UNIVERSITY OF CALGARY

Examining User Experience in Multi-Display Environments

by

Alemayehu Seyed

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF COMPUTER SCIENCE

CALGARY, ALBERTA

April, 2013

UNIVERSITY OF CALGARY

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled " Examining User Experience in Multi-Display Environments " submitted by Alemayehu Seyed in partial fulfillment of the requirements of the degree of Master of Computer Science.

Supervisor, Dr. Frank Oliver Maurer, Department of Computer Science

---

Co-Supervisor, Dr. Mario Costa Sousa, Department of Computer Science

---

Dr. Anthony Tang, Department of Computer Science

Mr. Ron Murch, Haskayne School of Business

Date

Abstract

With a large number of interactive displays and devices available for users today, multi-display environments are becoming both increasingly common and complex. This complexity also has an effect on a fundamental interaction that users frequently perform in multi-display environments – transferring content. The devices and displays in a multi-display environment – such as digital tabletops, tablets, and high resolution wall displays – now allow users to transfer content in a variety of different combinations.

A review of existing research literature revealed that many of the interactions designed for transferring content in multi-display environments were created by system designers and were not necessarily interactions that users would find usable in real-world multi-display environments. From a user experience perspective, these interactions in multi-display environments require a focus on users, whose real-world experiences and perceptions play a significant role in the interactions themselves. This thesis presents research that identifies better interaction design for multi-display environments. This is accomplished by performing an elicitation study to determine the interactions that users are both comfortable with and prefer for transferring content in a multi-display environment. The result is a set of interaction metaphors and guidelines for user experience professionals to draw upon when creating new gestures and interactions for transferring content in multi-display environments.

Acknowledgements

The research I have performed and this thesis, wasn't done by just me. It is a testament to the support, guidance and help I've received from many people over the last few years. I'd like to thank some of them here.

- To Dr. Frank Maurer – Thank you for giving me the opportunity to join the ASE Lab in undergrad. You have always supported, motivated and inspired me and this mere paragraph can't state enough how appreciative I am to be a part of your lab. I hope, one day, that I'll be able to return all that you have done for me.
- To Dr. Mario Costa Sousa – Thank you for continually supporting me throughout my research and providing numerous opportunities. I truly appreciate the freedom and encouragement you gave me in my MSc. research, as it allowed me to enjoy research and provided inspiration to continue to a PhD.
- To Dr. Anthony Tang – Thank you for the invaluable help and guidance, not just in this thesis work, but in general. Thank you for continually challenging me to think critically about research and HCI. Also, thanks for the chat about crazy research ideas ☺
- To Chris Burns (aka CBC) – Not going to get emotional here, this isn't how we roll. But it's been a long time and quite the ride from 2nd year undergrad to our MSc's. Thanks for all the help and support, but most importantly, the friendship. We've always worked well together and we will continue to do so in the future! Too much to say, not enough space, so nothing but love here brah!
- To Abraam Mikhail (aka Milosh) – Thanks for being a great friend through undergrad and MSc and always being helpful and inspiring to both myself and Chris. Long live Team Awesome! ☺

- To Abhishek Sharma (aka Sharma) and Tulio de Souza Alcantara (aka Senor Tulipa) – Damn guys, it's been a heck of a ride. Thanks for the great environment, the friendship and of course, the jokes, can't forget that! You guys will be titans of industry someday. Some specific shout outs now for you guys. Sharmful – thanks for helping me with my thesis and giving me feedback/guidance, truly appreciate it. Tulipa – Lari is an awesome gal and a wonderful person to be around. You guys are awesome, don't let each other go!! Well Sharma and Tulipa (+ Lari!), it's been a pleasure being friends with you guys, and thank you for all the help, support and shenanigans. And yes, they will continue!
- To Theodore Hellman (aka The Duke) – I truly thank you for seeing the potential in me in my undergrad and providing me the opportunity to join the lab. Without your friendship, help, guidance and feedback – from undergrad through MSc – I wouldn't have had the chance to grow professionally, academically, and personally.
- To Ali Khayat (aka Prince Ali) – Thanks for all of the help in both undergrad and MSc, and taking me under your wing. Without question, without your help and encouragement, I wouldn't feel as confident in my skills today. I can't express enough how much I learned from you and how helpful you were.
- To the ASE Lab Grad Students – Thanks Arlo, Shahbano, Zabed and Rojin for the funny chats/crazy moments, help and cheering me up at times during my MSc.
- To the ASE Lab Undergrads (currently) – Thanks Alaa, Sydney, Edwin, Kevin and Daniel, for being supportive during the writing of this thesis. Without your help, I wouldn't have been able to focus on this thesis and finish. Much like the others who have helped me from undergrad to MSc, I hope to be able to do the same for you guys! ☺

Dedications

I dedicate my MSc. work to my family. You have always been there for me throughout my life and made me into a person who is confident in himself and loves life. Thank you for being the best family in the world and the continued encouragement. I will always strive to make you all proud.

Publications from this Thesis

Portions of the materials and ideas presented in this thesis may have appeared previously in the following peer reviewed publications:

- *MRI Table Kinect: A multi-surface application for exploring volumetric medical imagery*. Teddy Seyed, Chris Burns, Patrick King, Francisco Marinho Rodrigues, Mario Costa Sousa, Frank Maurer. In proceedings of the Workshop on Safer Interaction in Medical Devices (MediCHI'13), Paris, France, 2013.
- *A Usable API for Multi-Surface Systems*. Chris Burns, Teddy Seyed, Theodore D. Hellmann, Mario Costa Sousa, Frank Maurer. In proceedings of the Workshop on Envisioning Future Collaborative Interactive Spaces (BLEND'13). Paris, France, 2013.
- *From Small Screens to Big Displays: Understanding Interaction in Multi-Display Environments*. Teddy Seyed, Chris Burns, Mario Costa Sousa, Frank Maurer. In proceedings of the ACM International Conference on Intelligent User Interfaces (IUI2013), Santa Monica, California, USA, 2013.
- *Eliciting Usable Gestures for Multi-Display Environments*. Teddy Seyed, Chris Burns Mario Costa Sousa, Frank Maurer, Anthony Tang. In proceedings of the ACM International Conference on Interactive Tabletops and Surfaces (ITS2012), Cambridge, Massachusetts, USA, 2012.
- *Towards A Usable API for Constructing Interactive Multi-Surface Systems*. Chris Burns, Teddy Seyed, Theodore D. Hellmann, Jennifer Ferreira, Frank Maurer. In proceedings of the Workshop on Infrastructure and Design Challenges of Coupled Display Visual Interfaces (PPD 2012), Capri, Italy, 2012.

Table of Contents

List of Figures

Chapter One: Introduction

"The ultimate job of design is to transform man's environment and tools and, by extension, man himself"—Victor Papanek [1]

In the traditional single or multi-monitor computing environments, how users interacted with their information (i.e.via keyboard and mouse) typically did not change and neither did the environment itself. For example, for a user switching between word processing applications at home and office environments, environmental changes had minimal impact as they didn't have to worry about changes in the devices they were using or how they were interacting with their information. For user experience professionals, understanding both how users interact with their information and the context of how they interact with it, are important in creating software systems and interfaces that users are comfortable using and understand.

More recently, these environments have advanced rapidly to move from these simple multi-monitor desktops environments to environments that now contain newer technologies such as smartphones, tablets, digital tabletops and wall displays. An environment "where interaction spans multiple input and output devices and can be performed by several users simultaneously" [2] can be called a multi-display environment, and encompasses many of these newer technologies. The interactions that users perform (e.g. touch, mouse and keyboard, and physical movements) with their content in multi-display environments highlights the notion that the environment can change, unlike the simple-multi-monitor environments. Returning to the word processing example, in a multi-display environment (in either home or office settings) a user now has several choices to make, such as where to do their word processing, what device to use – such as a tablet, laptop, or a laptop attached to a large display – and what type of word processing tasks they'd like to accomplish – reading, editing or document creation. Figure 1 highlights this example, showing the choices in devices between multi-monitor desktop environments and a multi-display environment that now includes a tablet, mobile phone and digital tabletop.

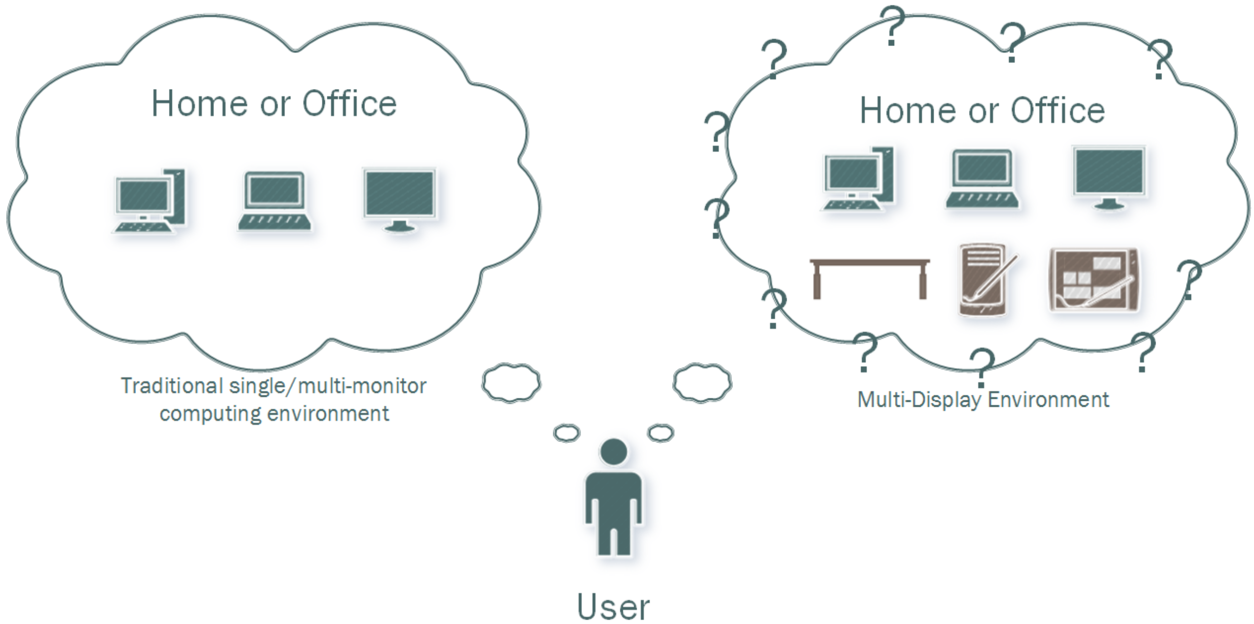

Figure 1 - Comparing the device decisions for a user in single/multi-monitor computing environments and multi-display environments

The number of choices a user is required to make in a multi-display environment and the impact the choice has on the user and their interaction with content, affects the usability of multi-display environments, as tasks are now directly correlated to types of displays and devices. This is of particular interest for user experience professionals and system designers. The research discussed in this thesis explores interactions with content for multi-display environments from the user experience perspective.

This chapter provides an introduction for this thesis. Section 1.1 provides a brief overview into multi-display environments, while Section 1.2 discusses gestures and interactions, and Section 1.3 discusses and usability and user experience. The motivation behind this thesis is then discussed in 1.4 and serves as the basis for the research questions in Section 1.5. Section 1.6 discusses the goals of this thesis. The contributions of this thesis are then detailed in Section 1.7, which is followed by an overview of the structure of this thesis in Section 1.8.

1.1 Multi-Display Environments

A multi-display environment is described by Nacenta et al. [3] as an “interactive computer system with two or more displays that are in the same general space (e.g., the same room) and that are related to one another in some way". This definition also includes "systems where multiple displays are connected to a single computer and systems where networked computers link their displays" [3]. In this definition, the displays are defined as mechanisms of output, and not necessarily mechanisms of input as well.

The definition of multi-display environments used for the work in this thesis, is the following:

> *“A spatially-aware environment where interaction is distributed over several different devices – such as digital tabletops, wall displays, tablets and mobile phones. This distribution of interaction means that multiple users in this environment can perform multiple activities with any number or combination of different display components.”*

A distinct difference between this definition and [3] is the focus on devices that provide different forms of input for interaction – such as multi-touch and gestures. Interaction in this definition also implies that the displays and devices in multi-display environments are not being treated as merely a means of output. As stated by Chen et al. [4], this “emphasizes the nature of many of today’s interactive walls, tables, Tablet PCs, desktop displays, laptops and PDAs that often can be interacted upon in addition to be merely the visual display”. Furthermore, interaction is now distributed, meaning applications and content are no longer on a single device, but instead are distributed amongst the devices in the environment. For user experience professionals, this provides a challenge in designing interactions for content transfer and for applications that are consistent across the displays and devices. Factors due to the heterogeneity of the displays and devices in multi-display environments – such as size, position, resolution and mobility – now impact how users interact with their content and perform their tasks. Content and tasks are now no longer

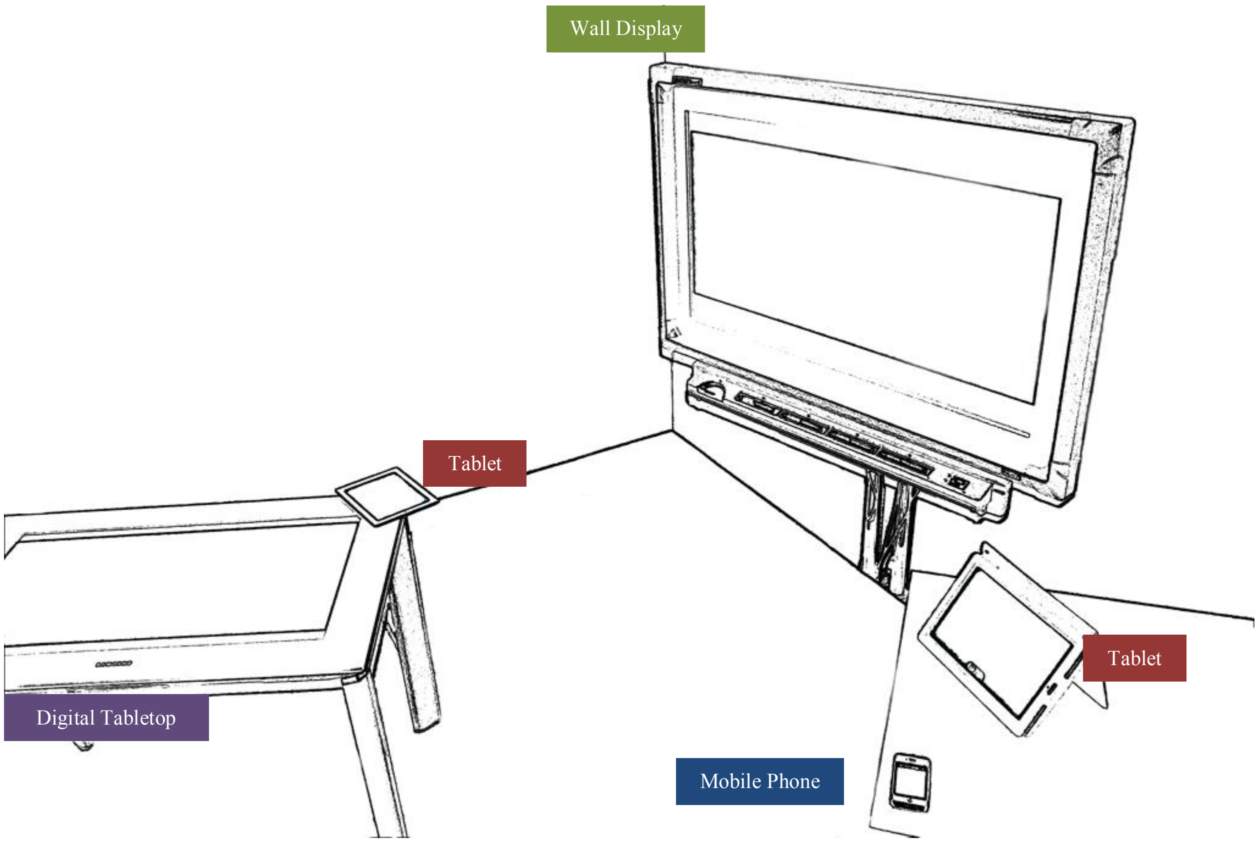

Figure 2 – An example of a heterogeneous multi-display environment

bound to fixed displays, but can be mobile and move in the environment, depending on the display or device used.

Figure 2, highlights a heterogeneous multi-display environment that contains a digital tabletop, wall display, tablets and a mobile phone. Users in this example environment, can freely choose a particular display or device for an appropriate task. Understandably though, not all tasks are suited for all the devices or displays. For example, in Figure 2, a user required to examine large quantities of information (i.e. map data), would most likely choose a larger display such as the digital tabletop or wall display, and not the mobile phone or tablet. This again, highlights the decisions that user experience professionals must make when designing applications and methods of interaction for users in multi-display environments.

A multi-display environment, that takes into account its surroundings, is called a spatially aware multi-display environment. Applying spatial awareness to Figure 2 means that the different displays are aware of each other's properties – such as position in the room or orientation. For example, a mobile phone can be aware of the location of the wall display, and if a user takes the mobile phone and points it at a wall display, the wall display knows that it is being pointed at with a mobile phone. This highlights the notion that how users interact with their content and applications, can now be influenced by the spatial layout of the environment. In this thesis work, multi-display environments are considered to be spatially-aware. This leads to many new methods of interaction in multi-display environments, which are discussed in detail in Chapter 2.

1.2 Gestures and Interaction

A noticeable difference between traditional desktop environments and multi-display environments, is the method by which users can interact with content. The traditional mouse and keyboard interface is now augmented with touch-enabled devices. As a result, depending on the task a user is performing, different interaction methods or combinations of interactions methods can be used.

One way to provide meaningful interactions for users, is to adopt interactions from everyday life instead of creating new ones. This has resulted in methods by which users can interact with computing systems that take advantage of the way they interact with each other and the world itself. These methods include speech [5], the use of objects (tangibles) [6] and gestures [7]. The definition of a gesture and examples, are best provided by Kurtenbach and Hulteen [8] as the following:

> *"A gesture is a motion of the body that contains information. Waving goodbye is a gesture. Pressing a key on a keyboard is not a gesture because the motion of a finger on its way to hitting a key is neither observed nor significant. All that matters is which key was pressed".*

This definition implies that the gesture is the motion of the body and that interaction is the communication of information with the gesture.

Gestures can be performed with or without objects typically. Gestures that don't involve objects include waving, pointing, and summoning. Gestures with objects include pointing at objects, touching or moving objects, changing object shape, activating objects such as controls, or handing objects to others [9]. For multi-display environments, the definition of gestures can change, as tapping on an object is both served and signficant.gestures can be performed with or without the devices or displays, and are discussed in further detail in Chapter 2. How gestures are used serve as the basis for a method of interaction in multi-display environments.

1.3 Usability and User-Experience

The ease of use of a software system or multi-display environment, describes the usability of the system or environment. This means that multi-display environments or software that ignore usability entirely, will be likely to fail, as they either do not meet the needs of the user or they negatively impact the efficiency of a user. Efficiency in the context of usability, refers to how easily users can accomplish their task, how fast the task is accomplished and the likelihood of failure in the task.

Focusing on the user and their tasks, is emphasized in user experience design. User experience, a term first coined by Norman [10], is defined as how a user feels when interacting with a system. For user experience professionals, this means examining factors such as the performance (or efficiency) of the system, its design, how users interact with it, their mental state when using it, as well as its utility, to name a few. Techniques that user experience professionals use to address usability of systems include contextual inquiry [11], prototyping [12] as well as usability testing [13].

One problem for user experience professionals face with these techniques is that it can be difficult to conduct usability examinations when a system is still being designed. This is typically the reason behind the creation of prototypes, as they provide a means of testing different components of a system, either individually or as a whole [14]. In the context of multi-display environments, this is even more problematic for user experience professionals, as the environments are still not well defined, which leaves a number of issues to be addressed. These include, what kind of tasks are suitable for multi-display environments, how should applications, tasks and content be distributed among and across the different displays and devices, and how should users move applications and content across the different displays and devices. In the context of this thesis, moving content across different displays and devices is particularly important, as it serves as the basis of improving the efficiency and usability of multi-display environments for users.

1.4 Motivation

Initial research into multi-display environments highlighted that for many interactions, the conceptual models didn't necessarily match the mental models of users [15]. A conceptual model with respect to interaction in multi-display environments, is the actual method of interaction provided to a user in the environment. A mental model refers to how a user thinks about something – such as a book, a piece of software or a method of interaction. Mental models for users are typically derived from prior experience with similar software, devices or interactions, assumptions they've made and things other users have said. Ultimately, mental models are what users refer to when predicting:

- How they interact with a piece of software, a device or an environment.
- The response of the software, device or environment to their interaction.
- What they should be doing with the software, device or environment.

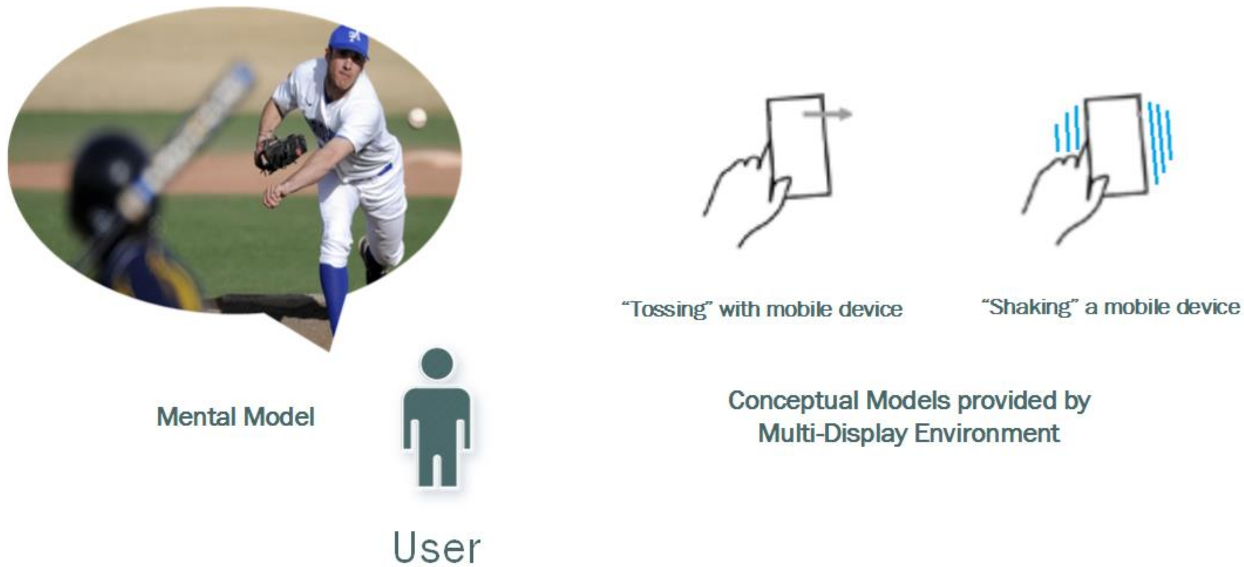

Figure 3 - Comparing Mental Models and Conceptual Models in Multi-Display Environments

Interactions in multi-display environments are most commonly used for tasks that involve the transfer of content. This means sending content to devices, as well as retrieving it. Take for example, the toss interaction. In a multi-display environment, a tossing gesture with a mobile device can be used as a method of sending content to different displays and devices. A user's mental model for tossing may draw upon prior experience of tossing objects, such as a baseball. In Figure 3, two conceptual models are presented to a user for a tossing interaction that sends data to a wall display – one that involves performing a tossing action with a mobile device in one-hand and the other shaking the mobile device. If a user were to draw upon prior experience, they would likely choose the first conceptual model, as the second conceptual model doesn't necessarily map to their mental model and may not make sense for them. This highlights the importance of matching a user's mental model with a conceptual model for content transfer interactions, particularly in a multi-display environment.

For user experience professionals, matching the user's mental model with the conceptual model is vital. This mismatch between conceptual and mental models for methods of interaction can lead to the following challenges in multi-display environments:

1. Users may find interactions difficult to learn or understand, simply because the conceptual model of the interaction doesn't match their mental model.
2. As there are many different types of users, having a conceptual model that only matches one user means that not all users will be able to learn or understand an interaction properly.
3. If interactions are designed as a result of the constraints presented by the multi-display environment (i.e. only digital tabletops exist in the environment) then the conceptual model of the interactions likely doesn't match the user's mental model.
4. If designers of multi-display environments don't consider the mental models being used by users, then it is likely the interactions they devise will be difficult to learn or understand for users.

Focusing on the third and fourth challenges, highlights the problems with the interactions for content transfer seen in the research literature thus far [15]. As the interactions in multi-display environments serve as the basis for transferring content in multi-display environments, an extremely common task for users, the goal of this thesis is to provide a better match for the conceptual models of interactions with the mental models of users. A better match leads to interactions for content transfer tasks that users find easy to understand and easy to use, thus improving the overall usability and efficiency of multi-display environments.

1.5 Research Questions

The first step in providing user experience professionals with interactions for content transfer tasks in multi-display environments, is to understand the existing research space. This provides the first research question for this thesis:

1. *What is the current state of research in multi-display environments, particularly within the context of interaction design?*

Understanding existing multi-display environments will help user experience professionals learn about methods of interaction in multi-display environments, the conceptual and mental models of users and system designers, as well as the what types of multi-display environments exist, if any.

