Embodied Interaction for Data Manipulation Tasks on Wall-sized Displays

Can Liu

NNT : 2015SACLS207

THESE DE DOCTORAT
DE
L'UNIVERSITE PARIS-SACLAY
PREPAREE A
L'UNIVERSITE PARIS-SUD

ECOLE DOCTORALE N° 580
Sciences et technologies de l'information et de la communication

Spécialité de doctorat : Informatique

Par

Mme Can Liu

Embodied Interaction for Data Manipulation Tasks on Wall-sized Displays

Thèse présentée et soutenue à Gif-sur-Yvette, le 17 décembre 2015:

Composition du Jury :

M. Martin, Jean-Claude	Professeur, Université Paris-Sud	Président et Examinateur
M. Schmidt, Albrecht	Professor, Universität Stuttgart	Rapporteur
M. Grisoni, Laurent	Professeur, Université Lille 1	Rapporteur
Mme McGrenere, Joanna	Professor, University of British Columbia	Examinatrice
M. Marquardt, Nicolai	Lecturer, University College London	Examinateur
M. Beaudouin-Lafon, Michel	Professeur, Université Paris-Sud	Directeur de thèse
M. Chapuis, Olivier	Chargé de Recherche, CNRS	Co-directeur de thèse
M. Lecolinet, Eric	Maître de Conférence, Telecom ParisTech	Co-directeur de thèse

ÉCOLE DOCTORALE
Sciences et technologies de l'information et de la communication (STIC)

Titre : Interaction incorporée pour des tâches de manipulation de données sur un mur d'écrans

Mots clés : mur d'écrans, interaction incorporée, manipulation de données

Résumé : Avec l'avènement des données massives, des experts dans différents domaines, comme la sociologie, la médecine, l'économie, etc., sont amenés à appréhender de grandes quantités de données. Ceci pose de nombreux défis pour l'informatique en terme de traitement des données, mais aussi aux êtres humains qui doivent pouvoir comprendre les données ainsi traitées.
Dans ce contexte, les grands murs d'écrans de très hautes résolution pourraient s'avérer très utiles car ils permettent d'afficher de grandes quantité de données, donnant à un utilisateur, se déplaçant, des vues d'ensemble et des vues de détail sur celles-ci. De plus, les murs d'écrans permettent à plusieurs utilisateurs de visualiser et d'interagir ensemble de manière concourante et simultané avec différentes parties d'un ensemble de données, favorisant ainsi le travail collaboratif.
Cette thèse étudie les avantages de l'utilisation des grands murs d'écrans pour interagir et manipuler une grande quantité d'éléments disperser sur un mur d'écrans dans le cas d'un seul utilisateur et dans un contexte collaboratif. La thèse apporte de nouvelles informations sur les phénomènes interactifs mises en jeux dans ce contexte grâce à des expériences en laboratoire, et propose de nouvelles techniques d'interaction pour améliorer le travail collaboratif.

Title : Embodied Interaction for Data Manipulation Tasks on Wall-sized Displays

Keywords : wall-sized display, data manipulation, embodied interaction

Abstract : Today big data is used in many professional domains, for analyzing social behavior, uncovering health or economic phenomena, etc. This raises challenges not only for computers to process the data, but also for people to view and understand it.
In this context, ultra-high resolution wall-sized displays can be very useful. Firstly because they can display a large amount of information all at once. This enables users to navigate data with large-scale body movements : seeing an overview or various levels of details by walking around. Secondly, they allow multiple users to view different part of a data set concurrently and enable co-located collaboration that facilitates discussion and data exchange.
This dissertation studies the benefits of navigating and manipulating data on wall-sized displays through the lens of embodied interaction, for both single-user and multi-user situations. It contributes with novel insights on the interaction phenomena found in this context through laboratory experiments, as well as with the design and prototyping of novel interaction techniques for supporting collaboration.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France

SYNOPSIS

Large data sets are increasingly used in various professional domains. This raises challenges in managing and using them for sense-making, searching and classification tasks. Not only does big data require advanced algorithms to process the data, it also needs users' judgment to correct and interpret it. This thesis explores the use of large, high-resolution wall-sized displays, which can display large amounts of information, to support user interaction with large data sets.

Humans interact with their environment and with other people using a number of physical and social capabilities. In Human-Computer Interaction (HCI), the concept of *embodied interaction* describes situations where users engage their body to interact with virtual information in order to take advantage of the skills they have acquired by interacting with the real world.

