaire time!

TILT Gesture

Begins at 13:55
(V) The participant gets ready for the final test. He reads out the instructions smiling, before making any attempts. What a jolly fellow.

14:15
(P) I think that going to mess up the tracking already before I've even tried it. Let's give it a try. I select, I tilt…(he makes a slight tilting motion, it doesn't work)...doesn't track that very well.
(Q) You might want to change your angle to more upward than - yeah - and then you have to throw a bit hard. Yes, exactly.
(P) Okay, that felt really hard to do, but...it has a learning curve, this.
(Q) Okay.
(P) Okay. I feel it's hard to both keep my hand steady and make this hard motion at the same time. I don't feel very confident that the image will land where I'm trying to throw it, because I feel the bop - the pointer moves when I do this gesture (tilt).
(Q) Yeah. So right now you actually have had a really great error <inaudible> throughout your progress even for...including the other techniques. How would you have it if you for instance were aiming at your target, but you missed the target right to the side, however the image that comes up on the screen were overlapping the target area - would you feel it was too far away or would it be close enough for you? If you were to...um.
(P) I don't quite understand. If I missed, would I feel okay with that or would I be…?
(Q) Yeah, if so if if...um.
(P) <scratches eye/discreetly facepalms> Depends on education. If <inaudible> this, and I tried putting up a document on the screen and I put it like on top of another document, I wouldn't feel very good about that, because then people wouldn't be able to see the other document. You'd have to somehow pick it up and move it again.
(Q) Okay. So, you don't feel that the system's ability to overlap elements is a good way?
(P) It is for some things. For example these pictures where it's just tiger and Marilyn Monroe, but if it's documents that one have to read or notices or something like that, then…
(Q) Yeah, something important. <P repeats>

(P) But if it’s just a collage where you show off all the nice pictures from the party last week, then that would be okay.
(Q) So you actually rather like to have a sort of a...how can we say...more control system that wants you to overlap something that’s rather new and important?
(P) Yeah. I think these interface methods are good, but I just think the whole software has to enable things like moving <inaudible> for example. If all these things were supported then one or two of these interaction methods could be good.
(Q) Yeah. Okay.
(P) For example, you already had the push thing (demonstrates the release from the pinch gesture), so the pull thing makes sense; to pick something up again and then put it down somewhere (on the large display).
(Q) Yeah, okay. So, the ability to move something around is something you’d like?
(P) Yeah, I’d like that.

(V) A bit of talk continues…

17:20
(V) Questionaire time!