As multi-display environments are considered to be spatially-aware in this thesis, this can have an impact for users and how they interact with content. This leads to the following question:

2. *What are the impacts of the spatial characteristics of the environment on interactions for tasks involving content transfer in multi-display environments?*

This question also leads to another question about interaction that is raised from the heterogeneous nature of multi-display environments, that is:

3. *What are the impacts on content transfer interactions based on the characteristics/affordances of the different displays and devices in a multi-display environment?*

Answers to questions two and three, has an impact on the mental models of users and the conceptual model provided by user experience professionals for transferring content in multi-display environments. This leads into the following question:

4. *To what extent is the conceptual model of interaction for users based upon spatial characteristics or device characteristics?*

The final research question for this thesis is focused on producing a set of interactions for content transfer tasks in multi-display environments. These interactions can aid user experience professionals in the design of future multi-display environments. This is inspired by prior work where a set of interactions were defined for users, and is as follows:

5. *Prior research has produced several gesture sets for digital tabletop interaction, but can gestures be elicited for multi-display environments in a similar fashion?*

These research questions are addressed by findings from a user-elicitation study and are discussed in the remainder of this thesis.

1.6 Research Goals

This thesis has two primary research goals. The first is to provide an extensive overview of the existing research on interactions for multi-display environments. This overview provides a structuring of the multi-display environment space, by collecting a number of different multi-display environments and highlighting their evolution. Also provided in the overview, is a classification of the different interactions for multi-display environments that have been created by system designers of multi-display environments.

The second goal of this thesis is to provide user experience professionals a set of interactions for content transfer tasks, as well as guidelines, based on heterogeneity, for multi-display environments. Chapter 3 discusses a user-elicitation study that addresses this goal. If future work for user experience professionals is to create new content transfer interactions that map to both the mental and conceptual models for users, as well as take into account the different factors that affect them, it should be possible after answering the research questions outlined in the previous section.

1.7 Thesis Contributions

The contributions of the work discussed in this thesis for multi-display environments are as follows:

1. An overview of the current research space of multi-display environments.
2. A set of interaction metaphors that are related to the mental models users have with the conceptual models of interactions for multi-display environments that involve content transfer. These

interaction metaphors can be used by user experience professionals as a tool to guide the creation of new interactions.

3. A set of recommendations for spatially aware multi-display environments. These can be used by user experience professionals for designing content transfer interactions for multi-display environments that take into account the heterogeneous and spatially aware nature of multi-display environments.

1.8 Thesis Structure

This introductory chapter presented a concise background about the research for this thesis. Additionally, the research questions and goals were discussed and thesis contributions outlined. The remaining chapters for this thesis are organized as follows:

Chapter Two: An Overview of Multi-Display Environments - provides a detailed overview of related research into multi-display environments, which includes current approaches for interaction, how to design interactions, supporting users and finally the different types of multi-display environments.

Chapter Three: Exploring Interactions for Multi-Display Environments – describes a study that explores interactions in multi-display environments and attempts to create a user-elicited gesture set for a set of content transfer tasks. Also discussed is the impact of factors such as distance and device type on gestures for users, and recommendations are provided for designing gestures in multi-display environments.

Chapter Four: Conclusion –discusses the conclusion of this thesis research work, including the results, contributions, and future work.

Chapter Two: Background and Related Work

Mark Weiser [16], a research pioneer in the field of computing, envisioned a modern computing environment to be comprised of numerous devices in a multitude of sizes and orientations. This type of computing environment was originally envisioned in his work, “The Computer for the 21st Century" (1991), and is called *Ubiquitous Computing* [17]. Weiser states that machines in these ubiquitous environments "will be interconnected in a ubiquitous network" and that the machines themselves "will also come in different sizes, each suited to a particular task" [17]. The stage for this sort of environment has been set by the gradual increase in the number of devices – such as tablets, mobile phones, and wall displays – in the consumer space, each evolving and being adopted at rapid and, sometimes, independent paces [18].

Within this context of *Ubiquitous Computing* and interconnected devices, Weiser also envisioned these computing environments as "a new way of thinking about computers in the world, one that takes into account the natural human environment” [17]. A quick look at modern video consumption reveals that Weiser's vision is coming closer to reality. Users are able to watch video content between work, mobile and home environments seamlessly, without any regard for device or location. Undoubtedly, with each new tablet, mobile phone, wall display or other new device that becomes interconnected, we come closer to the "natural human environment". However, in order to take into account the "natural human environment", these devices need particular information in order to be physically (or spatially) aware of their surroundings and each other. Information such as the location of other devices, location of people, and the orientation of each device in relation to the other devices, are all important pieces in building a "natural human environment". While much of this information is available individually, such as the location of people (via tracking systems) or the orientation of a device, it is their amalgamation into the "natural human environment" that has not fully occurred yet. This amalgamation is called a multi-display environment and is described in Section 2.1. Section 2.2 describes a particular problem space for interaction design in multi-display environments. Finally, Section 2.3 discusses the different types of multi-display environments.

2.1 Describing Multi-Display Environments

As described in Section 1.1, a multi-display environment is:

> *"A spatially-aware environment where interaction is distributed over several different devices – such as digital tabletops, wall displays, tablets and mobile phones. This distribution of interaction means that multiple users in this environment can perform multiple activities with any number or combination of different display components."*

The notion of a spatially-aware environment in is definition implies that displays and devices are aware of each other's location and other properties (such as orientation). Figure 4 is an example of spatial awareness in multi-display environments, and highlights the impact that spatial awareness has on interaction. In this example, a user is able to *pour* content from a tablet onto a digital tabletop easily. This works because each device is aware of the other and knows their respective orientations. Figure 5 shows another example of spatial awareness in a multi-display environment. Here, the spatial awareness in the environment is provided by a tracking camera which allows users to explore volumetric medical images using a tablet and digital tabletop, much like [19]. This notion of spatial awareness is important for several interactions discussed in further detail in Section 2.2.

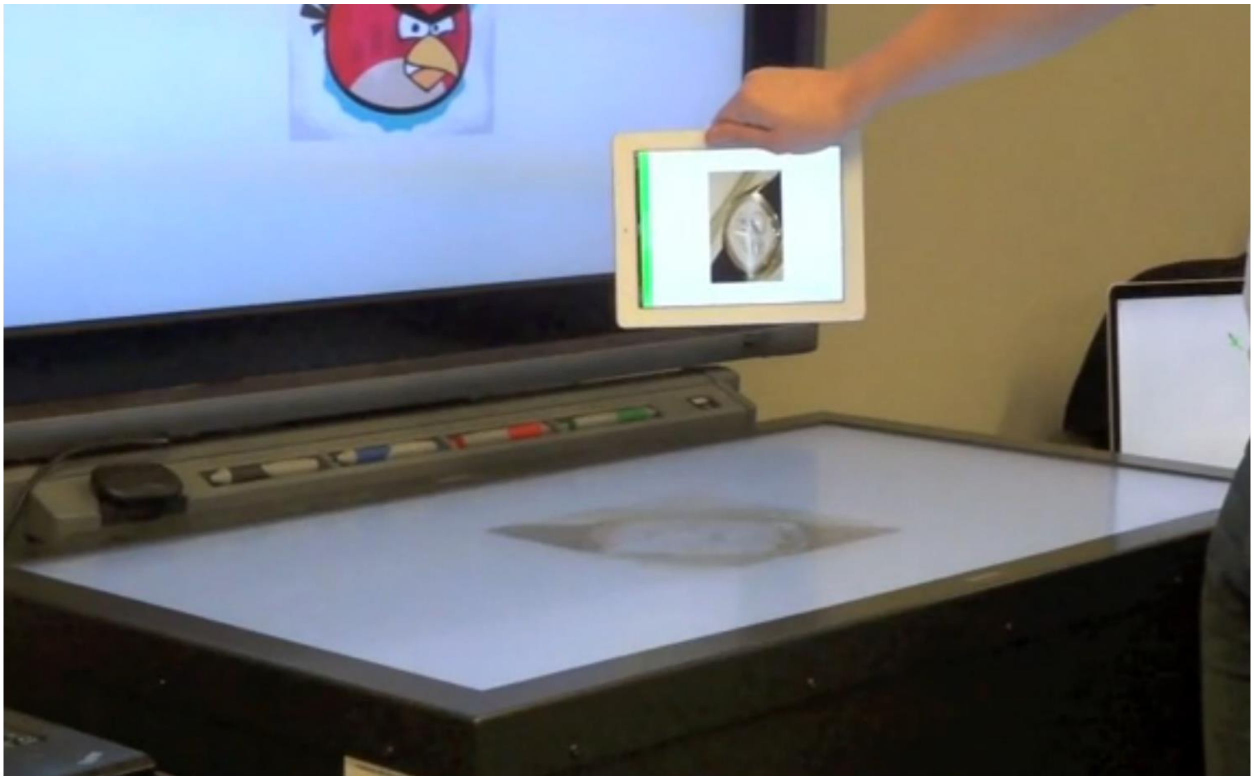

Figure 4 - Example of an interaction that is spatially aware in a multi-display environment [114]

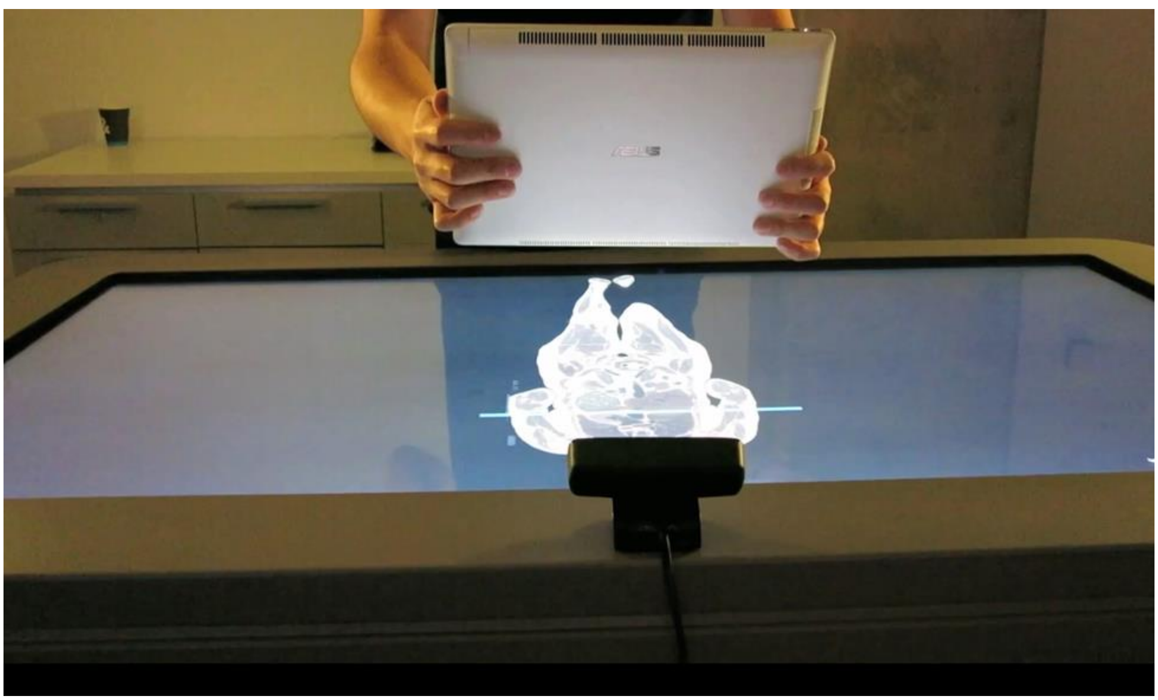

Figure 5 – An example of a multi-display environment

Figure 6 - Cyworld Control Room in Seoul, South Korea [20] – An example of a multi-display environment with workstations and a large wall display

2.1.1 The Evolution of Multi-Display Environments

The Cyworld Control Room [20] in Seoul, Korea (See Figure 6) is an example of an early multi-display environment. In this environment, there are numerous individual workstations (each with a display) and a large wall display that users can interact with through their workstations. The purpose of this multi-display environment is to allow for the monitoring of a virtual game world called Cyworld [20]. At individual workstations (also known as a *workspace)*, users are able to monitor individual user accounts while also being able to monitor various aspects of the virtual world on the large wall display. This early multi-display environment and others in the research literature were found in multi-user *meeting rooms* or *workspaces,* where each user typically had his/her own workstations [21] [22] [23] [24]. Another example of an early prototype multi-display environment with this focus was provided by Tani et al. [25] called

Courtyard, that supported co-operative work by integrating individual workstations with a large *shared* display that could be manipulated via keyboard and mouse input from individual workstations. The large display acted as a *shared workspace* where users could place content onto the wall display individually, to prompt collaborative discussions amongst other users in the environment. Prior to this type of environment, users typically needed to physically move to other individual workstations for collaboration, which wasn't feasible for large groups of users needing collaboration (such as city planning). This concept of a *shared workspace* was also extended to allow techniques such as allowing large displays for overview and smaller displays for details. These techniques augmented the tasks of the individual fixed displays in several research prototypes, as they allowed users to use multiple displays for work, simultaneously. [22] [23] [26].

An important benefit of including multiple heterogeneous screens in workspaces and meeting rooms – displays and large wall displays – is allowing users to divide and organize tasks appropriately and far more efficiently than any single display could, particularly in collaborative tasks [27]. Much of the early research literature and prototypes have focused on how multi-display environments are inherently suited for collaborative tasks, and were typically built around fixed workstations (or displays) and focused on supporting collaboration [21] [22] [23] [24] [25] [28] [29]. A formal description of a multi-display environment according to this early literature and research ([20] - [29]) is as follows:

> **Definition 1:** A multi-display environment is an environment where multiple displays, such as wall displays and computer monitors can be interacted with by users. Users typically interact with these heterogeneous displays via keyboard and mouse input. Additionally, the displays in the environment are typically in fixed, static locations.

An evident problem with the early prototypes of multi-display environments that fit Definition 1 is that many of the interactions users had were tied to fixed displays and restricted their movements [26].

Restricting the mobility of a user to a workstation and limiting their ability to move themselves according to their tasks can be extremely detrimental in these types of multi-display environments [30] [31]. To address this problem, several prototypes were created using cameras (both fixed and moving) to expand the workspace beyond individual workstations [30] [32]. These cameras were used to combine the individual workstations of users in separate locations into one *shared workspace* for all users. The research showed however, that these technologies had a negative effect on how users physically moved in the environment and how they collaborated with other users as they began to move uncomfortably due to the presence of cameras [30] [32]. In essence, mobility can play an important role in multi-display environments when collaboration is important.

Luff [31] addressed this lack of mobility in multi-display environments by suggesting the use of devices with mobile properties in combination with fixed workstations and wall displays. At this point, several different prototypes of multi-display environments began to appear in the research literature that incorporated mobile technologies – such as PDAs, handheld tablets and laptops. Rekimoto [33] [34] used the concept of mobility to introduce a prototype multi-display whiteboard application, where users could use a handheld tablet to draw upon and then later transfer their drawings to a shared wall display. The handheld served as private workspace for users to create drawings and the wall display acted as a public space where users could share their drawings. Koike et al. [35], with *EnhancedTable*, designed a prototype system that integrated laptops, a shared wall display, and a table. Users had distinct workspaces on their laptops and were able to transfer content to either the wall display or a table that could be interacted upon because of enhancements from a projector and a hand tracking system. It is important to note that in *EnhancedTable*, the table was the shared centre of collaboration – despite the mobility of laptops in the environment – and this was mirrored in other prototypes that addressed mobility within multi-display

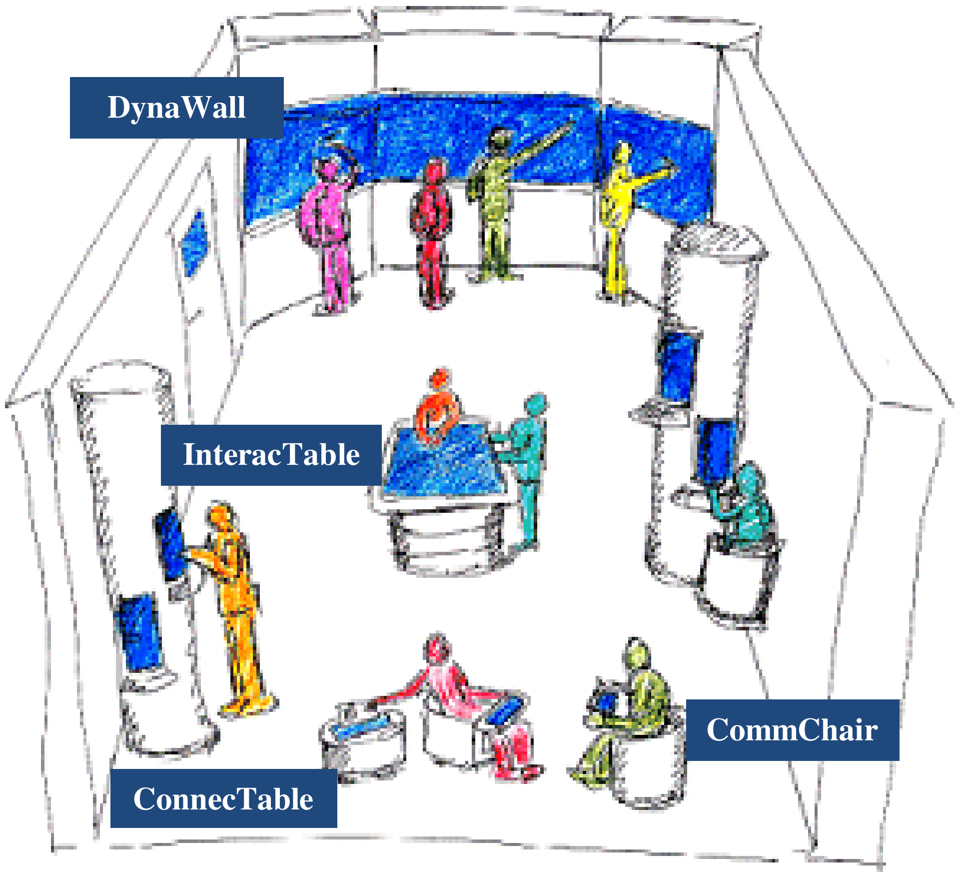

Figure 7 - i-Land from Streitz et al. [6]

environments [36]. This highlights a similarity to the early prototypes in which fixed displays were arranged around a table that indirectly served as a centre of collaboration [21] [22] [23] [24] [25] [28] [29].

Incorporating a digital tabletop into multi-display environments lends credence to the "natural human environment", by Weiser [16] [17], as users have been found to mirror and extend their existing work practices when using them [37]. This means that users use similar practices with digital tabletops as they do with physical tabletops. The *i-Land* project, by Streitz et al. [6] [38], is an early prototype multi-display environment that incorporated a digital tabletop into a multi-display environment (see Figure 7). The once-fixed workstations of early prototype multi-display environments were now integrated into chairs (called

CommChairs) which could be positioned anywhere in the environment and could also interact (through a custom interface) with a large, high resolution wall display (called the DynaWall). The digital tabletop (called InteracTable) was used as a collaboration centre where users were able to share and interact (via touch input) with content. A similar technique was used with the smaller, more portable digital tabletops (called ConnectTables), which could be connected to other surfaces to form a larger, shared digital tabletop. The mobility allowed within *i-Land* and the use of a digital tabletop as a shared space served as an inspiration for several other research prototypes [39] [40] [41] [42] [43] [44] [45] . Additionally, multi-display environment prototypes also began incorporating different methods of input, such as pointing gestures and voice commands, to create more natural environments [6] [38] [46].

The trend to begin using newer and mobile technologies – laptops, PDAs, tablets, and digital tabletops – along with different input modalities – keyboard and mouse, touch input, or gestures – highlights the need to force a change in the definition of a multi-display environment from Definition 1. An updated definition can be formalized as follows (based on [31] - [46]):

> **Definition 2:** A multi-display environment is an environment where multiple displays, some mobile and some fixed – wall displays, laptops, PDAs, tablets, digital tabletops and computer monitors – can be interacted with by users. Users interact with these heterogeneous displays through keyboard and mouse, touch input, gestures or voice commands.

Definition 2 accentuates two important changes from Definition 1: *mobility* and *variety of input*. The displays are no longer in fixed physical locations and the method of interaction is from different sources, such as touch, gestures or voice commands.

Spatial awareness is an important concept in many of the multi-display environments that fall under Definition 2 such as *i-Land* [6] [38], *Co-Mem-iRoom* [47], and other research prototypes [48] [49]. Spatial awareness in the context of multi-display environments refers to how aware the environment is of the location of its users and displays and the effect that the location of displays has on interactions. In *i-Land*, there is no spatial awareness as users can simply move around with displays and no interaction with a display requires knowledge of another display in the environment (see Figure 7 [6] [38]). For example, a user in *i-Land* can move their content from a CommChair to the DynaWall by simply sending it through a graphical menu that is located on the CommChair. This interaction implies that the CommChair doesn't require any spatial awareness of the DynaWall or vice versa. This lack of spatial awareness also means that a user in *i-Land* is unable to take a ConnecTable to the InteracTable or the DynaWall and transfer their content as the interaction is physically bound. As a result, in order to enable spatial awareness in *i-Land*, the displays would need knowledge of the physical location of other displays, which could be accomplished through methods that make the assumption the environment is spatially aware. In several prototype multi-display environments, this is accomplished with interfaces that have graphical representations of the displays (a "world in miniature approach"), menus with available displays in the environment (menu-based approach), or novel methods such as "portals" which allow users to drag content between displays [35] [38] [39] [40] [41] [47] [49] [50] [51] [52]. A more recent example of a multi-display environment that fits Definition 2 is Microsoft's SmartGlass application [53], which allows users to create a connected environment that consists of an Xbox 360 [54] (which is typically connected to a display) and any number of mobile phones or tablets. The tablets and mobile phones can be used to control a display that is connected to the Xbox 360, for example, a user can control content on a tablet from the Xbox 360 or vice versa.

Many examples in the research literature incorporated limited spatial awareness in multi-display environments to achieve implicit knowledge of the location of other displays or devices [36] [55] [56] [57] [58] [59] [60] [61] . Limited in this context means that limited spatial information – such as information from gyroscopes, accelerometers, RFID tags, QR codes – is used to simulate full spatial awareness for the displays in the environment. This means it is not possible for all the devices to know all the information of other devices in the environment. For example, in a multi-display environment that has limited spatial information, it is not possible to know the exact physical locations (or orientations) of other devices at all times. However, despite the limited spatial information in these multi-display environments, it has been shown that spatially aware displays provide natural complements or substitutes to direct interaction with other displays in multi-display environments [56] [58] [59] [62] [63] [64]. For example, Yee [59] used a sensor-equipped PDA to navigate through a custom drawing program, map viewer, and calendar application. The results of a usability study for the spatially aware PDA applications built by Yee found that users preferred this method of navigating digital information and was far more effective compared to the other input methods used in the study, such as keyboard and mouse or touch. This result was also found by Tesoriero et al. [58] who used RFID tags and mobile displays in applications designed for museums and art galleries.

An important result of spatially aware displays is the resulting changes in how users interact with the displays themselves. With a spatially aware display, a user can now use the display itself as a means of interaction. Ballagas et al. [65] [66] highlight numerous interaction techniques with mobile devices controlling content on a large digital display using simulated spatial awareness. These techniques included using the mobile display through different physical input mechanisms to transfer content such as *sweeping*, *pointing*, and *shooting*. Hassan et al. [67] propose a *chucking* interaction that combines touch with accelerometer input on a mobile device, letting users physically toss content to displays in different locations. Other proposed additional interactions with mobile devices include *throwing* and *tilting* to transfer content to a large digital display [68] [69]. These types of interactions – interactions with as opposed to on displays or devices – are only available due to the spatial awareness of the display or device itself and highlights a change from Definition 2, where the only input mechanisms for displays were either keyboard and mouse, touch input or gestures. As a result, a further updated definition is as follows:

Figure 8 - Nintendo Wii U showing a "shake" gesture on the handheld controller shaking content on the TV. An example of simulated spatial awareness in a multi-display environment [70]

Definition 3: A multi-display environment is an environment where multiple displays, either mobile or fixed – wall displays, laptops, PDAs, tablets and computer monitors – can be interacted with by users. Since some of these displays and devices are spatially aware, users can interact with these heterogeneous displays through different input modalities – keyboard and mouse, touch input – or with the displays themselves – via intuitive gestures such as throwing or pointing.

A recent example of Definition 3 is provided by the Wii U [70] game console by Nintendo[1], where users have the ability to control content on a screen using a secondary display. Figure 8, shows how users can interact with a video game that lets them shake content on the TV by shaking a secondary handheld display.

[1] Nintendo, www.nintendo.com

Although some displays are spatially aware in a multi-display environment (as per Definition 3), in order to reach a true "natural human environment" as described by Weiser [17], emphasis needs to be placed on "human". That is, users themselves should play a part in the environment by being spatially recognized. From the perspective of multi-display environments, this means the environment is spatially aware of all devices and users. Newer, commercially-available technologies, such as the Microsoft Kinect sensor [71] and Vicon Motion Capture System [72], allow for spatial tracking of individuals and when combined with spatial information from the devices themselves – such as orientation and position – it is possible to create multi-display environments where both users and devices are tracked. These multi-display environments consider proxemics – the spatial relationship between users and devices – meaning factors such as *position*, *orientation* and *movement* become important [73].

LightSpace, by Wilson and Benko [74], is an example of a prototype multi-display environment that utilizes spatial tracking of both users and displays (although the displays are projected surfaces) to allow for intuitive and novel interactions. In *LightSpace*, a user is able to manipulate content on a tabletop and then transfer it by touching the content and the target display simultaneously. Alternatively, a user in this system may also pick-up content with their hands and drop it onto another target display by touching it with their hands. Bragdon et al. [75] designed a multi-display environment called *Code Space* to support co-located, small group developer meetings. This multi-display environment used a combination of multi-touch screens, spatially aware mobile touch devices and the Microsoft Kinect sensor [71]. This allowed users to perform cross-device interactions with a combination of in-air pointing (with or without a device) and touch.