I argue that large wall-sized displays enable embodied interaction. They enable *physical navigation* by spreading a large amount of data over space, leveraging users' physical abilities such as walking to navigate data. They also provide a co-located multi-user space, where users are aware of each other's presence and actions and can easily communicate and coordinate with each other. These resources can be used to design embodied interactions that support collaboration.

In this dissertation, I begin by discussing users' needs for data manipulation with large data sets, as uncovered from interviews with and observations of real users. Then I introduce a series of controlled experiments that study user interaction with large wall-sized displays, in both single-user and collaborative situations. These experiments show that, for high information density, physical navigation in front of a large display outperforms virtual navigation on a desktop monitor, because large displays leverage users' whole-body skills to navigate and manipulate data. Another experiment shows that collaboration has a cost in terms of interaction efficiency due to multitasking and disruption. However, it also shows that a *shared interaction technique* that lets multiple users issue a command collaboratively encourages collaboration, improves interaction efficiency and reduces fatigue. Based on these results, I then explore the design space of shared interaction techniques and present the design and evaluation of *Collaborative Gestures*, which aim to facilitate data manipulation and exchange in various collaboration contexts. I also introduce *PoPle*, a technique that augments direct human-to-human communication for exchanging digital information. Altogether, these techniques explore embodied interactions that leverage users' spatial and social skills in a co-located environment.

SYNTHÈSE

De grands ensembles de données sont de plus en plus utilisés dans de nombreux domaines professionnels, comme par exemple la médecine, la sociologie ou l'économie. Leur exploitation pose de nombreux défis, en particulier pour classifier les données, faciliter leur compréhension et produire du sens, ou encore, aider à la prise de décision. Outre l'élaboration d'algorithmes avancés, ceci nécessite de permettre aux utilisateurs de visualiser et d'interagir avec les données afin qu'ils puissent les appréhender, vérifier les traitements et les corriger le cas échéant. Cette thèse explore cette problématique en étudiant l'interaction utilisateur avec de grands ensembles de données sur des murs d'écrans.

Le corps humain est fait pour interagir avec le monde physique ainsi qu'avec autrui. Nous pouvons naturellement voir, entendre, toucher et nous déplacer pour interagir avec l'environnement à diverses échelles et collaborer avec d'autres personnes en communicant et en nous coordonnant. En Interaction Homme-Machine (IHM), *l'Interaction Incorporée* concerne les situations où les utilisateurs exploitent l'expérience existante avec le monde physique pour interagir avec l'information numérique.

Dans ce document, je défends la thèse selon laquelle les grands espaces interactifs permettent une interaction utilisateur incorporée en répartissant les données dans l'espace et en tirant parti des capacités physiques des utilisateurs, par exemple en marchant pour naviguer dans l'espace des données. De plus, ces environnements permettent aussi à plusieurs utilisateurs d'interagir ensemble via la communication verbale ou gestuelle tout en ayant une conscience de la présence physique de chacun. De telles capacités peuvent être exploitées pour concevoir des interactions incorporées qui supportent la collaboration.

Dans cette thèse, je commence par une discussion sur les besoins des utilisateurs pour la manipulation de grands ensembles de données, en me basant sur des interviews et des observations effectuées dans le cadre de tâches réelles. J'introduis ensuite une série d'expériences contrôlées qui étudient l'interaction avec des grands murs d'écrans, à la fois dans un contexte mono-utilisateur et dans une situation de collaboration. Ces expériences montrent que, lorsque la densité informationnelle est élevée, la navigation physique avec un grand dispositif d'affichage surpasse la navigation virtuelle sur un ordinateur de bureau car elle permet aux utilisateurs d'exploiter leurs capacités dans le monde physique pour naviguer et manipuler les données. Une autre expérience montre que la collaboration a un coût en terme d'efficacité interactionnelle du fait des interruptions et du partage de tâches. Cependant, elle montre également qu'une *technique d'interaction partagée* permettant aux utilisateurs de réaliser des commandes de manière collaborative encourage la collaboration, augmente l'efficacité de l'interaction et réduit la fatigue. A partir de ces résultats, j'explore ensuite l'espace de conception des techniques d'interaction partagée et présente la conception et l'évaluation d'une technique intitulée *Collaborative Gestures* qui vise à faciliter la manipulation et l'échange des données dans divers contextes de collaboration. J'introduis enfin *PoPle*, une technique qui augmente la communication inter-utilisateurs pour faciliter l'échange d'information numérique. L'ensemble de ces techniques explore des interactions incorporées qui mettent à profit les capacités spatiales et sociales des utilisateurs dans un environnement co-localisé.