Appendix F: Pictures

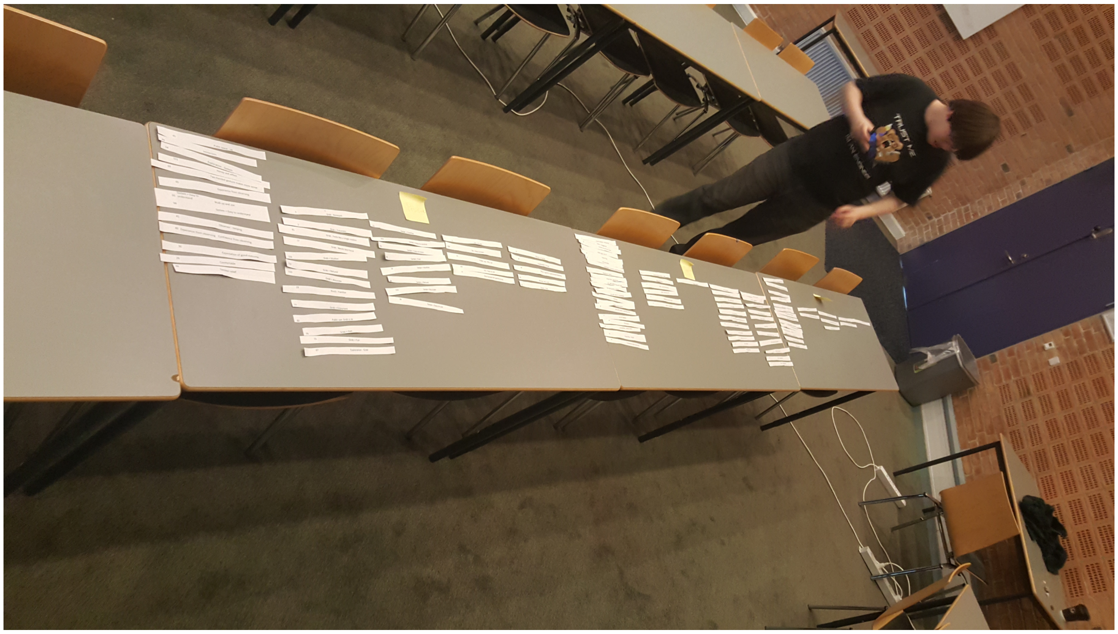
Setup

People showing interest

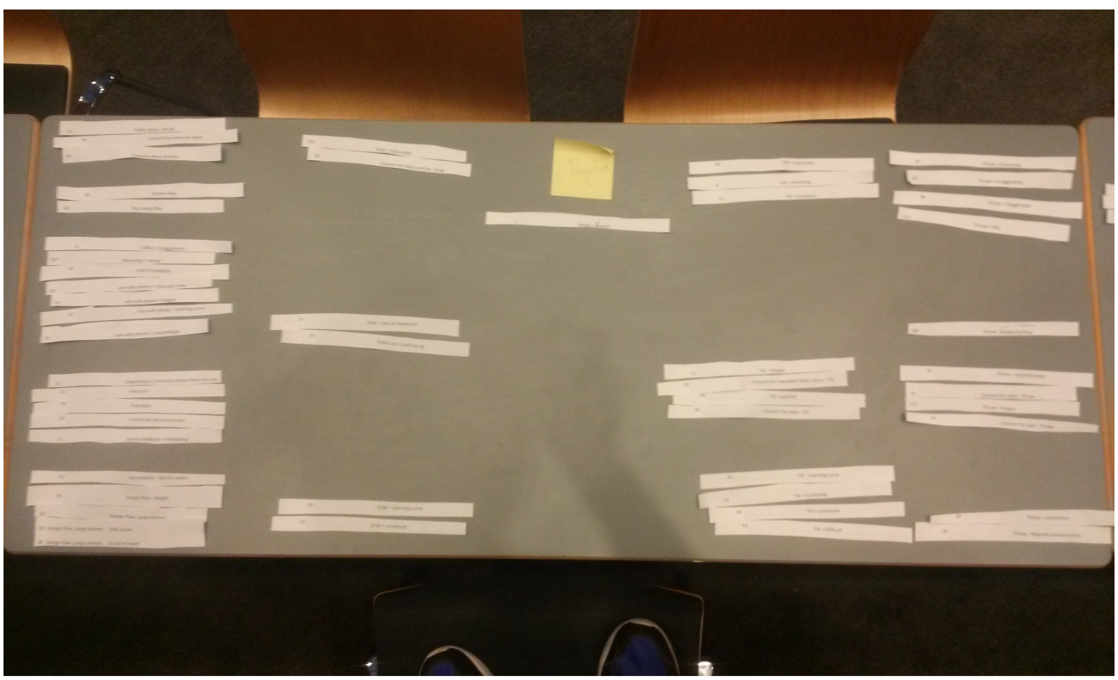
Observers observing

Normal trafic in canteen

Thematic analysis



A single technique



A normal experiment day

A major challenge when creating large interactive displays in a public setting is creating and defining the ways for users to interact with them. This thesis presents the outcomes of user evaluation sessions designed to test a series of different cross-device natural user interaction techniques for people to interact with large displays in public setting. It is an initial step towards the entire perspective, the essence of being able to establish a natural means for immersive interactions with public displays. This thesis proposes four cross-device natural user interaction techniques for transferring data from a mobile device to a public display in a realistic task scenario.

IVAN S. PENCHEV

ELIAS RINGHAUGE

Master's Thesis
Human-Computer Interaction
Aalborg University, 2016