Ballendat et al. [73] created a prototype multi-display environment to examine interaction design in proxemics relationships. In this multi-display environment, users can interact with a touch-enabled wall display that reacts to their distance, identity, physical location, orientation, and even devices such as a mobile phone or tablet. This multi-display environment and others highlight the potential of full spatial awareness in multi-display environments and leads to the definition used in this thesis. The thesis definition, unlike Definition 3, places an emphasis on the user and is the closest definition to the “natural human environment” envisioned by Weiser [17]. It is important to note that although multi-display environments exist that are entirely mobile and have varying degrees of spatial awareness, they are considered out of the scope of this thesis [76] [77].

2.2 Interaction Design in Multi-Display Environments

In an environment with any number of users and devices, an inherent problem for user experience professionals is how to allow users to move content across different displays. Movement of content in traditional PC-based environments with multiple monitors (Definition 1 in Section 2.1.1) is typically accomplished with a keyboard and mouse. In a multi-display environment where the “interaction is distributed over several different devices” (as per the definition used in this thesis), moving content has remained a primary design challenge in the research literature [78]. A number of unique approaches have been established and are discussed in Section 2.2.1. However, a problem with these approaches is that they are developed by system designers who were constrained by the technical limitations of their particular multi-display environments. As a result, this means that the interactions they designed may not be usable, as they may not map conceptually to those of the users, as mentioned in Section 1.4. Section 2.2.2 discusses the elicitation procedures that are used to develop user-defined gestures and interactions and solve this conceptual mapping problem.

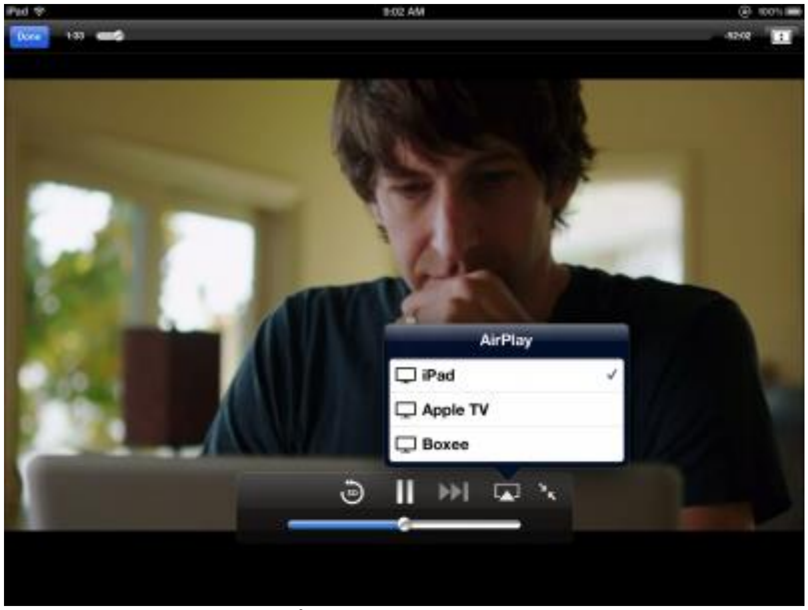

Figure 9 - Apple's Airplay[2], An example of using a GUI approach to select other devices

2.2.1 Approaches for Interaction

Interactions that have been established in the literature for multi-display environments are quite numerous; however, they can be generally categorized into approaches that use *graphical user interfaces* (GUI), *tangibles or objects (physical),* or *gestures and proxemics*.

2.2.1.1 Using a Graphical User Interface

For a majority of the interactions in the early multi-display environments described in Section 2.1, a GUI was the most common approach used to transfer content. This is because the GUI approach, regardless of the technology used in a multi-display environment, is traditionally the most familiar to users. The *Windows, Icons, Menu, and Pointer* (WIMP) GUI paradigm – which first began with Douglas Engelbart during the mid-1960s [79] and was later refined by Xerox's Parc Alto and Parc Star [80] as well as Microsoft's Windows operating systems [81] – is a well-established paradigm and each component in WIMP can be utilized in a number of different GUI approaches for multi-display environments, which are outlined in the following subsections.

2.2.1.1.1 Traditional Menus

A menu in a GUI provides a list of options or commands for an application or computer system. In a multi-display environment, a similar approach can be used to transfer content between devices and displays. *TeamSpace,* by Hebert and Chen [51], is comprised of several laptops and a shared wall display and utilizes a menu-based GUI approach to transfer content in a collaborative environment. Users are able to select other devices and displays from a menu and send content to the selected device or display. Users are also able to take control of the wall display through a menu. The menu-based approach is also used for driving interaction in a geospatial application designed for multi-display environments by Forlines et al. [82]. Menus on a tablet are used to control information (such as layers and drawings) that appears on a modified Google Earth application running on a digital tabletop and wall display. The menu contains information about what information a device or display is currently displaying and users can simply choose to add or remove information for a particular device or display by selection in the menu. Alternatively, *Dynamo* by Izadi et al. [43] uses a menu and icon-based GUI approach to let users transfer content. The icons represent the individual displays or devices in the environment and, in the menu, users are able to drag content (webpages, pictures, videos) to the appropriate targeted device icon.

In general, the GUI approaches for multi-display environments list the devices or displays in the environment in either a text-based menu form or use icons in a menu to represent the devices. Clicking or dragging is then used as a means of interaction to facilitate the transfer of content. This type of approach has also been seen in the retail space with Apple's Airplay[2] technology, letting users share music and video content through various Apple products – such as laptops, mp3 players, phones and tablets (see Figure 9).

[2] Apple, www.apple.com

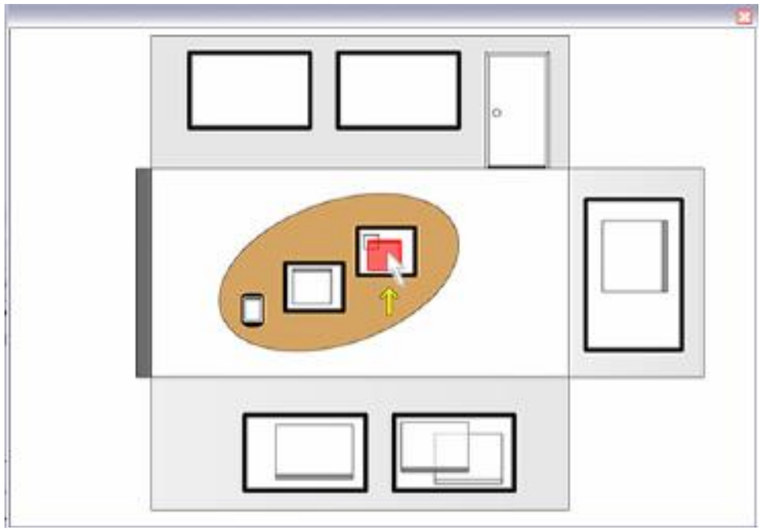

Figure 10 - Example of a WIM approach to selecting displays from [85]

2.2.1.1.2 World-In-Miniature

A *world-in-miniature* (WIM) approach for GUIs maps a physical space to a 2D GUI. These types of interfaces typically are presented in a "top-down" view of a physical space with icons or text spatially mapped to specific devices or users in an environment. This type of spatial mapping allows users to easily conceptualize spatial relationships of devices and other users in an environment. Accordingly, it is faster for users to send content to a targeted device compared to a menu-based approach where users select a target device from a list [50]. This is reinforced by prior multi-display environment research [83] which indicates that users are inclined to think of spatial relationships with devices either in terms of themselves or the environment, much like they would with real-world objects (such as a cup or a book). However, the WIM approach may not be faster in cases where the number of devices is either extremely low (e.g. 1 or 2 devices) or high enough to clutter an interface. Examples of a WIM approach for multi-display environment where users are able to select a display from a spatially-mapped WIM GUI and then select content or commands to send to a targeted display are provided by Ponnekanti et al. [84] and Biehl et al. [85] (Figure 10). Another approach to WIM where different interface elements are used to provide links

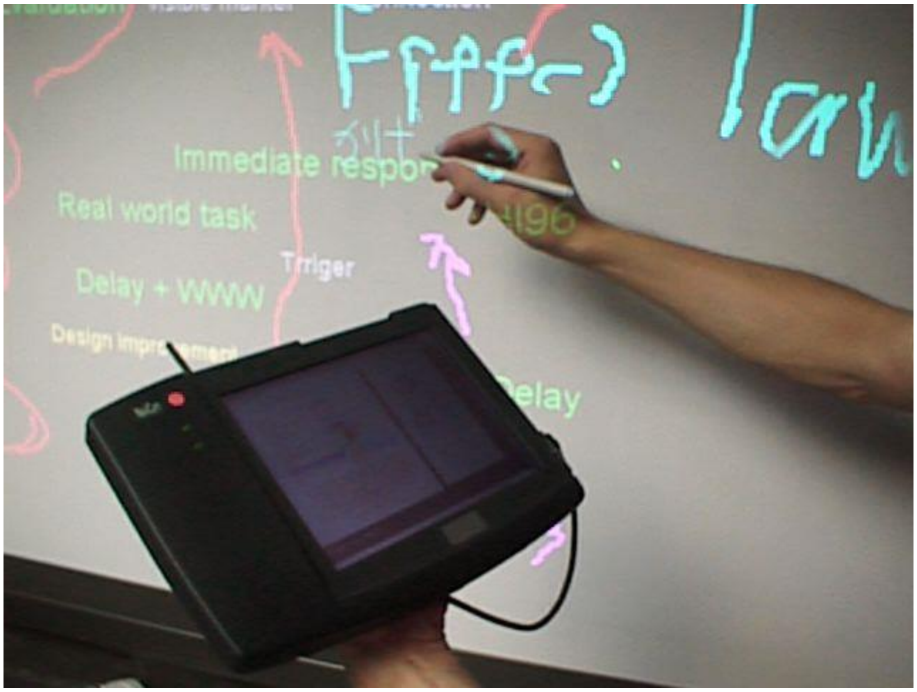

Figure 11 - Pick And Drop technique by Rekimoto [33]

between different displays is provided by Wigdor et al [49]. Users are either able to drag to conduits (called cords) that are mapped to different displays or to matching colors and shapes of the different displays.

2.2.1.2 Using Physical Approaches

A majority of the interactions that users have without technology can be considered forms of physical interactions and they may include physical objects – swinging a hockey stick while playing hockey or throwing a baseball. The mapping of these physical interactions with the available devices and users in a multi-display environment may result in interactions that physically and conceptually map to real-world interactions. The techniques in which devices are used in physical approaches can differ and they are outlined in the subsections below.

2.2.1.2.1 Physical Objects as a Medium

Utilizing a physical object as a means of transferring content in a multi-display environment is a useful technique for designers because, typically, this type of interaction maps to a conceptual model familiar to

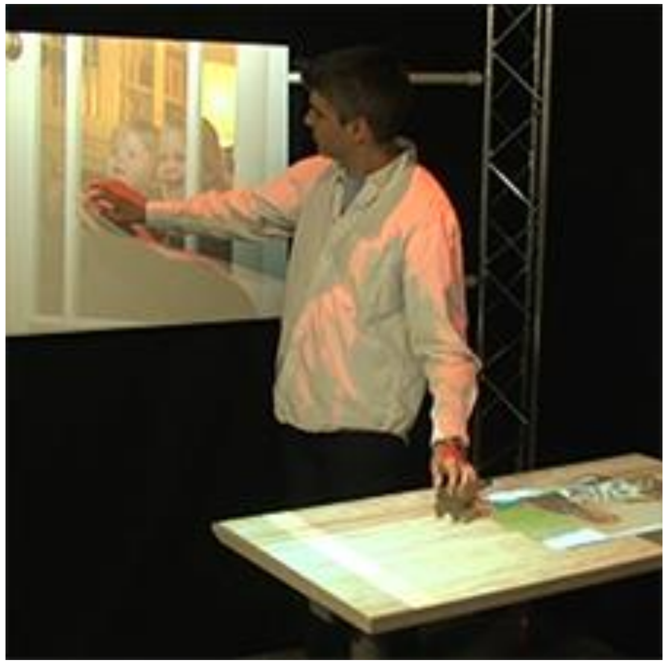

Figure 12 - A user acting as a direct conduit between two different displays via direct contact

users. For example, Rekimoto [33] provides a *pick and drop* interaction technique that allows users to select (or *"pick"*) content on one device by tapping it with a pen and place (or "*drop"*) it on another device by tapping it again with the same pen. The pen in this case provides a physical medium to transfer the information (as shown in Figure 11). Another example of using a physical medium as an interaction approach is provided by *i-Land* from Streitz et al. [6], mentioned in Section 2.1.1. Users are able to link information to any arbitrary object (a key, watch or block of wood, for example) that can be later placed onto digital tabletops that identify and transfer the data linked to the object.

2.2.1.2.2 Direct Contact

Direct contact for interactions in multi-display environments implies directly contacting one or more devices to transfer content. An example of this direct contact interaction approach is provided in *LightSpace* by Wilson and Benko [74] previously mentioned in Section 2.1.2. Users are able to simultaneously contact a wall display and a tabletop to transfer content. In essence, a user becomes a conduit by directly contacting the two displays simultaneously, as shown in Figure 12. Another technique is provided by Rekimoto with *Augmented Surfaces* [86] where users are able to directly contact devices in order to transfer content.

2.2.1.3 Using Gestures and Proxemics

Unlike the interaction approaches discussed in Sections 2.2.1.1 and 2.2.1.2, gestures provide additional advantages in interactions for multi-display environments, particularly because it is another input modality that can be combined with touch-based displays [87]. Proxemics is a sociological field that examines personal space and social interactions of people, and is defined by physical interaction zones based on distance [88]. Proxemic interactions are based on this field and are more concerned with the factors that affect spatial relationship between devices, such as *position, orientation, and movement* [73]. Combining the gestural approach and proxemic techniques allows for a powerful mechanism of transferring content, as the gestures (and devices) are now spatially aware of the environment. In the subsections below, a number of different gestural approaches involving devices are discussed that incorporate spatiality.

2.2.1.3.1 Gestures on Devices

Interacting directly on a display or device (typically performed via touch input) in a multi-display environment is an approach that users typically use to select and transfer content. For example, in *Code Space*, by Bragdon et al. [75], users are able to flick content from a mobile display to wall displays or other users with mobile displays. Users are also able to retrieve content from other displays by performing a downward swiping gesture. To retrieve content from a wall display, a user performs this gesture on a mobile display after selecting content to be retrieved. A reverse interaction for sending content is also supplied, wherein users are able to select content with a mobile display and then perform an upward swiping gesture to send to the wall display. Figure Figure 13 (a) highlights an example of a flick gesture done on a tablet to send content to a digital tabletop.

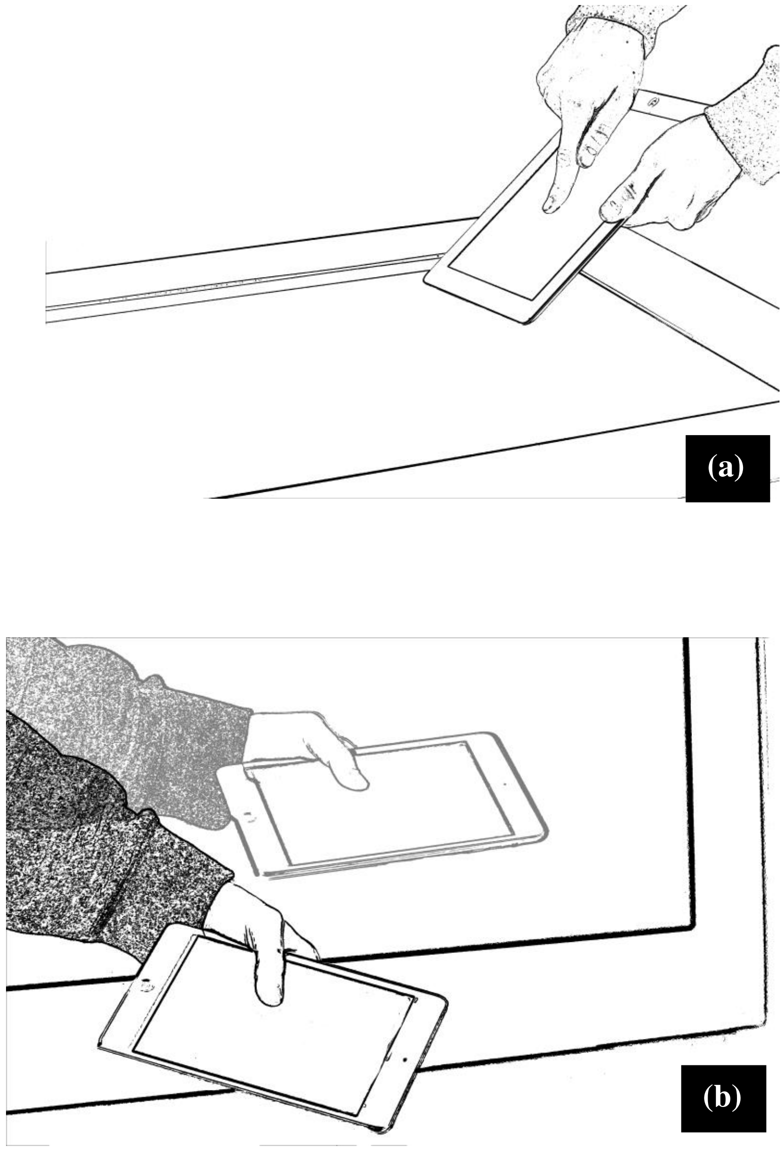

Figure 13 - Examples of Gestures on (a) and (b) Gestures with devices [115]

2.2.1.3.2 Gestures with Devices

Gestures with devices in multi-display environments typically involve the physical movement of mobile devices. The rotation of a device was one of the gestures first explored by Bhandari and Lim [89]. To

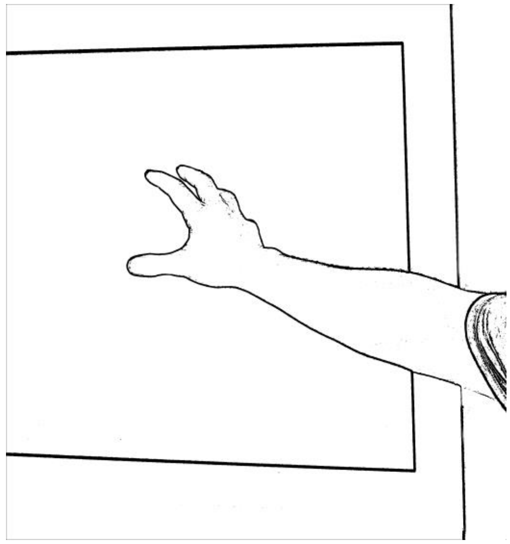

Figure 14 - Examples of Gesture without a device [115]

trigger this gesture, a user rotated a phone from a horizontal position to a vertical position to transfer content to another display. Another gesture explored was the shake gesture where a device is rocked back and forth, along its horizontal axis. This work was limited however, to trigger actions on the mobile devices themselves and not have an effect in the environment, as they were used for photo-editing or playing slideshows. Subsequent work by Doring et al. [69] linked these gestures to multi-display environments and added gestures such as throw from a mobile device to a digital tabletop, as well as pulling from a digital tabletop to a mobile device. The throw gesture and its subsequent iterations in the research literature – Daschelt and Buchholz [68], Hassan et al. [67] – is a primary example of using a device as a gesture mechanism in the research literature. Figure 13 (b) highlights an example of a toss gesture done with a device to send content to a wall display.

2.2.1.3.3 Gestures without Devices

Gestures without devices in a multi-display environment are typically done with hands and arms and can also be augmented with methods such as voice commands [26]. An example of this technique is also provided in *LightSpace,* by Wilson and Benko [74]. By using the positional tracking built into the system and roof-mounted projectors, a hold interaction is created that allows a user to *carry* a digital object in their hand from one location in the environment to another. *Pointing* is an interaction available in *Code Space* by Bragdon et al. [75], where users can point at objects with their fingers to select content from a source device and then point again at another target device to transfer. Figure 14 highlights an example of a grabbing gesture to grab content from a wall display, without a device.

2.2.2 Designing Interactions

A majority of the interactions that have been outlined in Section 2.2.1 have been created by system designers [78]. While these interactions can be effective and interesting, there exists a gap between the users they are designed for and the system designers themselves. This gap results in a potentially different conceptual and mental models of interaction for the system designers and the users. A technique used in the research literature is to create a set of gestures from the users themselves. This is called a user-elicitation technique.

Nielsen et al. [90] proposed a procedure for developing user defined gestures using elicitation. First, tasks for a system are first defined by system designers. Next, gestures are then elicited for these defined tasks from the users, and are then extracted into a gesture set for the system. Gestures are elicited using a guessability study methodology [91], which shows the effects of gestures to users and then asks them to perform actions that elicit the effect. Benchmarking is then done to validate the gestures chosen for the system and ensure that the gestures chosen are the most common, which is required to make a set of gestures for users. Wobbrock et al. first employed this user-defined gesture procedure to construct a gesture set for actions – such as copy, move and paste – on digital tabletops [7]. Participants are first shown the visual outcome of the action (called a *referent*), and then asked to produce a gesture that would result in the shown outcome. Results showed a high level of agreement among the gestures produced for any given task. The agreement value determines if there is similarity (or consensus) in the gestures produced amongst users and a higher value means that a user-elicited gesture set can be created. The gestures created in this set were evaluated in follow-up work by Morris et al. who showed that these elicited gestures were far more preferred by users than those created by expert designers, with previous experience in designing gestures for systems [15].

Frisch et al. [92] also employed this elicitation procedure to elicit a set of pen and touch gestures for diagram editing on tabletops. These gestures were then evaluated by Heydekorn et al [93] where the gestures were built into a real system and then evaluated with a usability study. It was found that implementing elicited gestures in a real world system requires careful consideration of how these gestures are applied, as some of the gestures were found to be difficult to implement in a system or were technically constrained (at the time).

Several researchers have applied a similar type of approach within the context of multi-display environments. For instance, Kray et al. explored how mobile phones could be used in a multi-display environment comprised of a tabletop and a wall display, focusing on a wide range of tasks and content types (e.g. voting for content, synchronizing devices, and rewinding interactions) [94]. In this work the elicited gestures were not categorized or labeled so no agreement statistics could be produced; however, the authors observed a wide variety of novel and diverse gestures which they reasoned was a result of the wide means in which a mobile phone could be used. Similarly, Kurdyukova et al. elicited a set of gestures from users specifically for transferring content between a tablet and fixed devices such as a wall display and tabletop [95] and included tasks such as tablet to tablet transfers, tablet to digital tabletop transfers and tablet to wall display transfers. These gestures were evaluated qualitatively however, so it was unknown if a gesture set could be created.

2.3 Displayless Space in Multi-Display Environments

There is a natural importance to the spatial layout of devices and users in a multi-display environment. The spatial layout of the environment and the inherent spatial separation between static displays, mobile

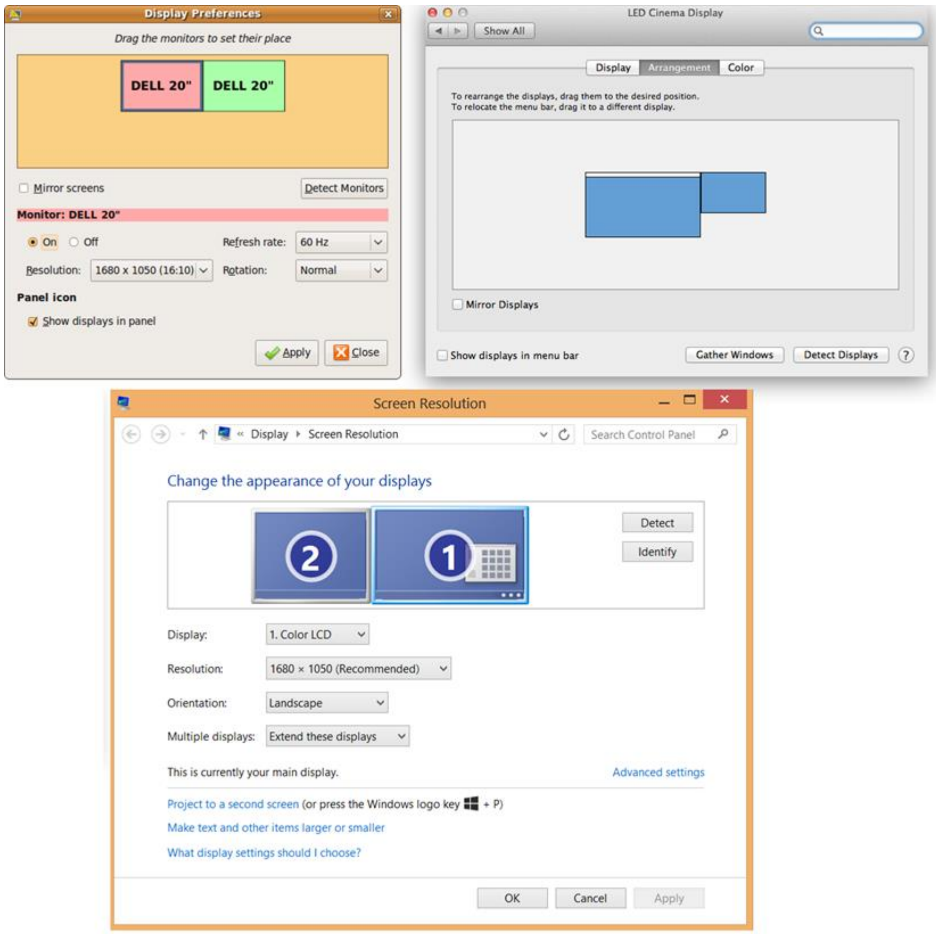

Figure 15 – Configuring a continuous desktop in various operating systems

displays, and users is important at an interaction level [96]. This separation is defined as *displayless space* [3] (the space between devices or displays) and how it is regarded in the spatial layout of an environment is a constraint. For example, devices (or displays) in a multi-display environment could be linked together in a common space, (also known as spatially continuous) considered separate (also known as spatially distinct) or a combination of both (a hybrid). This following subsections describe the differences between these types of multi-display environments.