ACKNOWLEDGMENTS

I would like to thank everybody who supported, influenced and accompanied me in the whole journey of my Ph.D.

First of all, I thank my advisors - Olivier Chapuis, Michel Beaudouin-Lafon and Eric Lecolinet for their guidance in the past three years. Olivier, thank you for always being there for me and being optimistic about what I do. You showed me many things by being hands-on in every project. You taught me what it is to do science and how to go deep in my research and find something strong. Michel, thank you for being the high-level mentor of my work. I appreciate your help when I tried to tell big stories but had difficulties supporting them. Thank you for improving my writing in every submission, including this long thesis. I learned to be rigorous and more rigorous from working with you. Eric, thank you for all the discussions whenever I have a bottle neck. They have been very helpful especially when my other two advisors agree with disagreement. Thank you for fighting with Latex and SVN for me. All in all, I have been lucky to have three of you with complementary skills and advising styles.

Special thanks to the best team I have been in - in|situ| team, the precursor of HCC teams. Almost everybody in this team had influence on me at the very beginning of my research path in 2011 as an intern. I still remember the impressive and enlightening moments that motivated me to start a PhD. For this I need to thank Stéphane Huot for being the first person who showed me the magic, and Wendy Mackay for the enlightening conversations that continued during my PhD. Our PhD seminar has been a great place to learn many things and get feedback. Although not as my advisor, Wendy has been a very influential mentor to me.

Furthermore, Anastasia Bezerianos is an amazing researcher I had a chance to work with and to learn from. Theophanis Tsandilas has been always very helpful in answering my questions from research to career. Thanks for all your constructive feedback. Pierre Dragicevic, although in a different team, is an incredibly inspiring person to me. Our several fruitful discussions led to cool perspectives for my research. Emmanuel Pietriga, thank you for helping me with extra funding and giving me great advices as well.

Also many thanks to all the members in the three teams I am involved - ex|situ| and ILDA in Orsay and VIA in Paris. The senior PhD students who graduated earlier set good examples for us and never hesitated to lend a hand. It has been nice to share the PhD struggles and conference trips with Quentin Roy, who finished his thesis at the same time. I enjoyed the random discussions with my officemates - Joseph Malloch, Thibault Raffaillac, Arnaud Prouzeau, Marc-Emmanuel Perrin and Thibaut Jacob. Thanks for all the tips from our designer Nolwenn Maudet, "secretary" Ignacio Avellino and paper work expert Oleksandr Zinenko. In addition, thank Olivier Gladin for making WILDER an almost convenient platform to work with.

In life, I need to thank my very dear friends who accompanied me through the good and bad time. Monireh Sanaei and Jessalyn Felice Alvina, thank you for listening to all my big and small problems and share feelings and stories with me. You have doubled the happiness and eased the pain. Christoph Bräunlich, thank you for being there for me no matter what, and thank you for being healthy. Mathieu Le Goc, thank you for saving my life many times and listening to my complain. I had fun arguing and fighting with you in the tea breaks and drinks.

Of course, I thank my mother for all the support she has given me during all these years abroad. The whole journey wouldn't have started without her.

Many thanks to my thesis jury members - Albrecht Schmidt, Laurent Grisoni, Joanna McGrenere, Jean-Claude Martin and Nicolai Marquardt, for reviewing my thesis and coming over to attend my defense.

This thesis was partially supported by Labex DigiCosme (Idex Paris-Saclay ANR-11-IDEX-0003-02) and the "équipement d'excellence" Digiscope (EquipEx ANR-10-EQPX-26-01), both operated by the French National Research Agency (ANR) as part of the program "Investissement d'Avenir".

CONTENTS

LIST OF FIGURES

LIST OF TABLES

"Technology is a double edged sword. Every technology enables us in certain ways while debilitating us in other ways. We have invented media that severely constrained our intellectual experience, that of all the capabilities we have, we have constrained ourselves to a tiny subset. "

Bret Victor – The Humane Representation of Thought (2014).