2.3.1 Continuous

In traditional PC-based environments where multiple displays can exist, it is possible to link the displays into a single and continuous display space. This allows input to be across displays, as a user can move a

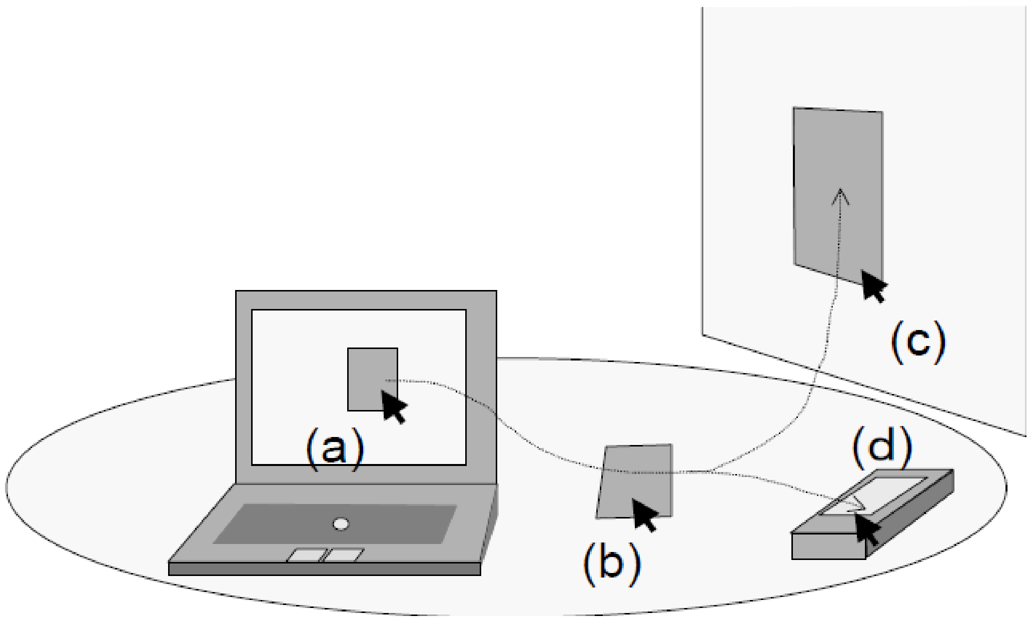

Figure 16 - A spatially continuous workspace, where a user can (a) drag an object from a laptop to a table (b) and then drag to a wall display (c) and finally to a physical object (d) [86]

mouse cursor continuously in the direction of any destination display. The displays are typically mapped to a conceptual physical location and several modern operating systems utilize this technique when linking multiple displays, as shown in Figure 15. In addition, software solutions also exist to allow separate PCs and laptops to create continuous display space with their respective displays as well [97] [98] [99].

A similar technique was applied to mobile displays by Kohtake et al. [48] in which users were able to physically attach display panels to other display panels resulting in a shared continuous space. An advantage of this system was that users were able to spatially configure the mobile displays for different tasks which is extremely typical in traditional PC-based environments (where users can arrange monitors based on a task or activity) [100]. Rekimoto et al. [86] followed this technique when developing a spatially-continuous interactive workspace that incorporated laptops, wall displays, and a desktop (See Figure 16). The interactive workspace mapped each of these components into a shared space that represented their associated physical location. Users in this prototype system (and subsequent prototypes from other researchers) were able to move content between displays by simply dragging it to the next display that was spatially-continuous [43] [101] [102] [103] [104] .

Unlike interactive workspaces, multi-display environments can be spatially-aware (devices or users or both can be tracked), which results in unique possibilities for a spatially continuous display space. For instance, a user in a multi-display environment could flick their content from one display to another display that is across the room but still part of the same continuous display space. Figure 17 highlights a multi-display environment that contains displays – a digital tabletop, wall display, 2 tablets and a mobile phone – that are configured in a continuous manner. The different displays show different parts of the same map highlighting the spatial continuity of the display space.

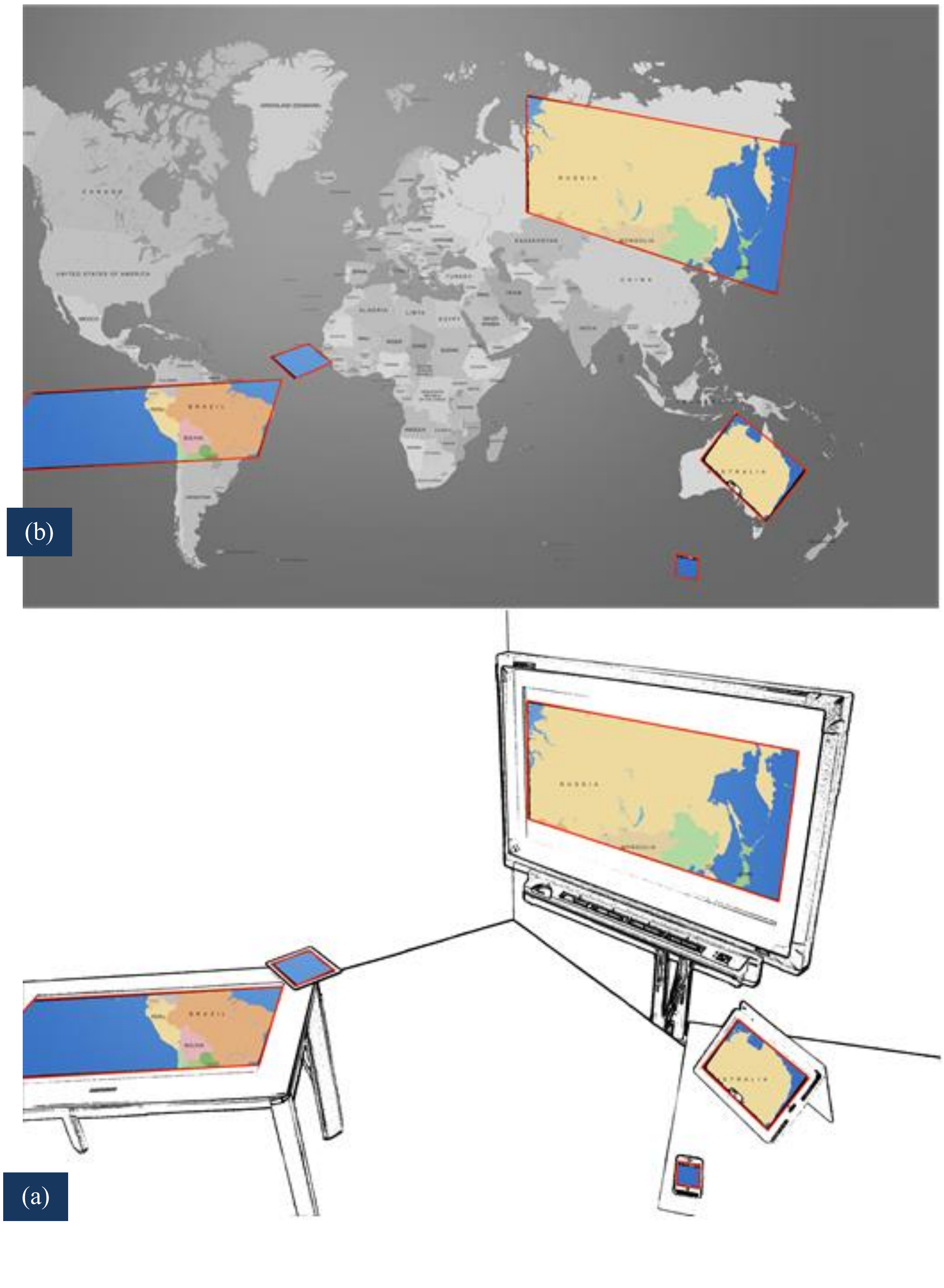

Figure 17 – A spatially-continuous MDE configuration. If a mobile device in (a) is moved, it would move to the appropriate location in the shared world map that is displayed in (b)

2.3.2 Discrete

In a traditional PC-based environment, it is also possible to have multiple displays that are not linked together but are related in some manner other than being in the same physical area. An example of this would be a workspace containing a laptop, mobile phone, and a desktop with a large display running a media application. Both the laptop and mobile phone each have a different means of controlling the media application on the large screen albeit each device contains different content, which highlights the nature of discrete displays that are related.

Aside from i-Land by Streitz et al. [6] [38], as described in Section 2.1.1, another example of a spatially-discrete workspace is provided by Wigdor et al. [49] where a digital tabletop is used to facilitate collaborative situation assessment and decision making for the New York Police Department. In this workspace, two wall displays and a digital tabletop are provided – each with distinct digital workspaces that have different menus and content. Although the displays are spatially discontinuous, users are able to share content between the displays through icons that represent each display (a WIM approach). Other means used for transferring content in this type of workspace include menus or widgets which allow users to drag the content to different displays [35] [6] [38] [39] [40] [41] [47] [49] [50] [51] [52].

In a multi-display environment where the environment is entirely spatially-aware, menus and portals may not be as intuitive to transfer content since the environment conceptually maps to a physical environment with objects that can be moved. As described in Section 2.1.1, *LightSpace* allows users to physically manipulate digital content stored on displays. Users are able to pick up content from a digital tabletop and place it on a target wall display, mimicking physical interactions performed in the real-world [74]. Figure 19 highlights a spatially-discrete multi-display environment where different types of maps are displayed on different displays and, unlike a spatially continuous multi-display environment (see Figure 17), there is no shared map. Interactions that are possible in this type of multi-display environment include picking and dropping, flicking, pouring, and throwing different maps to different displays, as the environment is entirely spatially aware [33] [68].

2.3.3 Hybrid

A hybrid multi-display environment is a combination of spatially-discrete and spatially-continuous displays. As shown in Figure 19, spatially discrete displays each showing different types of maps are combined with spatially continuous displays showing different parts of the same map.

The *WILD Room* [105] is a prototype multi-display environment that combines spatially-discrete displays and spatially-continuous displays into a unique interactive environment. The environment consists of an ultra-high resolution wall display, digital tabletop, and numerous mobile devices. This hybrid multi-display environment allows for users to have two distinct workspaces in the environment: a spatially-discrete private workspace where individual users can view and manipulate content and a spatially-continuous shared workspace where users can share content and collaborate. The major benefit of hybrid environments is providing users beneficial private and public workspaces [75] [106], however, this is not a primary focus of work presented in this thesis.

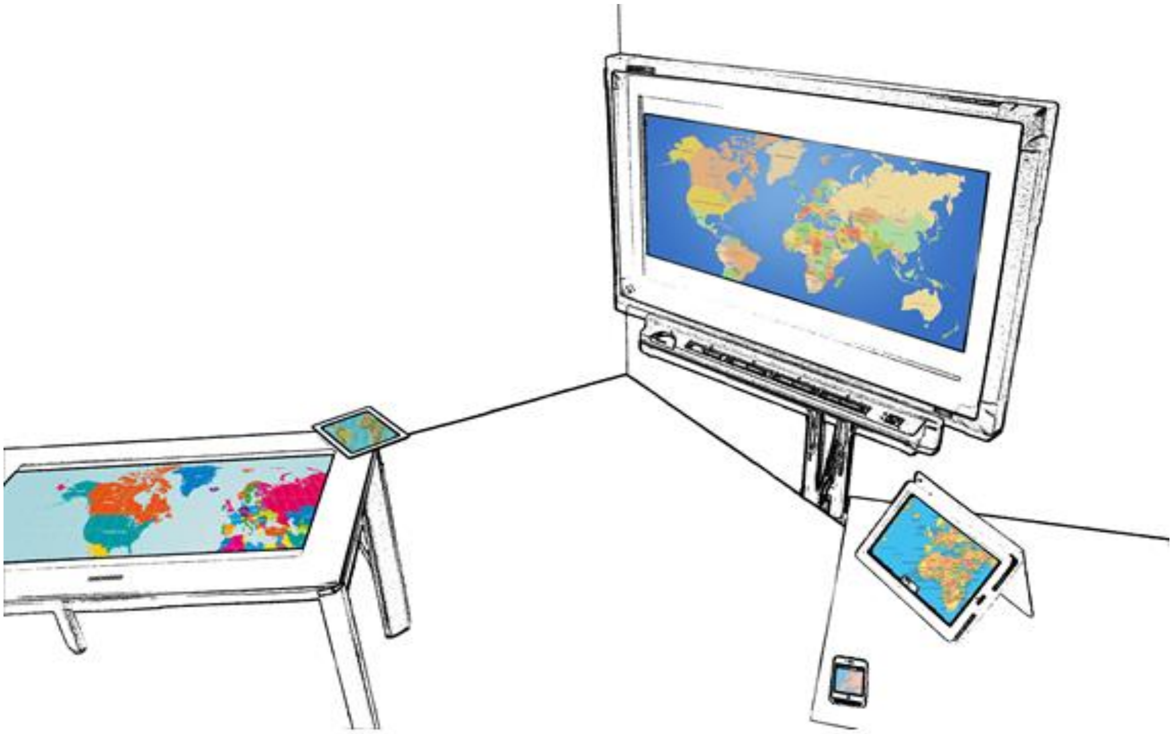

Figure 19 - A spatially discrete multi-display environment. Each display has a different type of map displayed, indicating each device display is discrete

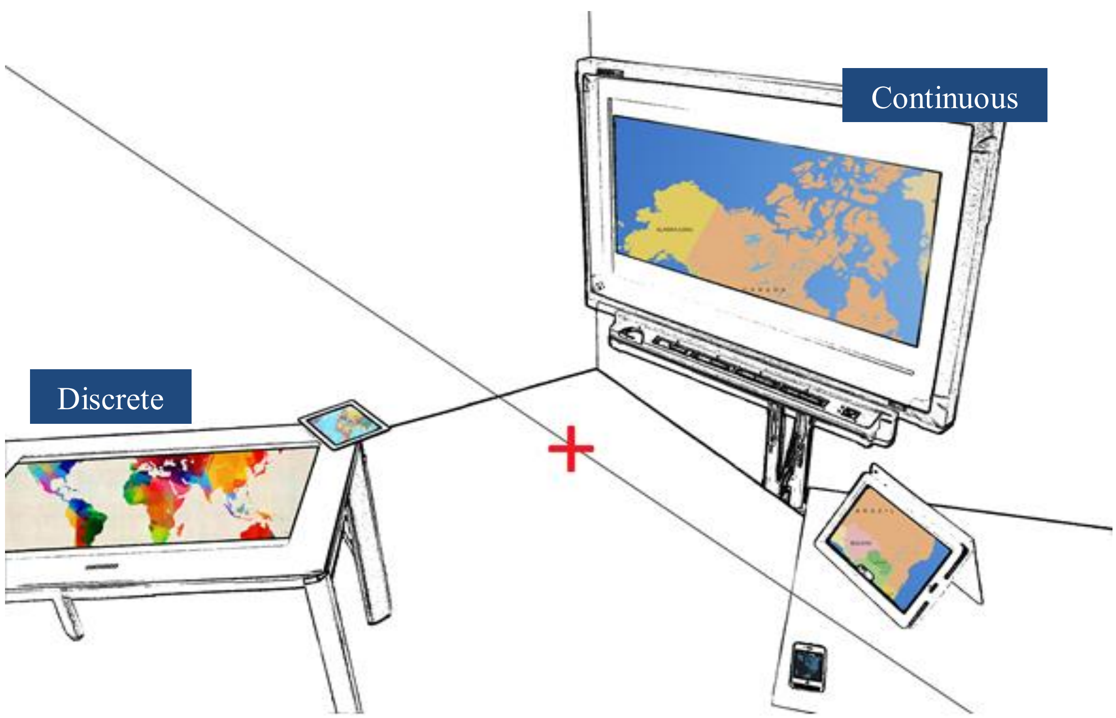

Figure 19 - A hybrid multi-display environment configuration. Some components of are in a spatially discrete space and some are in a spatially continuous space, indicated by the shared map and discrete maps

2.4 Summary

In this section, the evolution of the definition of multi-display environments was discussed with emphasis on the changes as displays evolved from stationary to mobile as well as the increase in availability of spatially-aware displays and tracking systems. This led to a categorization of multi-display environments in the research literature that is based upon mobility of the displays and the spatial awareness of the environment. With such a broad categorization of multi-display environments in the research literature, however, it was necessary to provide a definition that fit the scope of this thesis work. In this thesis, a multi-display environment, as discussed earlier (page 15), is:

> *"A spatially-aware environment where interaction is distributed over several different devices – such as digital tabletops, wall displays, tablets and mobile phones. This distribution of interaction means that multiple users in this environment can perform multiple activities with any number or combination of different display components."*

This definition implies that displays are not merely visual output but also a means of interaction. Also discussed are the different approaches for interactions in multi-display environments – such as using a graphical menu, physical approaches, as well as gestures and proxemics – and techniques to design gestures based on user feedback.

Another important component to interactions in multi-display environments is the notion of how the displays themselves are treated. These displays can be treated as though they are in the same shared space (*continuous*), entirely separate (*discrete*), or a combination of both shared and separate (*hybrid*). This brief examination of how displays are treated in interactions for multi-display environments serves as a concise introduction to the upcoming Chapter 3, where interactions specifically for transferring content are explored.

Chapter Three: Exploring Interactions for Multi-Display Environments[3]

Designing gestures and interactions for multi-display environments presents a number of challenges. It is important that users understand available interactions as they form the basis of how they use devices that have different capabilities such as size, position, resolution and mobility. A common task in interactions for multi-display environments involves users moving content among the different devices. With multiple devices with different capabilities, this can be a critical problem for user experience professionals in multi-display environments. For example, moving content with a mobile device provides different gestural possibilities than a static device, such as a digital tabletop. If users are unable to interact with content in a multi-display environment properly, then both its effectiveness and usability are reduced. Generally, the prior interaction approaches discussed in Section 2.2 – using a graphical user interface, physical approaches, gestures – have all been created by system designers. This raises the following issues:

1. *Many of the interactions designed by the system designers were built around technical constraints*. This results in interactions for transferring content that are easier to implement in a multi-display environment, but not necessarily easy to learn and remember by users.
2. *The conceptual model of the user and the system designers may not match*. As a result of the constraints of the system, the mental model users have for content transfer interactions may differ from the conceptual models of interactions provided by system designers. From a user experience perspective, this means that content transfer interactions created by system designers may not be ideal in real-world settings with everyday users.
3. *Interactions aren't necessarily generalizable across all devices and displays*. Another problem with designing content transfer interactions in technically constrained systems is that they do not necessarily generalize to all multi-display environments. For example, the tossing gesture allows

[3] The contents of this chapter are based upon [78]

a user to toss content to a wall display by mimicking the physical action of tossing an object. This interaction may work for mobile devices and tablets that are easy to move and have necessary technology (like accelerometers) built in; however, this would not work for, say, wall-mounted displays.

As part of this thesis, a user elicitation study was conducted to find a set of interactions and gestures from the users (as opposed to the designer-created interactions in Section 2.2) for transferring content in multi-display environments. The primary motive behind this study was to better understand content transfer interactions in multi-display environments and tackle the three aforementioned issues. This type of study not only allows for a set of user-created content transfer interactions, but can also lead to better interaction design, as the mental model of users is now the basis for the conceptual model for content transfer interactions. The user elicitation study and its design are discussed in Section 3.1, while Section 3.2 discusses the qualitative and quantitative techniques used for analyzing the data obtained from the study. The results are presented in Section 3.3 and followed up with a discussion in Section 3.4 that provides informed design suggestions for designers of multi-display environments. Section 3.5 discusses the limitations of the study which is then followed with a conclusion of the chapter in Section 3.6.

3.1 Study Design

The key focus of this study is understanding interactions for transferring content in multi-display environments. In designing these interactions for a heterogeneous multi-display environment, the capabilities of each device — size, position, resolution, and mobility —plays an important role. However, the extent that this affects a user's conceptual model of interaction in a multi-display environment isn't well understood and provides motivation for the study. Providing a user-defined gesture set for transferring content in multi-display environments is another motivation for this study. These gestures can provide insight into the conceptual models that users have for content transfer interactions, as they reflect their behavior and are a necessary step to creating real-world interactions for multi-display environments.

As transferring content can be a common task that users perform in multi-display environments, the participants in the study were asked to perform a series of tasks that relate to transferring content between devices that can exist in a multi-display environment. Additionally, users were asked to transfer content at different positions in a multi-display environment, highlighting two heterogeneous factors of multi-display environments – position and mobility. This allows for an examination into the impact of spatial and device characteristics on conceptual models of content transfer interactions in multi-display environments, which is a goal of this thesis.

The tasks used for the study are based on those by Voida et al. [107], where users were asked to create gestures for accessing multiple projected displays in an augmented reality office. The tasks users performed in the study were based on an exploration into how distance to an object, the type of an object and spatial models affected gestures for augmented reality. To examine the impact of spatial and device characteristics on interactions and gestures for content transfer in multi-display environments, this provides a good foundation to build upon for the study. As a result, the content transfer tasks examined the following:

- *Transfer Direction* – sending and retrieval of content from tablets, a tabletop and a wall display.
- *Source and Destination Devices* – which are tablets, a tabletop and a wall display.
- *Distance* – varying the distances, from close to far between tablets, a tabletop and a wall display.

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

Figure 20 - List of tasks used for the study

With these examinations in mind, a number of content transfer task combinations could be created and Figure 20 lists the final set of task combinations that were used in the study to elicit a set of gestures for content transfer in multi-display environments. A singular task in this study, refers to a singular row in Figure 20. Tablets were a major focus of the content transfer tasks in this study, as they are more likely to be used by real-world users and in business settings or collaborative environments than wall displays or digital tabletops [108]. The relative positions of devices and distances for tasks involving the sending and retrieval to wall displays and digital tabletops are shown in Figure 21 and Figure 23. For the content transfer tasks involving tablet to tablet interactions, a researcher was used as a secondary user to which a participant could send content. The relative positions of the researchers (for tablet to tablet tasks) and participant are shown for these tasks in Figure 22.

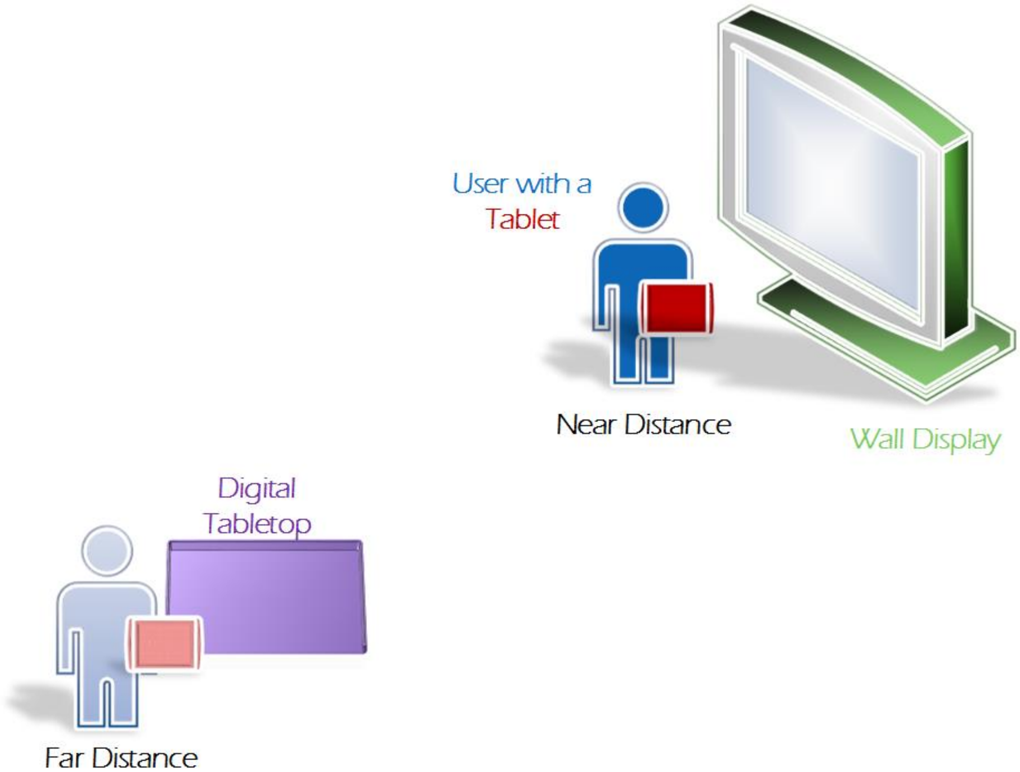

Figure 21 - Room schematic and distances for content transfer tasks involving tablet to wall display

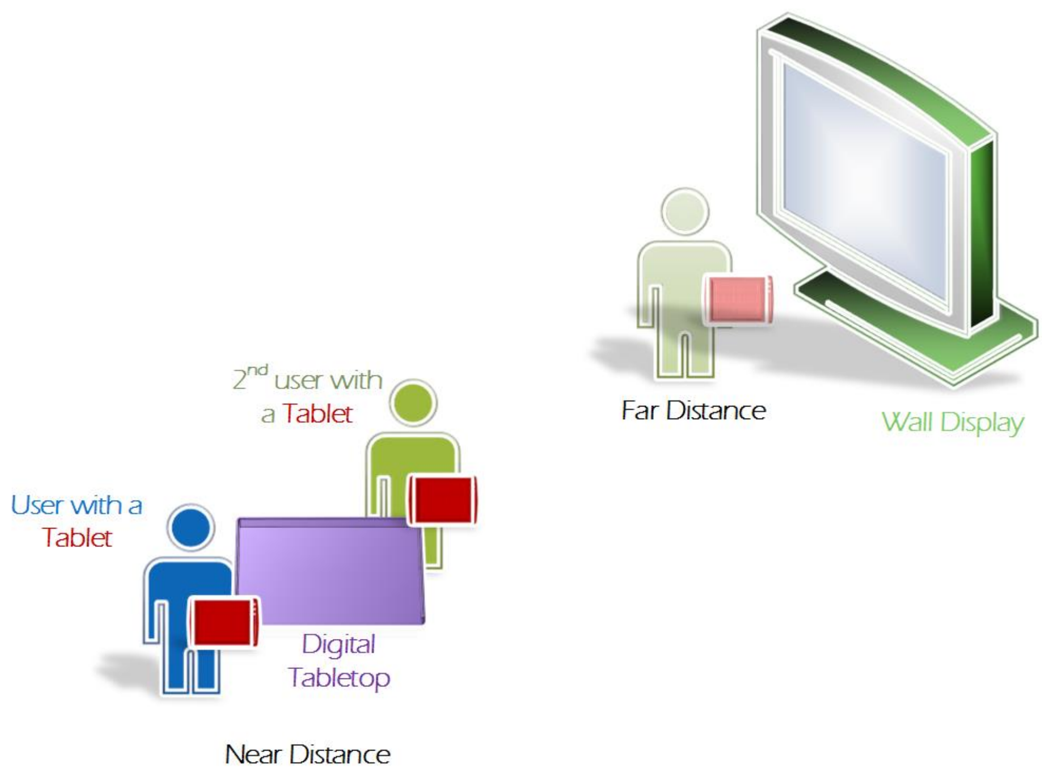

Figure 22 - Room schematic and distances for content transfer tasks involving tablet to tablet

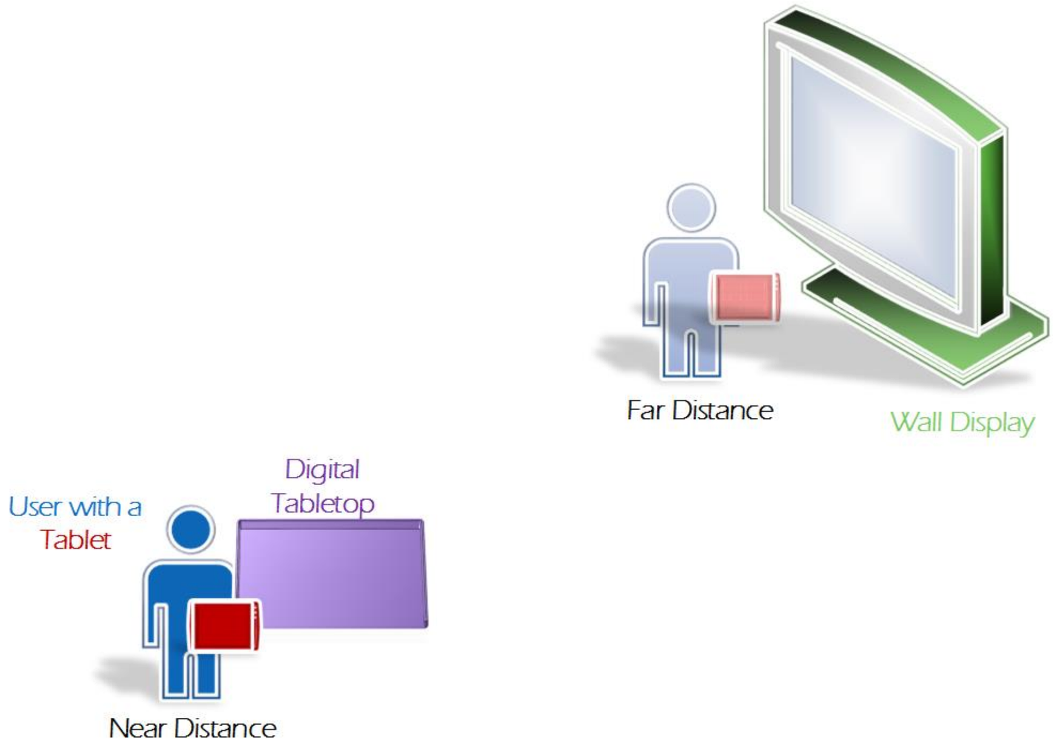

Figure 23 - Room schematic and distances for content transfer tasks involving tablet to digital tabletop

3.1.1 Participants and Descriptive Data

Using recruitment methods such as posters, email and word of mouth, a total of 17 paid participants were gathered for this study. A breakdown of the participants is provided in Figure 24 and Figure 25.

Of the 17 participants, 11 were male and 6 female, and their backgrounds included business, economics, engineering, and education. Ages of the participants ranged from 18 to 60 years of age, with the average age of males being 27 years of age and the average female age being 25 years of age. 11 of the participants were students at the time of the study, while 5 were professionals from industry and 1 participant was both a student and a professional from industry.

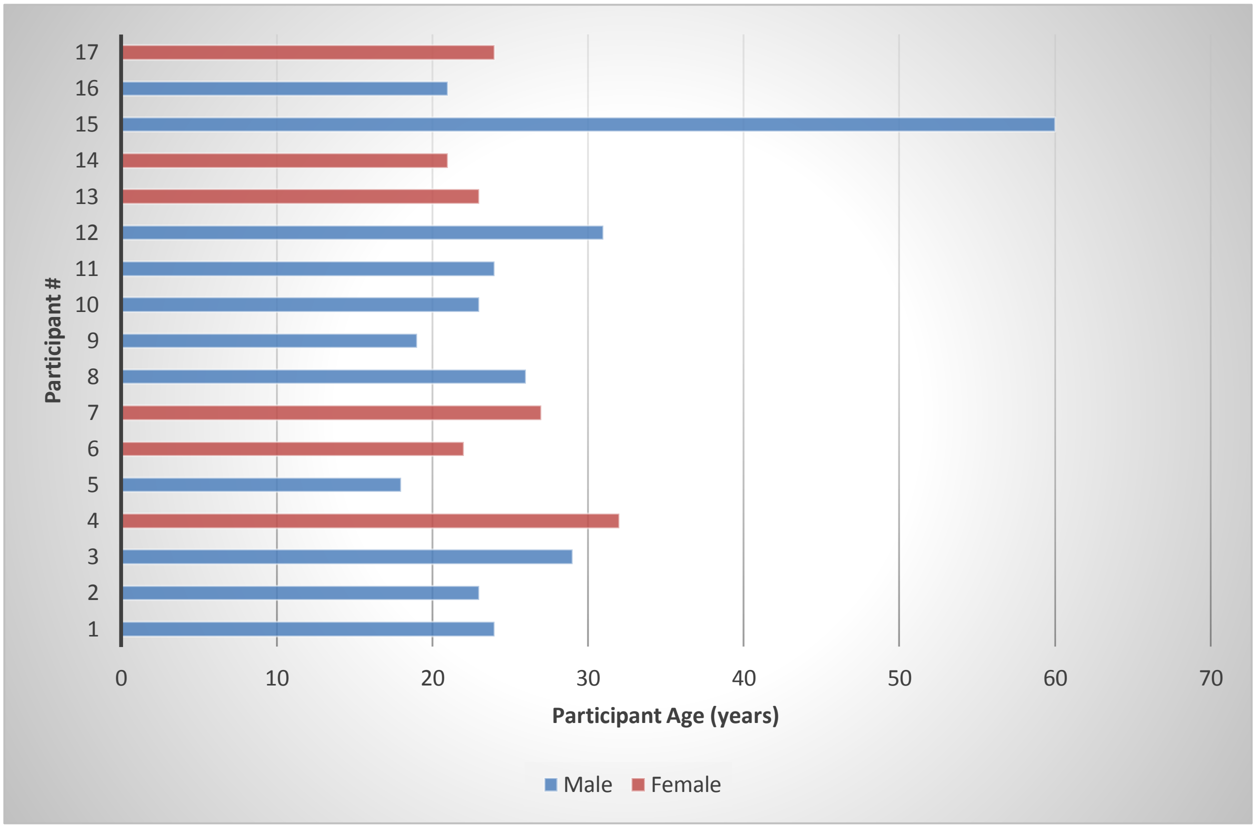

Figure 24 - Age and gender of participants

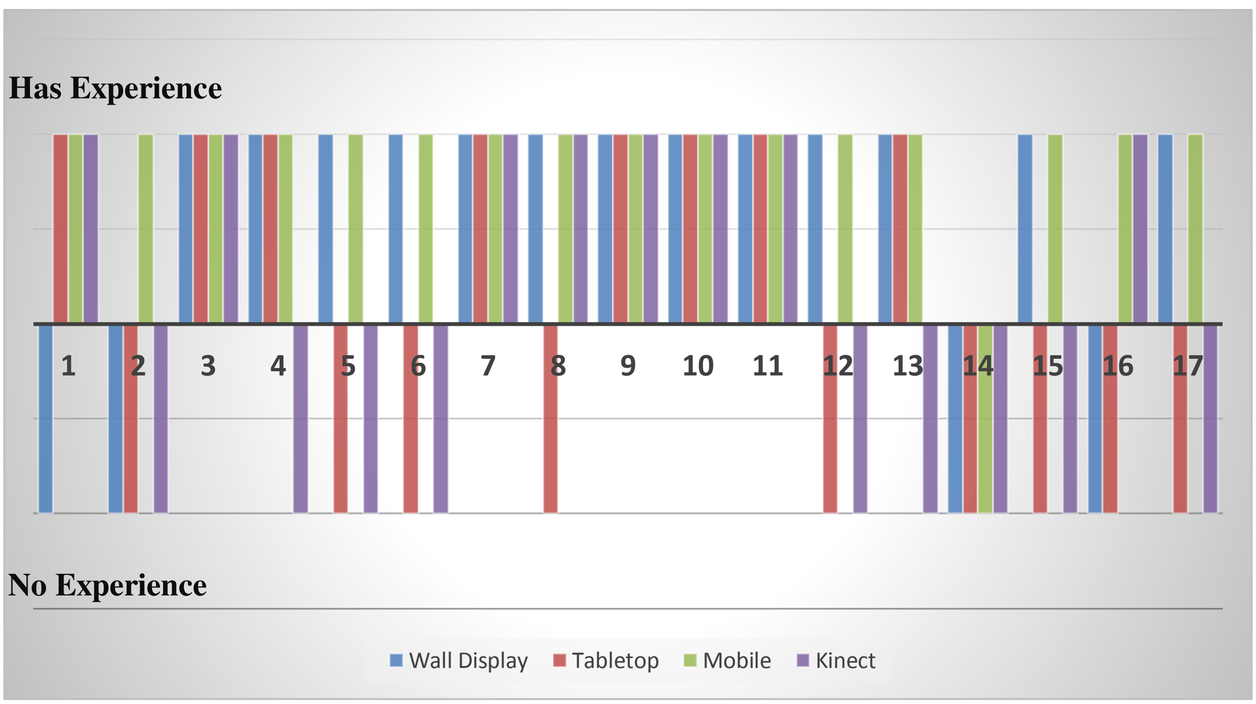

Figure 25 - Participant background experience with different technologies available in multi-display environments

Figure 25 provides a breakdown of the experience with the technologies that are available inside a multi-display environment – motion tracking systems, digital tabletops, tablets, wall displays – for each participant in the study. A percentage breakdown for the background experience with technologies for the participants in Figure 25 is as follows:

- 76% of the participants had previous experience with a wall display.
- 47% of the participants had previous experience with a digital tabletop.
- 94% of the participants had previous experience with touch-enabled mobile devices such as mobile phones or tablets.
- 47% of the participants had previous experience with the Microsoft Kinect or similar motion tracking technologies.

Noticeably, some participants had limited experience with the technologies available in multi-display environments. No participants were excluded based on experience, as participants with a wide range of experience was desired, to better represent the wider population that may use multi-display environments.

3.1.2 Apparatus

The study utilized the following apparatus:

- *Hardware:* Two Apple iPads[4], a Samsung SUR40 Digital Tabletop with PixelSense[5] and a SMART Board[6] were used to create a multi-display environment.
- *Software:* To visualize the outcomes of tasks that users performed, a specialized Node.js[7] web application was written. This application allowed for simulated transfer of images between the different devices in the environment, which represented the visualization of the outcome (also

---

[4] Apple iPad, http://www.apple.com/ca/ipad/
[5] Samsung SUR40, http://www.microsoft.com/en-us/pixelsense/default.aspx
[6] SMART Board, http://smarttech.com/smartboard
[7] Node.js, http://nodejs.org/

known as result) of a task. This was done by showing an image disappearing on an origin device and appearing on a target device after a short delay, caused by a secondary researcher operating the application.

- *Other:* A video camera recorded all gestures performed by users. To consider distance consistently between studies, specific locations were marked on the floor and participants were asked to move to these locations for certain tasks. The relative locations for the tasks in the study are indicated in Figure 21 to Figure 23.

3.1.3 Procedure

For all participants, the session began with a brief introduction that discussed the purpose of the study and video information from films such as Iron Man 2[8] and Minority Report[9], highlighting the possible types of gestures they could produce in a multi-display environment. This was done to provide participants with inspiration in the creation of the gestures for the upcoming content transfer tasks. The types of gestures participants witnessed in the videos included gestures performed by rotating or manipulating a device (where possible), on-device gestures (such as tapping or pinching), and gestures without devices (using arms, hands, or entire body).

After the introduction, participants then began to perform the first task in Figure 20. At the start of every task, the purpose of the task was explained verbally and the intended outcome of the task was displayed. At any point during a task, participants could ask questions to clarify the nature of the task. Participants were also continuously encouraged to create gestures without any regards to technical feasibility when creating gestures. This was done in order to limit the influence of gestures being created or chosen because they are easier for a system to recognize. Furthermore, as some participants also had experience with some

[8] Iron Man 2, http://ironmanmovie.marvel.com/
[9] Minority Report, http://movies.foxjapan.com/minority/index.html

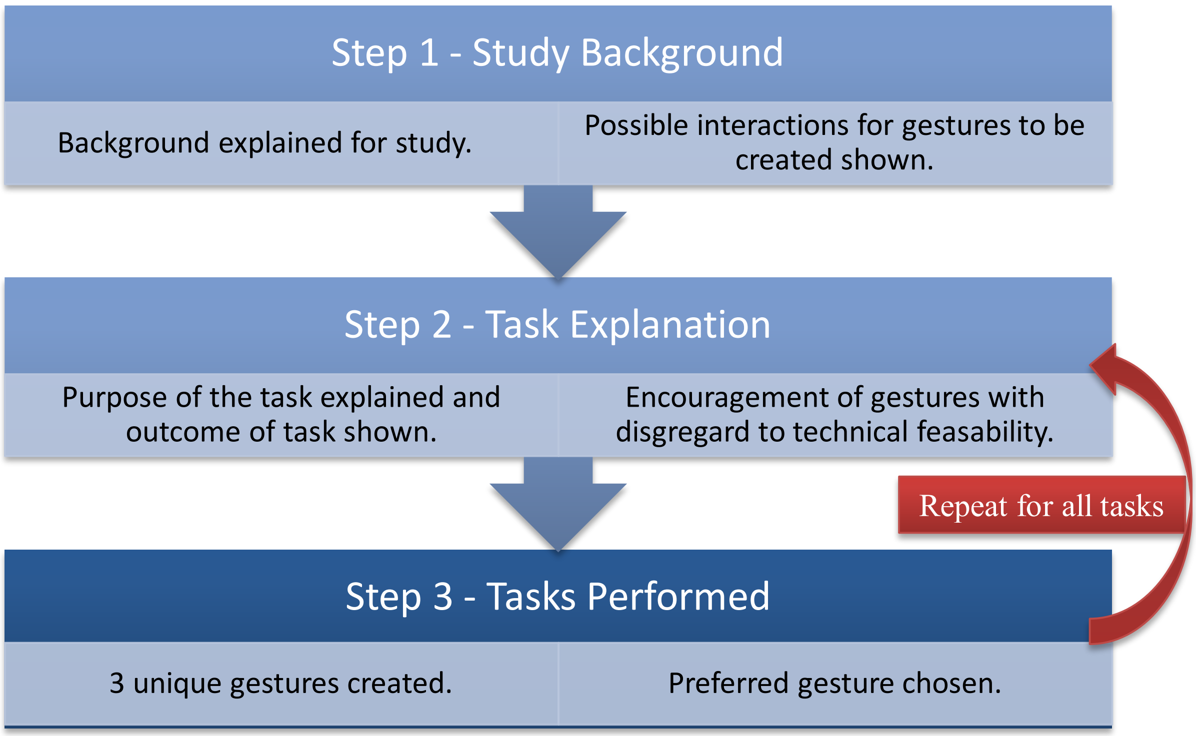

Figure 26 - Study Design

of the technologies in the multi-display environment, that meant they understood some of the technical limitations of the technology. Removing technical feasibility also served as a means to encourage imagination for participants who had knowledge of the technical limitations of the technologies in the multi-display environment.

For each task, participants were also asked to (shown in Figure 26):

1. Create three unique gestures (as a means to stimulate creativity).
2. Explain verbally the gestures being performed.
3. Choose a preferred gesture for the task.

Each participants performed all 16 tasks that are listed in Figure 20.

3.2 Analysis Techniques

The following sections describe the techniques used for both qualitative and quantitative analysis of the results of this study.

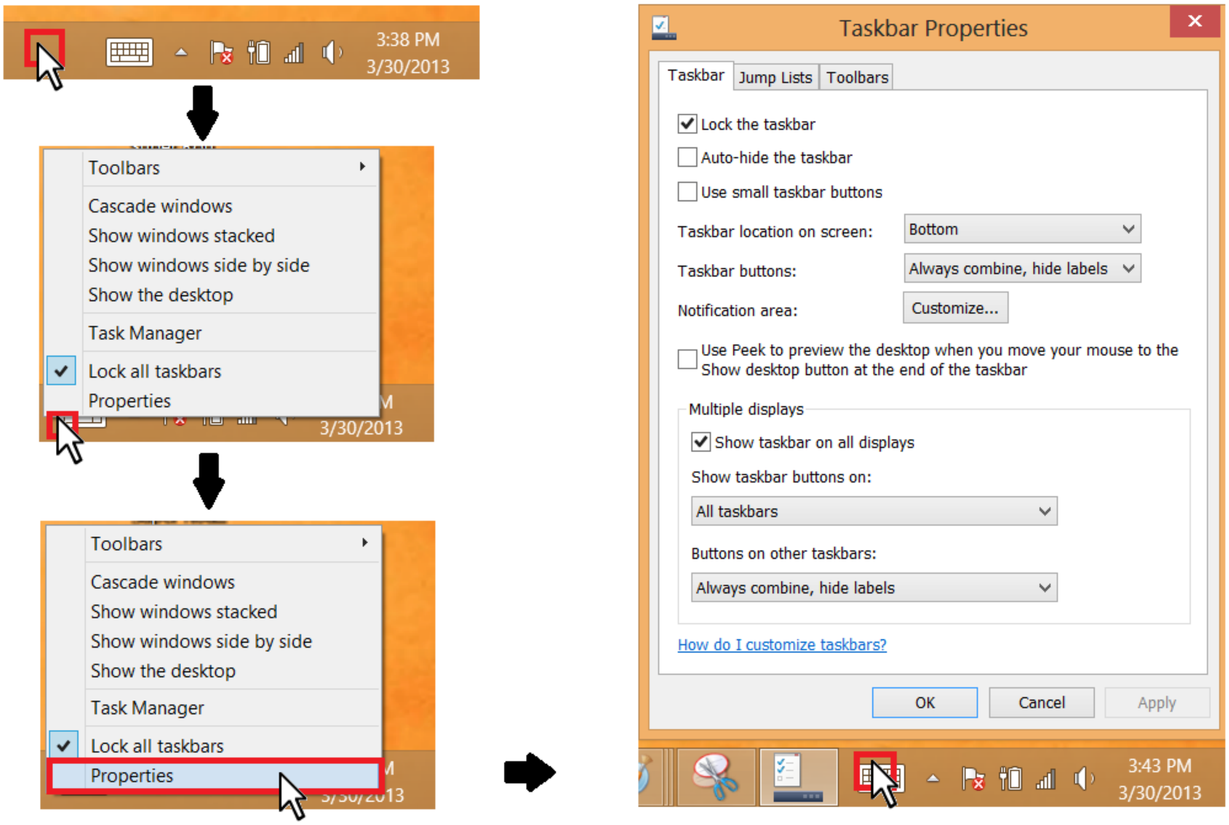

Figure 27 - Chunking and Phrasing [109] in Windows 8

3.2.1 Qualitative Analysis

The *Descriptive Labeling* technique described by Nielsen [90] was used to analyze the gestures collected and classify them. This technique describes gestures by their movement and action – postures, hand/finger positions, hand trajectories, and posture – and not by what a gesture communicates or its purpose (semantic meaning). Chunking and Phrasing by Buxton [109] was a technique used to take into account *atomic* gestures and *compound* gestures. An example of this technique can be shown by examining the events that occur when making a selection in the pop-up menu that appears when right-clicking an empty space in the taskbar in Microsoft Windows 8[10]. In the selection through the pop-up menu, there are a number of sub-tasks (shown in Figure 27) – right-clicking to bring up the pop-up menu, moving the mouse to select

[10] Microsoft Windows 8, http://windows.microsoft.com/en-ca/windows/home

an option and then clicking an option – that are grouped together to form a single logical action. These sub-tasks or actions are called *phrases* and the grouping of all of them together is called *chunking*. Typically, this grouping is related to the psychological process of the action, which in the case of the example discussed, is bringing up a menu and selecting an item. A *phrase* in a gesture is defined by periods of tension and relaxation. Thus, a gesture with a singular *phrase* is considered to be *atomic* and a gesture with *series of phrases* is considered *compound*.

This analysis technique was specifically chosen because it aligns with one of the motivations for this thesis – providing gestures and interactions for content transfer in multi-display environments. Putting an emphasis on the physical actions of a gesture allows user-experience professionals to better describe the nature of the actions and gestures that users are performing, as well as the mental models being drawn upon by users.

3.2.2 Quantitative Analysis

In a user-elicited gesture set, there needs to be a large degree of consensus among a set of gestures created by users. From a user experience perspective, it can be argued that a set of gestures for content transfer in multi-display environments can be limiting for users because they can have different preferences for gestures. However, one of the main goals of this thesis, is to understand the mental and conceptual models of all users and discovering gestures that are common and understandable between them. For this study, it means finding common gestures which have been created independently by different participants in the study for the same content transfer tasks. A large group of common gestures means there is a large degree of consensus among the participants. This metric was previously used by Wobbrock et al. [7] for the creation of a gesture set for digital tabletops and is defined below:

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

In the equation shown above, r is defined as a *referent* in the entire set of referents, represented by R. A referent, or outcomes of a task, is represented in the study by an image being transferred when a participant performs an action, as described in Section 3.1.2. With 16 different tasks and outcomes for participants, the entire set of referents are represented in Figure 20. P_r is defined as the set of gestures proposed for the referent r, while P_i contains the subset of gestures that are identical. A represents the agreement score. An agreement score is a quantitative measure of the variability of gestures produced (qualitative data) and in a single number, determines the degree of consensus for the gestures of a task [7].

To understand the impact that factors such as distance and device type have on the gestures chosen by participants, a statistical model also needs to be applied to the results of the study. Not applying a statistical model, means a quantitative analysis is performed only at an individual level of the data. This provides information such as repeatability of individual gestures and the range of gestures used for an individual participant. Although this is useful information and provided in the analysis of the data for this thesis, it is more useful when developing a machine learning approach to gesture recognition for a multi-display environment. Having a statistical model is more practical for designing a useable gesture set for a large group of people, which was one of the goals of Wobbrock et al [7] and a goal of this thesis as well.

As distance and device type can be considered variables that can impact gestures chosen by participants, a *Generalized Linear Model (GLM)* can be used [110]. The GLM allows for the prediction of a variable (typically called a *dependent* variable) from one or more other variables (typically called *independent* or *explanatory* variables), which in the case of the study are device type, distance and the gestures themselves. One problem with GLM is that it assumes that there is no correlation in the data observed, which in this study, is the gestures [110]. This isn't the case in the study, as gestures are extremely correlated to types of tasks (sending or retrieving), device types and distance. To handle data that is correlated, typically equations are used with GLM. The *Generalized Estimating Equation* (*GEE*) is one such equation that provides a means of determining correlation for longitudinal data that is measured at different points of time or can be classified [111]. In the study, the correlation to be examined is the effects of distance and device on the gestures, and the data is longitudinal as it occurs throughout the duration of the study, as well as classified, using the techniques described in Section 3.2.1.

To specifically examine the effects of distance and device type on the gestures with GEE, a distribution should be used. The distribution used is determined by the type of data in the GEE, which in the case of this study, is counted data. This counted data is determined by counting the number of gestures that have occurred throughout the study and as a result, a *Poisson distribution* is used [112]. With this distribution, it is then possible to determine the impact of specific variables (distance and device type) in GEE. A GEE with a Poisson distribution can thus be used for the examination of different factors such as distance, and device type on content transfer gestures in multi-display environments, which is a goal of this thesis.