1 INTRODUCTION

Wall-sized displays have one of the new forms of computation that extends over space. They break the screen boundaries of traditional desktop settings and allows people to interact with virtual content with body-scale and room-scale movements such as turning, leaning and walking. In this dissertation, I show the benefits of such interactive spaces as enabling embodied interaction in the way of capitalizing humans' existing skills learned in the physical world. I first show the benefits of physical navigation of large data sets with a high-resolution wall-sized display, compared to virtual navigation for data manipulation. Then in a collaborative data manipulation context, I show benefits of providing interaction techniques that take advantage of human-to-human interactions in co-located collaboration. I promote a view of technology that blends physical and virtual resources to smooth and empower interaction rather than computerize the physical world to provide interaction.

Modern societies have entered an age of big data. Various professional domains are producing and consuming massive amounts of information for predicting climate, analyzing social behavior, mining health and economic data, detecting potential terrorists, etc. Big data raises challenges for computation to process it, as well as people's cognitive ability to understand it.

Historically, from ancient weaving machine using punch cards (Figure 1) to recent advances in computer algorithms, technology has

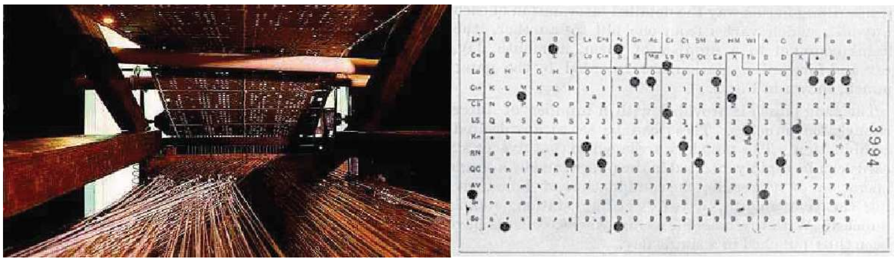


Figure 1: Left: Jacquard's Loom producing weave upon patterns from punched cards; Right: Hollerith's punched card for recording data.

progressed tremendously in processing data automatically. However, computers need human knowledge to guide the analysis as well as to adjust during the process. In the end, we often need human judgment to manually correct mistakes of the machine or interpret its results, when the problems are rather complex.

In this dissertation I discuss the large data challenge from the perspective of human, instead of machines. I am interested in the challenges and solutions when users directly interact with a large amount of data. Due to physical constraints, humans can see, perceive and reach only a limited amount of information at one time. For instance, handling hundreds of items is challenging for humans' cognitive skills.

Research in human-computer interaction tackles this problem mainly with two approaches - information visualization and high-resolution large displays.

Information Visualization uses visual elements and structures to perceptually facilitate users' understanding of data. By adding animation and interaction, the user's view space is more effectively used by changing the content or presentation over time or with direct manipulation. For instance, Fekete and Plaisant [36] present a tree maps interactive visualization technique capable of handling a million items. Blanch and Lecolinet [15] combine tree map and zoomable interfaces to facilitate the navigation in large hierarchical data sets. Hierarchical visual aggregation [34] allows existing visualization techniques to scale to massive datasets.

While visualization techniques help present an overview of data, users still need to zoom in to see details due to a limited screen space. Commonly applied focus + context techniques for viewing details [22] include pure pan and zoom (Figure 2), overview + detail (Figure 3) and Fisheye lenses [102]. Variations of these techniques include context-aware or topology-aware navigation (eg. JellyLense [89], Bring and Go [78]) and cue-based techniques (eg. Halo [9]).

Designing visualization techniques and presentations may require some prior knowledge about the data. Since high-resolution wall-sized displays feature large sizes and ultra-high resolutions, they can display massive amounts of data at once, which may facilitate the ex-

Figure 2: Pure pan and zoom interface: an overview (left) and a detail view after zooming in (right), image reproduced from [22].

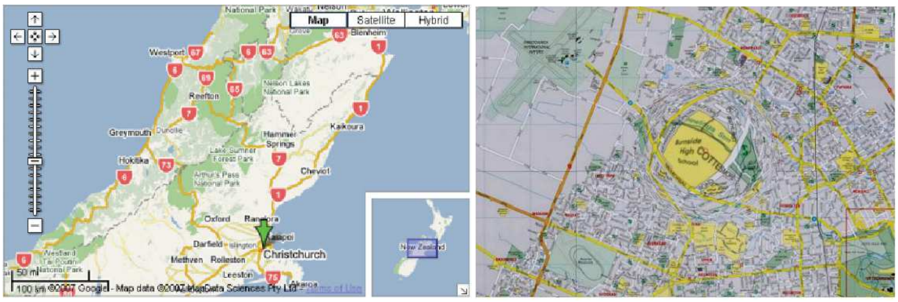


Figure 3: Overview + Detail technique (left) and fisheye lens technique (right), image reproduced from [22].

ploration of unknown data. Large amounts of data can be displayed over the space, allowing users to navigate by leaning and walking and to interact with their hands or other limbs.