3.3 Results

A total of 816 gestures was collected (17 participants × 16 tasks × 3 gestures) from 17 participants and these were classified based on the technique described in Section 3.2.1 (see Appendix D.1 for all gestures as a reference). The atomic gestures were grouped into a set of four conceptual interaction metaphors that address users' mental models of the multi-display environment. Following the grouping of gestures, the amount of agreement for the preferred gestures was calculated using the agreement metrics described in Section 3.2.2 and is discussed in Section 3.3.1. Section 3.3.2 discusses how the conceptual metaphors were derived, the four conceptual metaphors themselves – *Close Contact, Moving Objects, Selection, and Borrowed Interactions* – in detail and provides examples. Section 3.3.3 provides a brief quantitative

Task	Gesture 1	Gesture 2	Agreement Score
Send to Tablet (Near)	Swipe Up	-	0.26
Send to Tablet (Far)	Swipe Up	Touch Five (Pick) + Throw (Drop)	0.33
Retrieve from Tablet (Near)	-	-	0.11
Retrive from Tablet (Far)	Swipe Down	-	0.13
Send to Wall Display (Near)	Swipe Up	Bump	0.2
Send to Wall Display (Far)	Swipe Up	-	0.18
Retrive from Wall Display (Near)	-	-	0.07
Retrieve from Wall Display (Far)	Swipe Down	-	0.1
Send to Tabletop (Near)	Swipe Up	-	0.21
Send to Tabletop (Far)	Swipe Up	-	0.27
Retrieve from Tabletop (Near)	Position On + Swipe Across	-	0.09
Retrieve from Tabletop (Far)	Swipe Down	-	0.08
Send from Tabletop to Wall Display	Touch Five (Pick) + Touch Five (Drop)	-	0.08
Send from Wall Display to Tabletop	Camera	-	0.1
Retrieve from Wall Display with Multiple Images (Near)	Touch Five (Pick) + Throw (Drop)	-	0.1
Retrieve from Wall Display with Multiple Images (Far)	Touch Five (Pick) + Throw (Drop)	-	0.12

Figure 28 – Most common gesture for each task. Generated by removing all gestures in a task with fewer than three observations.

analysis of the metaphor set, while Section 3.3.4 discusses the compound gestures that were observed. Finally, the impact of distance and devices on the gestures is discussed in Section 3.3.5.

3.3.1 Agreement and Gesture Sets

To attempt to create a user-elicited gesture set, there needs to be a significant amount of consensus among the gestures created. The agreement score for this study was calculated using the preferred gestures that were chosen by each participant and the formula described in Section 3.2.2. The calculation of the agreement scores for the preferred gestures resulted in a value of $A_{Favorite}$=0.16.

The agreement score for this study is significantly lower than the agreement scores found by Wobbrock et al [7] in a study of user-defined gestures for digital tabletops. In their study, the agreement scores for one-handed gestures was A_{1H} =0.32 and the two-handed gestures were A_{2H} =0.28. Comparatively, the overall agreement score for this study was $A_{Favorite}$=0.16, showing minimal consensus. This is similar to qualitative results from earlier elicitation studies designed for multi-display environments [94] [95] . A lower agreement score also implies that it is not possible to propose a single user-defined gesture set, unlike prior studies [7]. Figure 28 provides the agreement scores calculated for all 16 tasks. Gestures that were observed less than 3 times for a task were removed, as a significant amount of gestures had minimal observations, which again highlights the lack of consensus amongst the gestures performed by participants.

3.3.2 Conceptual Metaphors

As it was not possible to provide a gesture set due to low agreement scores, a set of conceptual metaphors are instead provided that group the gestures observed throughout the study and provide a basis from which user experience professionals can draw upon for designing new content transfer gestures in multi-display environments. Using the technique described in Section 3.2.1, the gestures were classified based on the following factors (among others):

a) Position of users relative to the devices
b) How devices of users were positioned and/or repositioned.
c) Numbers of fingers or hands used when a user touches a device.
d) Type of motions being made (with or without devices).

Position of Users Relative to Device	
NEAR	A user is close to a device. This is based on the assigned positions in the environment.
FAR	A user is far away from a device. This is based on the assigned positions in the environment.
How devices of users were positioned and/or repositioned.	
POSITION_ON	Performed with a device which is placed down on top of a target and left there.
POSITION_NEAR	Performed with a device, which is placed very close to the target device, but not above or exactly beside or on top of.
POSITION_BESIDE	Performed with a device, which is placed exactly beside, such that the two devices line up, a target device.
POSITION_ABOVE	Performed with a device, with is placed over top or directly above it's target.
Number of fingers or hands used when a user touches a device	
TOUCH_FIVE	Performed on a device with five fingers.
TOUCH_HAND	Performed on a device with a flat hand.
DOUBLETAP_2F	Performed on a device with two fingers, with two distinct touches.
SINGLETAP_2F	Performed on a device with two fingers, momentarily touching.
GRAB_2H	Grab gesture using two hands.
THROW_2H	Throw gesture using two hands.
SWIPE_HAND_UP	Performed on a device, swipe up but performed with a whole hand.
SWIPE_2HAND_UP	Performed on a device, swipe up but performed two whole hands.
SWIPE_2HAND_DOWN	Performed on a device,swipe down but performed with two whole hands.
TOUCH_2H	Performed on a device with two flat hands.

Figure 29 - Possible values for factors (a), (b) and (c)

Types of motions being made	
BUMP	Performed with a device, physically contacting another device for a brief period of time.
POINT	Performed in the air, where a user points with a single finger at the target.
SINGLETAP	Performed on a device with one finger momentarily touching.
POINT_DEVICE	Performed with a device, which is held away from the body in order to point towards a target.
SWIPE_DOWN	Performed on a device, with an arbitary number of fingers, in the downwards direction.
SWIPE_UP	Performed on a device, with an arbitary number of fingers, in the upwards direction.
SWIPE_LEFT	Performed on a device, where the users makes a swipe in the left direction .
SWIPE_RIGHT	Performed on a device, where the users makes a swipe in the right direction.
SWIPE_ACROSS	Performed over two devices, a swipe which continues across a second device.
SWIPE_TOWARDS	Specific to instances where the user has placed a device down, the user swipes toward the direction of the placed down device.
SWIPE_AWAY	Specific to instances where the user has placed a device down, the user swipes away from the placed down device.
SHAKE_DEVICE	Performed with a device which is shaken in a rapid manner.
ROCKING_DEVICE	Performed with a device which is moved up and down in a rocking motion along it's horizontal axis.
CHUCKING_DEVICE	Performed with a device, which is moved in a manner similar to the action required to throw an object.
POUR_DEVICE	Performed with a device, which is rotated slight, paused momentarily and restored original orientation.
SCOOP_DEVICE	Performed with a device by by mimicking a scooping motion.
SCRAPE_DEVICE	Performed with a device, which is tilted at angle and dragged slowly in a scraping like motion.
ROTATE_DEVICE	Performed with a device with is rotated for a longer period of time.
GRAB	Performed in the air, a gesture where the user mimics grabbing an object.
BLOW	Performed when a user blows on the device.
SWEEP	Performed in the air, a gesture with mimics sweeping across a wide area.
PUSH	Performed the air, a gesture where the user appears to push and object with their hands.
SNAP	Performed in the air, a gesture where the user snaps their fingers.
WAVE	Performed in the air, a gesture where the user makes a flat hand and rocks it side to side.
PINCH	Performed on a device, with two finger motion which closes until fingers are touching, mimics a pinch.
KNOCK	Performed on a device, the users takes their fists and strikes the device.
THROW	Performed in the air, a gesture mimic the action of tossing or throwing an object.
CAMERA	Performed by holding a device vertically, with the screen towards the user, mimic the action of taking a picture with a camera.
MIRROR	Performed with a device which is held vertically and inverted, exposing the screen to the target.
LASSO	Performed in the air, a gesture were the user closes a distinct loop in the air with a finger.
TOUCH_LASSO	Performed on a device where the users creates a near circle or lasso with a single finger.
TOUCH_BOTH	Performed on two devices, where the users simulatenouly performs a touch on both devices.
DOUBLETAP	Performed on a device one finger, with two distinct touches.

Figure 30 - Possible values for factor (d)

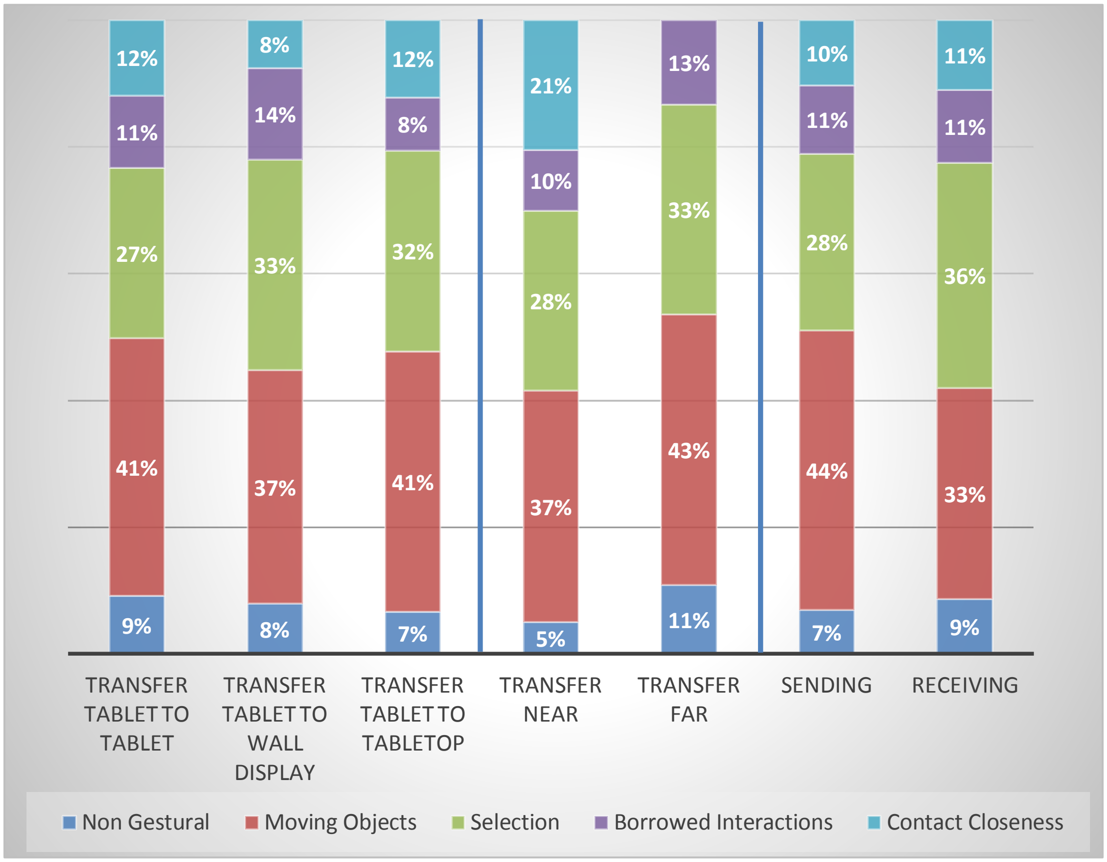

Figure 31 - Percentage breakdown for metaphors for device conditions

All possible values for the gestures being coded according to these different factors are highlighted in Figure 29 and Figure 30. With these factors in mind when coding the gestures in an iterative process, four distinct conceptual interaction metaphors were created that grouped similar gestures together. They are defined and discussed in the upcoming subsections.

3.3.2.1 Close Contact

When the positions of devices are close or there is contact during a gesture.

As per the definition above, these gestures incorporated either physical contact, closeness, or both for devices. Closeness in this metaphor specifically refers to how close a device is to another device. In the study, closeness occurred for participants when they were in the near position (as shown in Section 3.1) for a task and close to a target device. As close contact gestures required a secondary device, these

Close Contact Metaphor		
Gesture Name	# of times observed	% occurred
Bump	36	38%
Position On	26	27%
Position Near	12	13%
Position Beside	11	12%
Position Above	10	11%

Figure 32 - *Close Contact* gesture occurrences (throughout study)

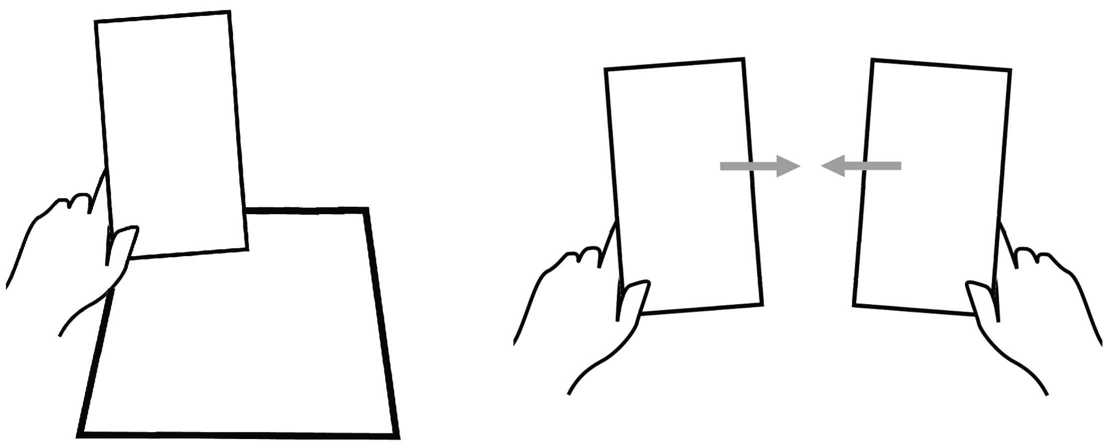

Figure 33 - Examples of gestures from the *Close Contact* metaphor. (a) *Position Above* and (b) *Bump*

occurred with tasks involving tablets. Participants could either choose to make contact with a target device or position their device close to the target device in a variety of ways.

Figure 33 highlights a selection of gestures from the *Close Contact* metaphor. The full set for this metaphor are provided in Appendix C.1 for reference. The variations of the gestures performed for this metaphor and their occurrences is highlighted in Figure 32. *Bump* (see Figure 33b) was the most frequent gesture occurring in this theme at 38%, while Position Above (see Figure 33a) was the lowest at 10%.

A limitation of these types of gestures is the proximity they require, as being physically close to a device is not generalizable across different distances. Examples of gestures in the *Close Contact* metaphor include *Position Above*, where participants physically place one device directly over top of another, or *Bump*, where physical contact is made with another device for a short period of time.

Moving Objects Metaphor		
Gesture Name	# of times observed	% occurred
THROW	99	22%
SWIPE_UP	96	22%
GRAB	67	15%
SWIPE_DOWN	45	10%
SHAKE_DEVICE	30	7%
SWIPE_ACROSS	21	5%
ROCKING_DEVICE	16	4%
CHUCKING_DEVICE	12	3%
GRAB_2H	8	2%
SWIPE_LEFT	6	1%
THROW_2H	6	1%
SWIPE_HAND_UP	5	1%
SWEEP	5	1%
POUR_DEVICE	4	1%
SCOOP_DEVICE	4	1%
SCRAPE_DEVICE	4	1%
SWIPE_2HAND_UP	4	1%
ROTATE_DEVICE	3	1%
SWIPE_RIGHT	3	1%
PUSH	3	1%
SWIPE_TOWARDS	2	0%
SWIPE_2HAND_DOWN	1	0%
SWIPE_AWAY	1	0%
BLOW	1	0%

Figure 34 - Moving Objects gesture occurrences (throughout study)

3.3.2.2 Moving Objects

When a gesture mimics the real-world actions of moving an object in physical space.

The *Moving Objects* metaphor mimics the interactions of moving objects in physical space. Prior experience that participants had with interactions that involve physical objects in the real word, such as sliding, flicking or throwing led to 24 variations for this metaphor, as shown in Figure 34. *Throw* and *Swipe Up* were the most common gestures at 22% each, while *Swipe Towards, Swipe 2Hand Down, Swipe Away and Blow* were the least common.

Examples of these gestures include the *Pour* gesture, where a user mimics pouring content onto another device (shown in Figure 36) and the *Chucking* gesture, where a user simulates the tossing of a device (shown in Figure 35). The full set of gestures for this metaphor are provided in Appendix C.4 for reference.

Many participants also stated that repeatedly performing these types of gestures was fatiguing. This is due to the fact that many gestures in the *Moving Objects* metaphor require physical actions to be completed, and as with any activity that is physical and repeated over a period of time, fatigue becomes an issue.

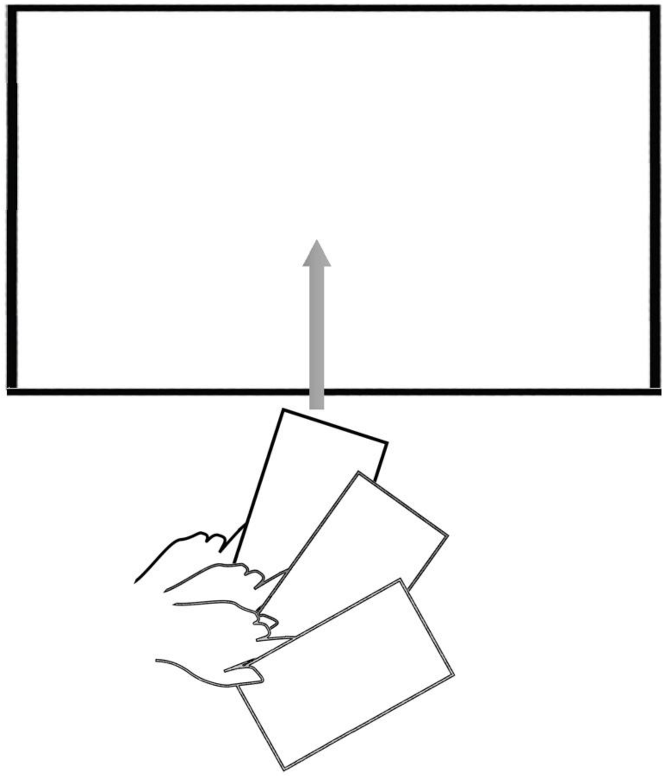

Figure 35 - *Chucking* gesture with a tablet towards a wall display, from the *Moving Objects* metaphor

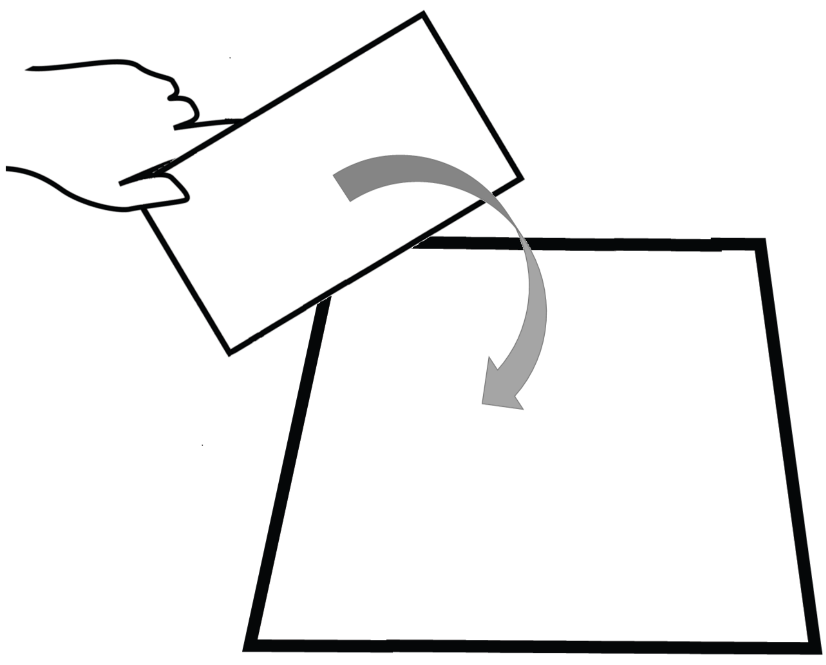

Figure 36 - *Pour* gesture with a tablet to a digital tabletop, from the *Moving Objects* metaphor

Selection Metaphor		
Gesture Name	# of times observed	% occurred
TOUCH_FIVE	117	30%
POINT	106	27%
SINGLETAP	63	16%
POINT_DEVICE	29	7%
TOUCH_HAND	20	5%
DOUBLETAP	19	5%
DOUBLETAP_2F	9	2%
SINGLETAP_2F	8	2%
TOUCH_LASSO	7	2%
SNAP	6	2%
WAVE	4	1%
TOUCH_BOTH	3	1%
TOUCH_2H	1	0%

Figure 37 - Selection gesture occurrences (throughout study)

3.3.2.3 Selection

When a gesture is used to actively select another object, which can include other persons or devices.

The *Selection* metaphor was used by participants to select either objects, other devices, or other participants. This was done typically through methods such as touching (with a finger, a whole hand, or two hands) among others. 13 different variations occurred for the *Selection* metaphor with *Touch Five* being the most common gesture at 30% and *Touch 2H* being the least common, as shown in Figure 37.

Examples of this gesture include the *Point Device* gesture where users would also point at a target device or person with their own device (shown in Figure 39) and the *Point* gesture where participants would point at a target device or person (shown in Figure 38). The full set of gestures for this metaphor are provided in Appendix C.3 for reference. An interesting observation for the gestures in this metaphor, was that they were typically a part of compound gestures, which are discussed in Section 3.3.3.

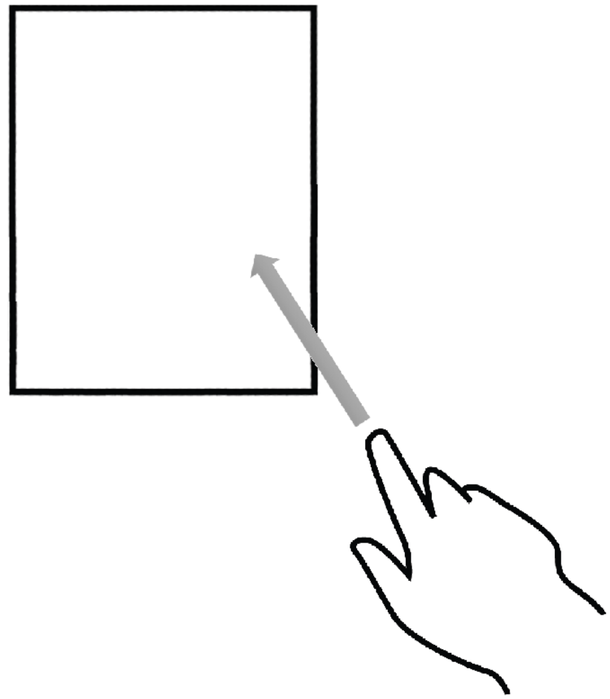

Figure 38 - *Point* gesture from the *Selection* metaphor

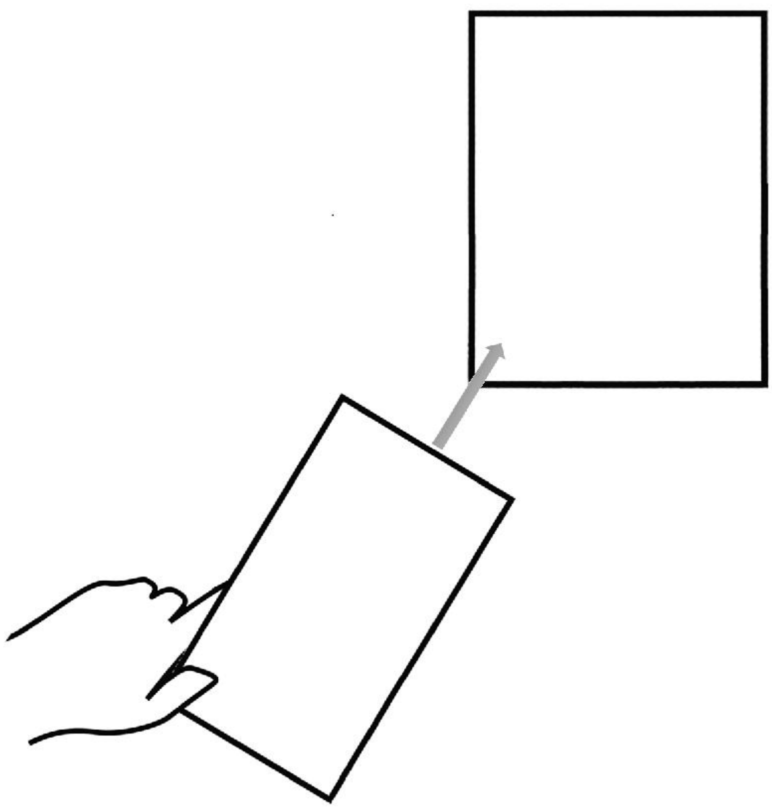

Figure 39 - *Point Device* gesture with a tablet towards a wall display, from the *Selection* metaphor

Borrowed Interactions Metaphor		
Gesture Name	# of times observed	% occurred
CAMERA	52	42%
MIRROR	51	41%
LASSO	11	9%
PINCH	9	7%
KNOCK	1	1%

Figure 40 - *Borrowed Interactions* gesture occurrences (throughout study)

3.3.2.4 Borrowed Interactions

When a gesture mimics or borrows interactions that already exist with other technologies or physical objects in the real-world.

The gestures that didn't fit in the *Close Contact*, *Moving Objects,* and *Selection* metaphors were clearly borrowed from users' previous experiences with existing technologies or physical objects in the real-world. This is a distinct change from the prior metaphors, which were based on activities and interactions from everyday life. There were 5 different variations observed for this metaphor, with *Camera* being the most common gesture at 42% and *Knock* being the least at 1%, as shown in Figure 40. An example of a gesture that was borrowed from previous experiences, is the *Camera* gesture where a user would hold a device in a fashion similar to that of holding a camera (see Figure 42). The *Camera* gesture was typically used to retrieve data while the *Mirror* gesture was used for both sending and retrieving data. The *Mirror* gesture, is performed by reversing the screen of a device and exposing it towards a target device (see Figure 41).

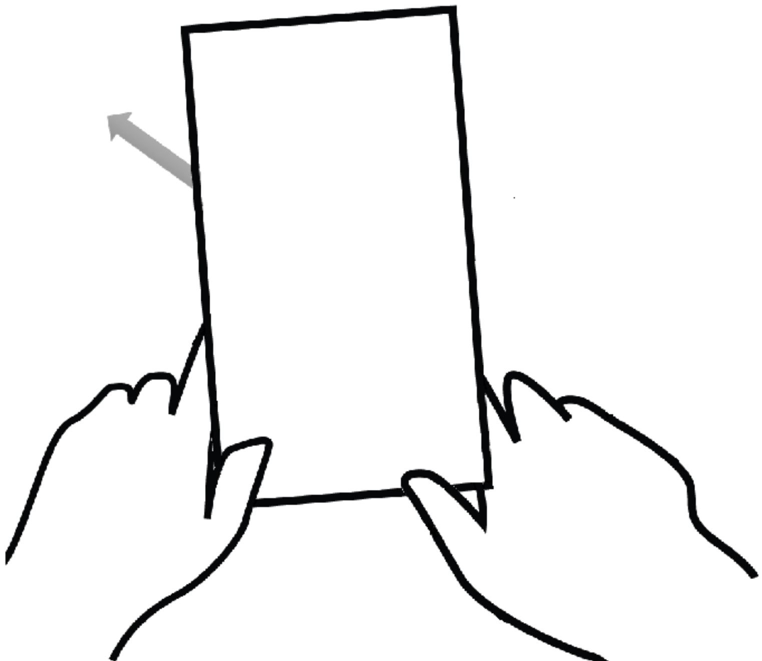

Figure 41 - *Mirror* gesture with a tablet, from the *Borrowed Interactions* metaphor

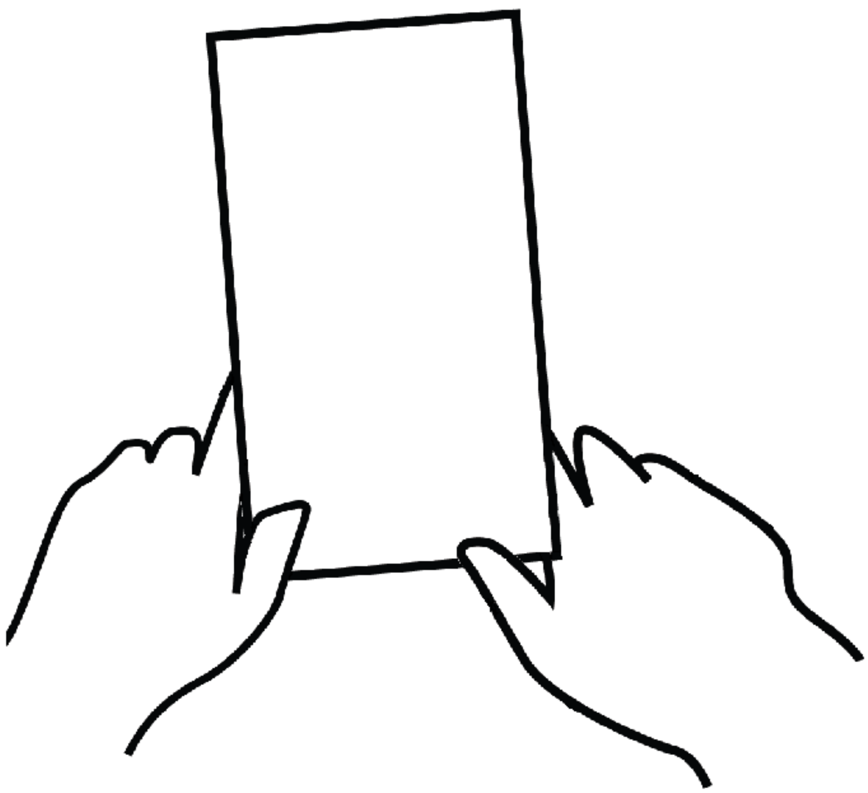

Figure 42 - *Camera* gesture with a tablet, from the *Borrowed Interactions* metaphor

Metaphor Name	# of Gestures Total	% occurred
Moving Objects	446	42%
Selection	392	37%
Borrowed Interactions	124	12%
Close Contact	95	9%

Figure 43 – A comparison of all the metaphors in the metaphor set

3.3.3 Comparing the Conceptual Metaphors

Figure 43 highlights a comparison of the metaphors and gestures found in the study. Noticeably, the *Moving Objects* and *Selection* metaphors are the most common metaphors observed, at 42% and 37% respectively. This indicates that participants tended to draw upon experiences that were more natural as opposed to drawing upon borrowed experiences (such as with technologies), highlighted by the *Borrowed Interactions* metaphor only comprising 12% of the overall metaphor set.

It is also interesting to note that the *Close Contact* metaphor was only 9% of the overall set. As 7 of the 16 tasks listed in Figure 20 were at close distances to either other devices or a secondary researcher, this suggests that participants weren't comfortable making physical contact or using gestures based on physical proximity.

3.3.4 Compound Gestures

In some of the gestures observed by the participants, distinct gestures were performed in succession, which is considered to be *compound* by Buxton [109] as described in Section 3.2.1. Some of the gestures performed by users followed a pattern that considered multiple users and these were also considered *compound* gestures. This pattern, called *Two-Party Interactions*, typically occurred when a participant was required to perform a task involving a secondary researcher's personal device (tablet). For example, in tasks where an image was to be retrieved from another participant's tablet, participants mentioned that their actions should involve another participant's device and that the transfer should be acknowledged by the other participant instead of simply occurring. A participant performing a *Mirror* gesture with their tablet and then asking another participant to perform a *Mirror* gesture, is an example of the *Two-Party*

Interactions pattern. In this example, the distinct atomic components are the *Mirror* gestures and the *Two-Party* aspect comes from the two participants involved in the interaction.

Another commonly observed compound gesture pattern mimicked the *pick-and-drop* metaphor from Rekimoto [33]. This pattern, called the *Pick & Drop Interaction* pattern, involved a distinct gesture for selecting content on a *source* device and then another distinct gesture for selecting on a *target* device. Unlike *Two-Party* interactions where tasks involved personal devices, tasks for the *Pick & Drop Interaction* involved all devices and gestures were extremely varied and influenced by distance. An example of this pattern is when a participants touches all five fingers (*Touch Five* gesture) on a tablet to select an image, and then follows with a *Throw* gesture towards a target device. A participant would perform this gesture when the target device is far away; however, if the target device is close, then the same participant performs a five finger touch (*Touch Five* gesture) on the tablet, followed with a five finger touch (*Touch Five* gesture) on the target device. This highlights the notion of distance playing a factor in gestures, and is discussed in further detail in the next section. Overall, *compound* gestures composed 43% of the total gestures performed for commands.

3.3.5 Impact of Factors

Another goal of this study was to examine the impact that certain factors had on the choice of gestures for participants. These factors were:

1. *Device Type – Does the device used play a role in the choice of gesture?*
2. *Distance – Does the distance at which the task is performed play a role in the choice of gesture?*

To focus the qualitative analysis on these two factors, the gestures that users identified as their favorite in the study were used with the *Generalized Estimating Equations* (GEE) and a *Poisson distribution*, as described in Section 3.2.2.

3.3.5.1 The Impact of Device Types

After analyzing the data with GEE, it was found that only the *Close Contact* metaphor was statistically significantly impacted by the overall device factor ($\chi 2$ (2)=15.144, p=0.001). As a result, a closer look was taken at the impact of the individual device types – iPad, wall display, and tabletop. The results of this closer analysis of the *Close Contact* metaphor showed that it was negatively, but not significantly, predicted by the iPad device type (B=-3.26, p=0.117) unlike the wall display device type, which was significantly and negatively predicted (B=-0.622, p<0.001). This indicates that the *Close Contact* metaphor should be used for the suitable device types, as it is simply not possible to achieve contact with devices at certain distances (i.e. wall displays and digital tabletops).

3.3.5.2 The Impact of Distance

GEE showed that distance did have a significant statistical impact on the gestures performed by users. Specifically, the *Borrowed Interactions* and *Non-Gestural* metaphors were significantly and positively predicted by far distances (B = .253, p = 0.045 and B = 0.799, p = 0.001), unlike the *Close Contact* metaphor which was significantly and negatively predicted by far distances (B=-3.773, p<.001). This highlights an importance of gestures and interactions in multi-display environments, as the environment typically encompasses a room and content transfer gestures that work on all displays might remove some of the disparities found.

3.4 Discussion

One of the primary goals of the study was to elicit a set of gestures for moving content in multi-display environments. As a result, the study was designed around prior procedures for eliciting gesture sets [7]; however, unlike those studies, the results showed minimal agreement. The implications of low agreement are discussed in Section 3.4.1. Despite this low agreement however, the set of interaction metaphors presented in Section 3.3.2 can provide a basis for user experience professionals to create new content transfer gestures, and applying them to real world multi-display environments is discussed in Section 3.4.2.

3.4.1 Implications of Low Agreement Scores

An agreement score, as mentioned earlier, is a quantitative measure of the variability of gestures produced (qualitative data) and ultimately determines the level of consensus [7]. High agreement would mean that participants tended to produce the same or similar gestures for the same command while low agreement would indicate that participants tended to produce different gestures for the same command. Most tasks in the present study had low agreement scores. As a result, it was not possible to create a set of user-defined gestures for the tasks presented in the study.

Prior elicitation studies for multi-display environments noted little consensus in gestures produced for tasks [55] [94]; however, these studies measured consensus qualitatively through questionnaires. This study adds value to previous studies in that it was able to quantitatively confirm this lack of consensus. This is in contradiction with elicitation studies that have been done for digital tabletop gestures in which higher levels of agreement were found [7] [113]. The difference might be explained as follows:

Interaction Possibilities:

A primary difference between interactions in a multi-display environment and interactions on a digital tabletop is the bounds of interaction. At a digital tabletop, gestures occur on or near the tabletop, typically horizontally or vertically. In essence, they are restricted to specific regions of the digital tabletop. This is not the case in a multi-display environment, where numerous displays and devices can exist. Not only are the restrictions increased with each device and display in the environment, but interactions are no longer necessarily horizontal or vertical, as displays and devices can be mobile or of different sizes. Interactions are now also possible without the constraint of devices and displays as well. With multiple displays and devices, it is also possible to treat the space between them (*displayless space* [3] from Section 2.3) in different ways. This means that users have a wider space to perform gestures in the environment.

Participant Background and Experience:

Participants in the study had a wide variety of experience with the technologies involved in multi-display environments, as shown earlier in Figure 25 and discussed in Section 3.1.1. As a result, it is natural that as a whole, participants did not have a consistent mental model for multi-display interactions simply because they had varying degrees of experience with technologies in multi-display environments. The interaction metaphors provided in Section 3.3.2 provide user experience professionals with knowledge to draw upon when creating new content transfer gestures for multi-display environments, despite no consistent mental model for users, as it provides a grouping for a wide variety of conceptual models.

3.4.2 Recommendations for Applying Gestures and Metaphors

With the large number of gestures created for the study and the resulting interaction metaphors presented, user experience professionals can utilize these metaphors in designing content transfer interactions for multi-display environments in a number of different ways. Recommendations for user experience professionals are described in the following subsections.

3.4.2.1 Aliasing Gestures

Aliasing gestures allows for multiple gestures to support a specific task. An example of aliasing is highlighted by the task to *save* a document in Microsoft Word 2013[11]. To *save* a document, a user has several methods to complete this task. One method is the Control-S keyboard shortcut, another is to click the save button in the toolbar and another is to select the save option from the File menu, as shown in Figure 44. Providing multiple options for a user to save a document, provides them with choices in

[11] Microsoft Word 2013, http://office.microsoft.com/en-us/word/

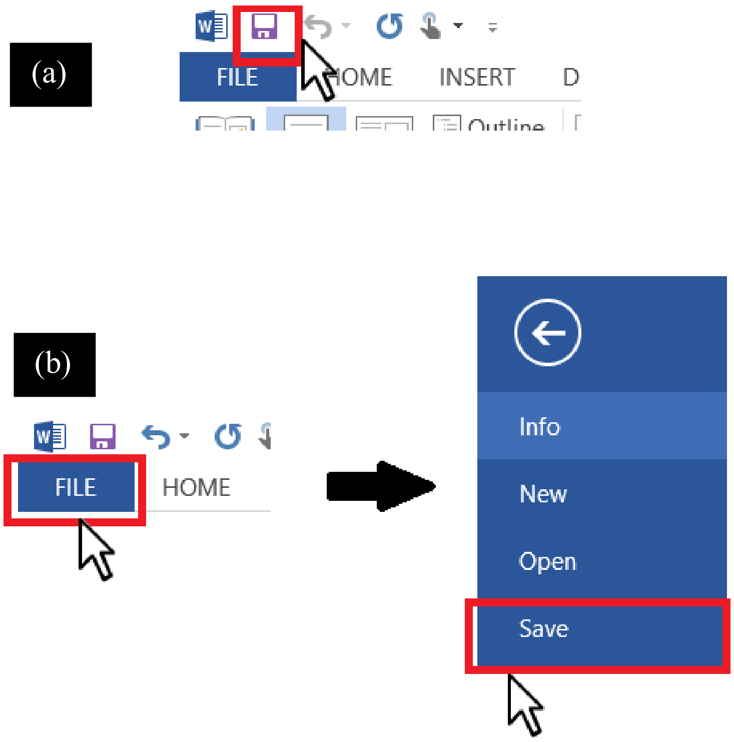

Figure 44 – 2 aliased options for saving a document in Microsoft Word 2013. (a) Using a toolbar option (b) Using the File menu option

whichever option they feel most comfortable with, quickest to use or easiest to perform, which are goals of user experience.

This technique can also be applied to gestures in multi-display environments as a way of handling the different conceptual models users have, which was illustrated by the lack of agreement discussed in Section 3.4.1. When analyzing the gestures in the study, no instances were found where gestures conflicted with one another in terms of their meaning. This indicates that despite different gestures having different meanings, different gestures were chosen by users for the same task. An example of this was seen with the *Position On*, *Position Near*, *Position Beside* and *Position Above* gestures, which were expected by users to produce similar results, despite the meaning of each gesture being inherently different.

This was also seen with several different gestures such as *Chucking*, *Swiping* and *Throwing*, despite the atomic nature of the gestures (with-device, on-device, and without-device, respectively) being very different. As a result, providing an alias for the gestures would allow user experience professionals to accommodate divergent conceptual models for users in multi-display environments. Aliasing gestures can also serve as a solution for the fatigue issue mentioned by users with gestures in the *Moving Objects* metaphor.

3.4.2.2 Gestures and Metaphors

The nature of the atomic gestures that were observed in the study served as the basis for designing the metaphors presented in Section 3.3.2. These metaphors are also based upon implicit understandings of spatial relationships in an environment – objects are independent from one another, have location, and can (sometimes) move [73].

The metaphors serve as a conceptual guide for user experience professionals to allow them to design gestures that are more easily discoverable and understandable, as they categorize a multitude of conceptual models users can have in multi-display environments.

3.4.2.3 Different Gestures for Distance and Device

The results presented in Section 3.3.5 showed that distance and device were important factors in users choosing gestures in multi-display environments. As a result, user experience professionals can opt for metaphors that map to distances and device type or choose independent gestures entirely. For example, users in the study often chose gestures in the *Borrowed Interaction* metaphor for far distances and gesture metaphors from the *Close Contact* metaphor when at a closer distance. A user experience professional could then choose to alias gestures, as discussed in Section 3.4.2.1, and use a *Camera* gesture (among others) for far distances, a *Bump* gesture for close distances, and a device- and distance-independent gesture (such as *Swipe Up*) for content transfer interactions in a multi-display environment. *Close Contact* gestures weren't utilized when a faraway wall display was the target device, which indicates that different gestures or metaphors for content transfer should be used for a multi-display environment where a wall-display is a primary component.

3.5 Study Limitations

In the user-elicitation study, participants were frequently encouraged to create gestures and interactions without regard for technological limitations. Despite this freedom, it appeared that their current and existing experiences with other technologies either influenced or limited the types of gestures they created. This was seen with the *Borrowed Interactions* interaction metaphor, where participants mimicked interactions from existing experiences with technologies such as digital cameras and objects such as mirrors.

It is also recognized that more comprehensive user studies need to be performed that involve a wider range of participants. As mentioned in Section 3.1.1, the average age of male participants and female participants were 27 and 25 years, respectively, which doesn't represent a large population of potential users for multi-display environments, which suggests that the findings from the study can't be generalized.

Likewise, many of the participants had prior experience with all or some of the technologies in multi-display environments, as indicated in Section 3.1.1. As the target users for multi-display environments should have any level of experience with the associated technologies, it makes sense for a wider range of experience for users in the study. This means that the study provided limited insights into users with little to no experience in technologies related to multi-display environments and accordingly how factors such as distance and device affected them when performing interactions for transferring content.

Another limitation of the study was the nature of the multi-display environment itself. For the purposes of the study, common multi-display hardware was used, however, other technologies such as mobile phones can potentially be used in interactions for transferring content. Furthermore, multi-display environments can exist with simply one of the technologies used in this study (i.e. just digital tabletops, just wall displays or just tablets). However, given that multi-display environments are still relatively new and a goal of the thesis was to provide user experience professionals a set of interactions for content transfer tasks and guidelines, the findings can be used to inform the design of future gestures and interactions in multi-display environments. Undoubtedly, given a wider range of participants and multi-display environments, more robust and generalizable insights about the interaction metaphors and guidelines used by user experience professionals can be provided.

3.6 Conclusion

The study presented in this chapter was designed to elicit content transfer gestures from users of multi-display environments. To elicit gestures, the study focused on tasks related to content transfer as such tasks are extremely frequent and form the basis of interactions for users in multi-display environments. Quantifying the level of agreement with the gathered gestures indicated that it was not possible to extract a single gesture set for user-experience professionals to use. As a result, four conceptual interaction metaphors were presented to form a basis upon which user-experience professionals can create gestures and interactions for multi-display environments. Also presented in this chapter are the implications for user-experience professionals when considering distance and device types in creating content transfer gestures for users. These implications lead to recommendations that user-experience professionals can consider when creating content transfer interactions for multi-display environments.

Chapter Four: Conclusion & Future Work

The work discussed in this thesis examined interactions for content transfer tasks, which are common for users in multi-display environments. This provides insight for future user experience professionals and system designers, so they can improve the usability and user experience of multi-display environments. Chapter 1 described the motivations and research questions behind this thesis work. An overview of multi-display environments, designing interactions for them, and the different types of multi-display environments are provided in Chapter 2. This was to provide background that is required to understand the challenges that user experience professionals face when designing content transfer interactions for users in multi-display environments. From this understanding, a user-elicitation study was then performed to answer the research questions presented in Section 1.5 and provide insight into improving user experience for multi-display environments. This was presented in Chapter 3.

4.1 Contributions

The first contribution of this thesis is an overview of the current research space of interaction design in multi-display environments. This is provided in the form of a literature review that discusses the background of multi-display environments, existing multi-display interactions, and the different types of multi-display environments. Also included is a categorization of past interaction techniques for multi-display environments. The categorizations are based upon whether or not graphical menus, physical approaches or gestures that may or may not involve devices, are used. This categorization should help user experience professionals determine what challenges a new interaction technique may face, based on these prior interactions. This answered the first research question, "*What is the current state of research in multi-display environments, particularly within the context of interaction design?*" that was directed towards understanding current research in multi-display environments and interaction design.

The second contribution of this thesis is its exploration of user experience for multi-display environments through investigations into the four remaining research questions presented in Section 1.5. The second research question, "*What are the impacts of the spatial characteristics of the environment on interactions for tasks involving content transfer in multi-display environments?"* was investigated in the user-elicitation study presented in Chapter 3 and discussed in detail in Section 3.3.5. The results from the study showed that distance had a significant impact on interactions for content transfer tasks in multi-display environments, particularly on interactions that required users to make physical contact with different displays. The third research question, *"What are the impacts on content transfer interactions based on the characteristics/affordances of the different displays and devices in a multi-display environment?"* was also discussed in Section 3.3.5. The elicitation study similarly showed that interactions that required closeness or physical contact are impacted significantly, particularly since some of these interactions are display-specific. The answers to these two questions demonstrate that interactions for content transfer tasks do change based on distance and display type. As a result, a set of recommendations is provided in Section 3.4.2 for user experience professionals to utilize when designing interactions for content transfer tasks in multi-display environments.

The fourth research question, *"To what extent is the conceptual model of interaction for users based upon spatial characteristics or device characteristics?"* and fifth research question "*Prior research has produced several gesture sets for digital tabletop interaction, but can gestures be elicited for multi-display environments in a similar fashion?"* are co-dependent and are also investigated in the user-elicitation study presented in Chapter 3. The fifth research question revealed that it was not possible to create a single, user-elicited interaction set for multi-display environments. This was due to a wide variety of interactions created by users and indicated that no consistent mental or conceptual model exists for users in multi-display environments. However, a categorization of these interactions did lead to the creation of a set of interaction metaphors provided in Section 3.3.2 that can guide user experience professionals in the design of new interactions for transferring content in multi-display environments. Many of the interactions in this set of interaction metaphors were influenced by spatial and device characteristics, both directly or implicitly, which answers the fourth research question.

4.2 Future Work

Multi-display environments in the context of the definition used in this thesis are still a relatively new research area, particularly with respect to user experience. As this thesis focused on a specific part of user experience in a multi-display environment and interactions for content transfer, a number of directions could be taken for future work. They are as follows:

The first direction for future work is to provide an evaluation of the interaction metaphor set provided in Section 3.3.2. This evaluation can follow the work by Morris et al. [15] and evaluate the interaction metaphor set either by working closely with user experience professionals to create new interactions or evaluating selected interactions for content transfer in the interaction metaphor set. These interaction metaphors could also be incorporated into real-world, multi-display environments and can be qualitatively assessed with users and multiple case studies. This would allow for real-world user feedback that can inform the design of future multi-display environments.

A second possible direction for future work is to examine the performance measures of the interaction metaphor set. This can be done by comparing user preference in the interaction metaphor set and performances measures such as distance, speed, accuracy and display type for moving content across the different displays in a multi-display environment. This examination can establish the efficiency and usefulness of certain interactions for moving content in multi-display environments and thus inform future interaction design standards.

Examining the impact of multiple users and collaboration on these interaction metaphors is a third interesting direction for future research. The scope of this current thesis research was limited to single or two-user interactions for content transfer in multi-display environments, but examining the impact of multiple users in a multi-display environment and the social dynamics the interactions entail can lead to meaningful results for designing group interactions.

4.3 Conclusion

Multi-display environments are becoming increasingly common. Nearly all technology tradeshows present technologies that are integrated into and considered parts of multi-display environments, much like the upcoming consoles by Microsoft[12] and Sony[13]. To maximize the inherent potential these environments contain in a wide variety of fields – such as medicine, oil and gas, and emergency response planning – it is important to have useful interactions that make sense not only to the users but also for the tasks they are performing. These interactions are the first step in creating interfaces that are designed for multi-display environments.

This thesis presented an exploration into user experience in multi-display environments by focusing on common task that users perform in multi-display environments – moving content across different displays and devices. The impacts that factors such as distance and type of device have on content transfer interactions for users, are discussed. Also provided is a set of interaction metaphors for user experience professionals to use, as well as approaches to deal with the factors mentioned previously.

Using previous research into multi-display environments, this thesis work provides information for future user experience professionals and researchers in the field of multi-display environments to ultimately produce more useful computer tools and transform man's environment, as stated by Victor Papanek in the first line of this thesis.

---

[12] Microsoft, www.microsoft.com
[13] Sony, www.sony.com

References

[1] Victor Papanek, *Design for the Real World: Human Ecology and Social Change*. New York, USA: Van Nostrand Reinhold Co., 1963.

[2] Tony Gjerlufsen, Clemens Nylandsted Klokmose, Clément Pillias, and Michel Beaudouin-Lafon, "Shared substance: developing flexible multi-surface applications," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'11)*, Vancouver, BC, Canada, 2011, pp. 3383-3392.

[3] Miguel Nacenta, Carl Gutwin, Dzimitri Aliakseyeu, and Sriram Subramanian, "There and Back again: Cross-Display Object Movement in Multi-Display Environment," *Journal of Human-Computer Interaction*, vol. 24, no. 1, pp. 170-229, 2009.

[4] Chia Shen, Alan Esenther, Clifton Forlines, and Kathy Ryall, "Three Modes of Multi-Surface Interaction and Visualization," in *In Information Visualization and Interaction Techniques for Collaboration across Multiple Displays Workshop associated with CHI'06 International Conference*, Montreal, Quebec, Canada, 2006.

[5] Michael A. Grasso, David S. Ebert, and W. Timothy Finn, "The integrality of speech in multimodal interfaces," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 5, no. 4, pp. 303-325, December 1998.

[6] Norbert A. Streitz et al., "i-LAND: an interactive landscape for creativity and innovation," in *Proceedings of the SIGCHI conference on Human Factors in Computing Systems (CHI'99)*, Pittsburgh, Pennsylvania, USA, 1999, pp. 120-127.

[7] Jacob O. Wobbrock, Meredith Ringel Morris, and Andrew D. Wilson, "User-defined gestures for surface computing," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'09)*, Boston, MA, USA, 2009, pp. 1083-1092.

[8] Gord Kurtenbach and Eric Hulteen, "Gestures in Human-Computer Communication," in *The Art and Science of Interface Design*. Reading, Massachusetts, USA: Addison-Wesley Publishing Company, 1990, pp. 309-317.

[9] Mark Billinghurst, "Gesture Based Interaction," in *Human Input to Computer Systems: Theories, Techniques and Technology*., 1999.

[10] Donald Norman, Jim Miller, and Austin Henderson, "What you see, some of what's in the future, and how we go about doing it: HI at Apple Computer," in *Conference Companion on Human Factors in Computing Systems (CHI'95)*, Denver, Colorado, USA, 1995, p. 155.

[11] Usability.gov. Contextual Interview. [Online]. http://www.usability.gov/methods/analyze_current/learn/contextual.html

[12] Jacob Nielsen, *Usability Engineering*.: Morgan Kaufmann, 1993.

[13] C. Bamum, *Usability Testing and Research*. New York: Pearson Education, 2002.

[14] Christopher D. Wickens, Sallie E. Gordon, and Yili Liu, *An Introduction to Human Factors Engineering*, 2nd ed. Upper Saddle River, NJ, USA: Pearson Education, Inc., 2002.

[15] Meredith Ringel Morris, Jacob O. Wobbrock, and Andrew D. Wilson, "Understanding users' preferences for surface gestures," in *Proceedings of Graphics Interface 2010 (GI'10)*, Ottawa, Ontario, Canada, 2010, pp. 261-268.

[16] Department of Electrical Engineering and Computer Sciences, University of California, Berkeley. (2013, February) Mark Weiser Biography. [Online]. http://www.cs.berkeley.edu/Weiser/bio.shtml

[17] Mark Weiser, "The computer for the 21st Century," *IEEE Pervasive Computing*, vol. 1, no. 1, pp. 19-25, 2002.

[18] International Data Group (IDG), "How Mobility is Disrupting Technology and Information Consumption," 2012.

[19] Teddy Seyed, Chris Burns, Patrick King, Mario Costa Sousa, and Frank Maurer, "MRI Table Kinect: A multi-surface application for exploring volumetric medical imagery," in *Proceedings of the Workshop on Safer Interaction in Medical Devices (MediCHI'13)*, Paris, France, 2013.

[20] Wikipedia. (2013, February) Cyworld. [Online]. http://en.wikipedia.org/wiki/Cyworld

[21] Douglas C. Engelbart and William K. English, "A research center for augmenting human intellect," in *Proceedings of the December 9-11, 1968, fall joint computer conference, part I (AFIPS '68)*, San Francisco, California, 1968, pp. 395-410.

[22] Peter Cook, Clarence Ellis, Mike Graf, Gail Rein, and Tom Smith, "Project Nick: meetings augmentation and analysis," *ACM Transactions on Information Systems (TOIS)*, vol. 5, no. 2, pp. 132-146, April 1987.

[23] M. Stefik, D. G. Bobrow, S. Lanning, D. Tatar, and G. Foster, "WYSIWIS revised: early experiences with multi-user interfaces," in *Proceedings of the 1986 ACM conference on Computer-supported cooperative work (CSCW'86)*, Austin, Texas, 1986, pp. 276-290.

[24] Mark Stefik et al., "Beyond the chalkboard: computer support for collaboration and problem solving in meetings," *Communications of the ACM*, vol. 30, no. 1, pp. 32-47, January 1987.

[25] Masayuki Tani, Masato Horita, Kimiya Yamaashi, Koichiro Tanikoshi, and Masayasu Futakawa, "Courtyard: integrating shared overview on a large screen and per-user detail on individual screens," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'94)*, Boston, Massachusetts, USA, 1994, pp. 44-50.

[26] Richard A. Bolt, *Human Interface: Where People and Computers Meet.*: John Wiley & Sons, Inc., 1984.

[27] Renata Bandelloni and Fabio Paternò, "Flexible interface migration," in *Proceedings of the 9th international conference on Intelligent user interfaces (IUI'04)*, Funchal, Madeira, Portugal, 2004, pp. 148-155.

[28] Jan Borchers, Meredith Ringel, Joshua Tyler, and Armando Fox, "Stanford interactive workspaces: a framework for physical and graphical user interface prototyping," *IEEE Wireless Communications*, vol. 9, no. 6, pp. 64-69, December 2002.

[29] Mary Elwart-Keys, David Halonen, Marjorie Horton, Robert Kass, and Paul Scott, "User interface requirements for face to face groupware," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'90)*, Seattle, Washington, USA, 1990, pp. 295-301.

[30] Christian Heath, Paul Luff, and Abigail Sellen, "Reconsidering the virtual workplace: flexible support for collaborative activity," in *Proceedings of the fourth conference on European Conference on Computer-Supported Cooperative Work*, Stockholm, Sweden, 1995, pp. 83-99.

[31] Paul Luff and Christian Heath, "Mobility in collaboration," in *Proceedings of the 1998 ACM conference on Computer supported cooperative work (CSCW'98)*, Seattle, Washington, USA, 1998, pp. 305-314.

[32] William W. Gaver, Gerda Smets, and Kees Overbeeke, "A Virtual Window on media space," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'95)*, Denver, Colorado, USA, 1995, pp. 257-264.

[33] Jun Rekimoto, "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)*, Banff, Alberta, Canada, 1997, pp. 31-39.

[34] Jun Rekimoto, "A multiple device approach for supporting whiteboard-based interactions," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'98)*, Los Angeles, California, USA, 1998, pp. 344-351.

[35] Hideki Koike, Shinichiro Nagashima, Yasuto Nakanishi, and Yoichi Sato, "EnhancedTable: an augmented table system for supporting face-to-face meeting in ubiquitous environment," in *Proceedings of the Second international conference on Ubiquitous Computing Systems (UBICOMP'05)*, Tokyo, Japan, 2005, pp. 117-130.

[36] Masanori Sugimoto, Kazuhiro Hosoi, and Hiromichi Hashizume, "Caretta: a system for supporting face-to-face collaboration by integrating personal and shared spaces," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'04)*, Vienna, Austria, 2004, pp. 41-48.

[37] Harry Brignull, Shahram Izadi, Geraldine Fitzpatrick, Yvonne Rogers, and Tom Rodden, "The introduction of a shared interactive surface into a communal space," in *Proceedings of the 2004 ACM conference on Computer supported cooperative work (CSCW'04)*, Chicago, Illinois, USA, 2004, pp. 49-58.

[38] Thorsten Prante, Norbert Streitz, and Peter Tandler, "Roomware: Computers Disappear and Interaction Evolves," *Computer*, vol. 37, no. 12, pp. 47-54, December 2004.

[39] Brad Johanson, Armando Fox, and Terry Winograd, "The Interactive Workspaces Project: Experiences with Ubiquitous Computing Rooms," *IEEE Pervasive Computing*, vol. 1, no. 2, pp. 67-74, 2002.

[40] Katherine Everitt, Chia Shen, Kathy Ryall, and Clifton Forlines, "MultiSpace: Enabling Electronic Document Micro-mobility in Table-Centric, Multi-Device Environments," in

Proceedings of the First IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TABLETOP2006), 2006, pp. 27-34.

[41] Jacob T. Biehl and Brian P. Bailey, "ARIS: an interface for application relocation in an interactive space," in *Proceedings of Graphics Interface (GI'04)*, London, Ontario, Canada, 2004, pp. 107-116.

[42] Tom Rodden, Yvonne Rogers, John Halloran, and Ian Taylor, "Designing novel interactional workspaces to support face to face consultations," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'03)*, Ft. Lauderdale, Florida, USA, 2003, pp. 57-64.

[43] Shahram Izadi, Harry Brignull, Tom Rodden, Yvonne Rogers, and Mia Underwood, "Dynamo: a public interactive surface supporting the cooperative sharing and exchange of media," in *Proceedings of the 16th annual ACM symposium on User interface software and technology (UIST'03)*, Vancouver, Canada, 2003, pp. 159-168.

[44] Hao Jiang, Daniel Wigdor, Clifton Forlines, and Chia Shen, "System design for the WeSpace: Linking personal devices to a table-centered multi-user, multi-surface environment," in *Proceedsing of the 3rd IEEE International Workshop onHorizontal Interactive Human Computer Systems (TABLETOP'08)*, Amsterdam, Netherlands, 2008, pp. 97-104.

[45] Daniel Wigdor, Hao Jiang, Clifton Forlines, Michelle Borkin, and Chia Shen, "WeSpace: the design development and deployment of a walk-up and share multi-surface visual collaboration system," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'09)*, Boston, MA, USA, 2009, pp. 1237-1246.

[46] Hannah Slay, Bruce Thomas, and Rudi Vernik, "An interaction model for universal interaction and control in multi display environments," in *Proceedings of the 1st international symposium on Information and communication technologies (ICPS'03)*, Dublin, Ireland, 2003, pp. 220-225.

[47] Renate Fruchter, Kushagra Saxena, Matt Breidenthal, and Peter Demian, "Collaborative design exploration in an interactive workspace," *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, vol. 21, no. 3, pp. 279-293, June 2007.

[48] Naohiko Kohtake et al., "u-Texture: self-organizable universal panels for creating smart surroundings," in *Proceedings of the 7th international conference on Ubiquitous Computing*, Tokyo, Japan, 2005, pp. 19-36.

[49] Daniel Wigdor, Chia Shen, Clifton Forlines, and Ravin Balakrishnan, "Table-centric interactive spaces for real-time collaboration," in *Proceedings of the working conference on Advanced Visual Interfaces (AVI'06)*, Venice, Italy, 2006, pp. 103-107.

[50] Roswitha Gostner, Enrico Rukzio, and Hans Gellersen, "Usage of spatial information for selection of co-located devices," in *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services (MobileHCI'08)*, Amsterdam, The Netherlands, 2008, pp. 427-430.

[51] Alan Hebert and Allan Chen, "A new collaborative software package: TeamSpace at Stanford University," in *Proceedings of the 33rd annual ACM SIGUCCS fall conference*, Monterey, CA, USA, 2005, pp. 109-112.

[52] Jim Wallace, Vicki Ha, Ryder Ziola, and Kori Inkpen, "Swordfish: user tailored workspaces in multi-display environments," , Montreal, Quebec, Canada, 2006, pp. 1487-1492.

[53] Microsoft. (2013, February) Microsoft SmartGlass. [Online]. http://www.xbox.com/en-US/smartglass

[54] Microsoft. (2013, February) Xbox 360 + Kinect. [Online]. http://www.xbox.com/en-CA/Xbox360?xr=shellnav

[55] Ekaterina Kurdyukova, Elisabeth André, and Karin Leichtenstern, "Introducing Multiple Interaction Devices to Interactive Storytelling: Experiences from Practice," in *Proceedings of the 2nd Joint International Conference on Interactive Digital Storytelling: Interactive Storytelling (ICIDS'09)*, Guimarães, Portugal, 2009, pp. 134-139.

[56] George W. Fitzmaurice, "Situated information spaces and spatially aware palmtop computers," *Communications of the ACM - Special issue on computer augmented environments: back to the real world*, vol. 36, no. 7, pp. 39-49, July 1993.

[57] Jun Rekimoto and Katashi Nagao, "The world through the computer: computer augmented interaction with real world environments," in *Proceedings of the 8th annual ACM symposium on User interface and software technology (UIST'95)*, Pittsburgh, Pennsylvania, USA, 1995, pp. 29-36.

[58] R. Tesoriero, J. A. Gallud, M. Lozano, and V. M. R. Penichet, "A Location-Aware System Using RFID and Mobile Devices for Art Museums," in *Proceedings of the Fourth International Conference on Autonomic and Autonomous Systems (ICAS'08)*, Gosier, Guadeloupe, 2008, pp. 76-81.

[59] Ka-Ping Yee, "Peephole displays: pen interaction on spatially aware handheld computers," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Ft. Lauderdale, Florida, USA, 2003, pp. 1-8.

[60] Robert Hardy and Enrico Rukzio, "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*, Amsterdam, The Netherlands, 2008, pp. 245-254.

[61] Andrew D. Wilson and Raman Sarin, "BlueTable: connecting wireless mobile devices on interactive surfaces using vision-based handshaking," in *Proceedings of Graphics Interface 2007 (GI'07)*, Montreal, Canada, 2007, pp. 119-125.

[62] George W. Fitzmaurice, Shumin Zhai, and Mark H. Chignell, "Virtual reality for palmtop computers," *ACM Transactions on Information Systems (TOIS)*, vol. 11, no. 3, pp. 197-218, July 1993.

[63] Alex Olwal and Steven Feiner, "Spatially aware handhelds for high-precision tangible interaction with large displays," in *Proceedings of the 3rd International Conference on Tangible and Embedded Interaction (TUI'09)*, Cambridge, United Kingdom, 2009, pp. 181-188.

[64] Andreas Lorenz, Clara Fernandez De Castro, and Enrico Rukzio, "Using handheld devices for mobile interaction with displays in home environments," in *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services*, Bonn, Germany, 2009, pp. 1-10.

[65] Rafael Ballagas, Michael Rohs, and Jennifer G. Sheridan, "Sweep and point and shoot: phonecam-based interactions for large public displays," in *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, Portland, OR, USA, 2005, pp. 1200-1203.

[66] Rafael Ballagas, Jan Borchers, Michael Rohs, and Jennifer G. Sheridan, "The Smart Phone: A Ubiquitous Input Device," *IEEE Pervasive Computing*, vol. 5, no. 1, pp. 70-77, January 2006.

[67] Nabeel Hassan, Md. Mahfuzur Rahman, Pourang Irani, and Peter Graham, "Chucking: A One-Handed Document Sharing Technique," in *Proceedings of the 12th IFIP TC 13 International Conference on Human-Computer Interaction: Part II*, Uppsala, Sweden, 2009, pp. 264-278.

[68] Raimund Dachselt and Robert Buchholz, "Natural throw and tilt interaction between mobile phones and distant displays," in *CHI '09 Extended Abstracts on Human Factors in Computing Systems*, Boston, MA, USA, 2009, pp. 3253-3258.

[69] Tanja Döring, Sahami Alireza Shirazi, and Albrecht Schmidt, "Exploring gesture-based interaction techniques in multi-display environments with mobile phones and a multi-touch table," in *Proceedings of the 2010 International Conference on Advanced Visual Interfaces (AVI'10)*, Rome, Italy, 2010, pp. 419-419.

[70] Nintendo. (2013, February) Nintendo Wii U. [Online]. http://www.nintendo.com/wiiu

[71] Microsoft. (2013, February) Microsoft Kinect. [Online]. http://www.microsoft.com/en-us/kinectforwindows/

[72] Vicon. Vicon Tracking Systems. [Online]. http://www.vicon.com/products/system.html

[73] Till Ballendat, Nicolai Marquardt, and Saul Greenberg, "Proxemic interaction: designing for a proximity and orientation-aware environment," in *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces (ITS'10)*, Saarbrücken, Germany, 2010, pp. 121-130.

[74] Andrew D. Wilson and Hrvoje Benko, "Combining multiple depth cameras and projectors for interactions on, above and between surfaces," in *Proceedings of the 23nd annual ACM symposium on User interface software and technology (UIST'10)*, New York, New York, USA, 2010, pp. 273-282.

[75] Andrew Bragdon, Rob DeLine, Ken Hinckley, and Meredith Ringel Morris, "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)*, Kobe, Japan, 2011, pp. 212-221.

[76] Ken Hinckley, Morgan Dixon, Raman Sarin, Francois Guimbretiere, and Ravin Balakrishnan, "Codex: a dual screen tablet computer," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'09)*, Boston, MA, USA, 2009, pp. 1933-1942.

[77] Jessica R. Cauchard, "Mobile multi-display environments," in *Proceedings of the 24th annual ACM symposium adjunct on User interface software and technology (UIST'11)*, Santa Barbara, California, USA, 2011, pp. 39-42.

[78] Teddy Seyed, Chris Burns, Mario Costa Sousa, Frank Maurer, and Anthony Tang, "Eliciting usable gestures for multi-display environments," in *Proceedings of the 2012 ACM international conference on Interactive tabletops and surfaces (ITS'12)*, Cambridge, Massachusetts, USA, 2012, pp. 41-50.

[79] Douglas C. Engelbart, "Augmenting Human Intellect: A Conceptual Framework," Washington, D.C., USA, 1962.

[80] Douglas K. Smith and Robert C. Alexander, *Fumbling The Future: How Xerox invented then ignored the first personal computer*. New York, USA: William Morrow, 1999.

[81] Frederic P. Miller, Agnes F. Vandome, and John McBrewster, *%T History of Microsoft Windows: History of Microsoft Windows. Windows 1. 0, Windows 2. 0, Windows 2. 1x, Windows 3. 0, OS/ 2, Windows 3. 1x, Windows NT, Windows NT 3. 1, Windows 95*.: Alpha Press, 2009.

[82] Clifton Forlines, Alan Esenther, Chia Shen, Daniel Wigdor, and Kathy Ryall, "Multi-user, multi-display interaction with a single-user, single-display geospatial application," in *Proceedings of the 19th annual ACM symposium on User interface software and technology (UIST'06)*, Montreux, Switzerland, 2006, pp. 273-276.

[83] Vicki Ha, Jim Wallace, Ryder Ziola, and Kori Inkpen, "My MDE: configuring virtual workspaces in multi-display environments," in *CHI '06 Extended Abstracts on Human Factors in Computing Systems*, Montreal, Quebec, Canada, 2006, p. 14811486.

[84] Shankar Ponnekanti, Brian Lee, Armando Fox, Pat Hanrahan, and Terry Winograd, "ICrafter: A Service Framework for Ubiquitous Computing Environments," in *Proceedings of the 3rd international conference on Ubiquitous Computing (UBICOMP'01)*, Atlanta, Georgia, USA, 2001, p. 56*75.

[85] Jacob T. Biehl and Brian P. Bailey, "Improving interfaces for managing applications in multiple-device environments," in *Proceedings of the working conference on Advanced visual interfaces*, Venezia, Italy, 2006, pp. 35-42.

[86] Jun Rekimoto and Masanori Saitoh, "Augmented surfaces: a spatially continuous work space for hybrid computing environments," in *Proceedings of the SIGCHI conference on Human Factors in Computing Systems (CHI'06)*, Pittsburgh, Pennsylvania, USA, 1999, pp. 378-385.

[87] N. Krahnstoever, S. Kettebekov, M. Yeasin, and R. Sharma, "A Real-Time Framework for Natural Multimodal Interaction with Large Screen Displays," in *Proceedings of the 4th IEEE International Conference on Multimodal Interfaces (ICMI'02)*, Pittsburgh, PA, 2002, p. 349.

[88] Edward T. Hall, "Proxemics," *Current Anthropologist*, vol. 2, no. 3, pp. 83-100, April - June 1968.

[89] Shruti Bhandari and Youn-Kyung Lim, "Exploring gestural mode of interaction with mobile phones," in *CHI '08 Extended Abstracts on Human Factors in Computing Systems*, Florence, Italy, 2008, pp. 2979-2984.

[90] Michael Nielsen, Moritz Störring, Thomas B. Moeslund, and Erik Granum, "A Procedure for Developing Intuitive and Ergonomic Gesture Interfaces for HCI," in *Gesture-Based*

Communication in Human-Computer Interaction, 5th International Gesture Workshop (GW'03), Genova, Italy, 2003, pp. 409-420.

[91] J.O. Wobbrock, Htet Htet Aung, Brandon Rothrock, and Brad A. Myers, "Maximimizing the guessability of symbolic input," in *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, New York, 2005, pp. 1869-1872.

[92] Mathias Frisch, Jens Heydekorn, and Raimund Dachselt, "Investigating multi-touch and pen gestures for diagram editing on interactive surfaces," in *Proceedings of the 2009 ACM International Conference on Interactive Tabletops and Surfaces (ITS'09)*, Banff, Alberta, Canada, 2009, pp. 149-156.

[93] Jens Heydekorn, Mathias Frisch, and Raimund Dachselt, "Evaluating a User-Elicited Gesture Set for Interactive Displays," in *Mensch & Computer 2011*. Munich, Germany: Oldenbourg Verlag, 2011, pp. 191-200.

[94] Christian Kray, Daniel Nesbitt, John Dawson, and Michael Rohs, "User-defined gestures for connecting mobile phones, public displays, and tabletops," in *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services (MobileHCI'10)*, Lisbon, Portugal, 2010, pp. 239-248.

[95] Ekaterina Kurdyukova, Matthias Redlin, and Elisabeth Andr, "Studying user-defined iPad gestures for interaction in multi-display environment," in *Proceedings of the 2012 ACM international conference on Intelligent User Interfaces (IUI'12)*, Lisbon, Portugal, 2012, pp. 93-96.

[96] Ramona E. Su and Brian P. Bailey, "Put them where? towards guidelines for positioning large displays in interactive workspaces," in *Proceedings of the 2005 IFIP TC13 international conference on Human-Computer Interaction (INTERACT'05)*, Rome, Italy, 2005, pp. 337-349.

[97] Bolton Software. Synergy Foss. [Online]. http://synergy-foss.org/

[98] Microsoft. Microsoft Garage Mouse without Borders. [Online]. http://www.microsoft.com/en-us/download/details.aspx?id=35460

[99] Julien Robert. teleport. [Online]. http://www.abyssoft.com/software/teleport/

[100] Dugald Ralph Hutchings, Greg Smith, Brian Meyers, Mary Czerwinski, and George Robertson, "Display space usage and window management operation comparisons between single monitor and multiple monitor users," in *Proceedings of the working conference on Advanced visual interfaces (AVI'04)*, Gallipoli, Italy, 2004, pp. 32-39.

[101] Peter Tandler, Thorsten Prante, Christian Müller-Tomfelde, Norbert Streitz, and Ralf Steinmetz, "Connectables: dynamic coupling of displays for the flexible creation of shared workspaces," in *Proceedings of the 14th annual ACM symposium on User interface software and technology (UIST'01)*, Orlando, Florida, USA, 2001, pp. 11-20.

[102] Ken Hinckley, Gonzalo Ramos, Francois Guimbretiere, Patrick Baudisch, and Marc Smith, "Stitching: pen gestures that span multiple displays," in *Proceedings of the working conference on Advanced Visual Interfaces (AVI'04)*, Gallipoli, Italy, 2004, pp. 23-31.

[103] Kris Luyten, Kristof Verpoorten, and Karin Coninx, "Ad-hoc co-located collaborative work with mobile devices," in *Proceedings of the 9th international conference on Human computer interaction with mobile devices and services (MOBILEHCI'07)*, Singapore, 2007, pp. 507-514.

[104] Pedro González Villanueva , Ricardo Tesoriero, and Jose A. Gallud, "Multi-pointer and collaborative system for mobile devices," in *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services (MobileHCI'10)*, Lisbon, Portugal, 2010, pp. 435-438.

[105] Michel Beaudouin-Lafon, "Lessons learned from the WILD room, a multisurface interactive environment," in *23rd French Speaking Conference on Human-Computer Interaction (IHM'11)*, Sophia Antipolis, France, 2011, pp. 1-8.

[106] Jae Yeol Lee et al., "Dual interactions between multi-display and smartphone for collaborative design and sharing," in *Proceedings of the 2011 IEEE Virtual Reality Conference (VR'11)*, Singapore, 2011, pp. 221-222.

[107] Stephen Voida, Mark Podlaseck, Rick Kjeldsen, and Claudio Pinhanez, "A study on the manipulation of 2D objects in a projector/camera-based augmented reality environment," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Portland, Oregon, USA, 2005, pp. 611-620.

[108] European Travel Commission. (2013) Mobile Devices. [Online]. http://www.newmediatrendwatch.com/markets-by-country/17-usa/855-mobile-devices?start=1

[109] William A. S. Buxton, "Chunking and phrasing and the design of human-computer dialogues," in *Human-computer interaction*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1995, pp. 494-499.

[110] Liang Kung-Yee and Scott Zeger, "Longitudinal data analysis using generalized linear models," *Biometrika*, vol. 73, no. 1, pp. 13-22, April 1986.

[111] James W. Hardin and Joseph M. Hilbe, *Generalized Estimating Equations*, 1st ed. London, UK: Chapman and Hall/CRC, 2002.

[112] Frank A. Haight, *Handbook of the Poisson Distribution*. New York, USA: John Wiley & Sons, 1967.

[113] Julien Epps, Serge Lichman, and Mike Wu, "A study of hand shape use in tabletop gesture interaction," in *Proceedings of the 2006 ACM Conference on Human Factors in Computing Systems (CHI'06) Extended Abstracts*, Montreal, Quebec, Canada, 2006, pp. 748-753.

[114] Chris Burns, Teddy Seyed, Theodore D. Hellmann, and Jennifer Ferreira, "Towards a Useable API for Constructing Interactive Multi-Surface Systems," in *Workshop on Infrastructure and Design Challenges of Coupled Display Visual Interfaces (PPD'12)*, Capri, Italy, 2012.

[115] Chris Burns, Teddy Seyed, Theodore D. Hellman, Mario Costa Sousa, and Frank Maurer, "A Useable API for Multi-Surface Systems," in *Workshop on Envisioning Future Collaborative Interactive Spaces (BLEND'13)*, Paris, France, 2013.