Computation technology has taken different physical form factors and scales. It ranges from small to large sizes and to distributed locations in 2D to 3D. Computational devices can be put into pockets, carried around, blended in fabric or attached to the skin. On the other hand, immersive experience can be provided by interactive surfaces covering the wall, table or floor, or virtual imagery can be overlaid on the real world through a lens or projection. The "disappearing computer" from Weiser's vision of ubiquitous computing [113] is technically realized.

With all these possibilities, users are no longer fixed in a sitting position in front of a screen. They can interact with technology in the same ways as how they interact with the physical world. This leads us to the concept of "*Embodiment Interaction*".

Embodiment is a term coming from philosophy and cognitive science. It was introduced into HCI with the advancement of technology and has been given a variety of meanings. In this dissertation, I use the term *embodied interaction* to describe the interaction enabled by technology that capitalizes the human skills learned from interacting with the real-world. Our body is "made for" interacting with the environment and with each other using the skills developed over thousands of years. Many post-WIMP (window, icon, menu, pointing device) interaction styles draw their strength from embodiment. For example, tangible interaction lets users interact with digital information with their existing knowledge of manipulating physical objects, such as picking up, moving, rotating and giving to another person.

I am interested in exploring embodied interaction enabled by new forms of technology. In this dissertation, I study the use of high-resolution wall-sized displays as one way of enabling embodied interaction. For single users, it lets them navigate in a large data set by moving closer or further away to see various levels of detail or an overview. The interaction takes advantage of users' kinesthetic and spatial skills. The standard way of navigating big data on a desktop computer is to use Pan-and-Zoom or Focus+Context techniques [22]. With a wall-sized display, the pan-and-zoom interaction is "outsourced" into the real world and replaced by physical navigation. Is such "physical navigation" better than the virtual navigation on a desktop computer?

For multiple users, wall-sized displays provide a co-located environment for working together, while bringing challenges due to possible large distance between users. They afford multiple users working together while viewing different content. They take advantage of social skills in co-located environments, such as users communicating face-to-face and coordinating their work with verbal and gestural actions. However, while working in a large space, users may be far away from each other when they want to exchange data or delegate tasks. They may or may not be aware of each other's action, depending on where their focus is at the moment. They can communicate spontaneously, but they may not know what another user is talking about if a deictic gesture was not seen. Both availability of co-located communication and distance problems exist at the same time, resulting in both opportunities and challenges. How do people interact with data and with each other when performing a collaborative task involving data manipulation? How is embodiment characterized in collaborative situations and how can we design interaction techniques to take advantage of it?

Existing techniques that let users exchange digital data rely mainly on traditional user interfaces for remote communication, such as email or instant messages. In an office we often see people talking about content and then send each other emails to actually exchange the data. Tabletop or public displays facilitate such digital exchanges by providing graphical representations or territories for each user, but this may not apply in wall display environments as users may move around a lot. Why not support a more direct way to exchange information when users are physically present in co-located environments?

1.1 THESIS STATEMENT

Conventional thinking about interaction design is that it is about how to receive input from users and how to provide output from the system. Everything we design happens in the computer as an external agency. I argue that interaction can be "outsourced" to the physical world and technology should blend with users' real-world practices. I believe that interaction with technology can benefit from letting users use their learned physical and social skills to carry out a task with virtual artifacts.

In this dissertation, I show that the benefits of using wall-sized displays for data manipulation come from taking advantage of users' physical skills of walking around to explore a physical space. I show the benefits of providing collaborative techniques that take advantage of users' communication and coordination skills for interacting with each other in a co-located space. I design new interaction techniques to support data exchanges in collaborative tasks involving data manipulation, and explore embodiment in such situations.

1.2 RESEARCH APPROACH

Common ways of uncovering interaction problems start with getting insight from real users' experience of using the system. Since wall-sized displays are not yet widely deployed in industry, it is difficult to study interaction phenomena due to the lack of real users. My research started with gaining some insights by interviewing a few potential users and observing extreme users of wall-sized displays. Based on that, I designed artificial or game tasks in order to simulate interaction situations similar to the real tasks. With different experiment design, I focus on different parts of a complex real task depending on our research questions.

As shown in Figure 4, my work includes the following general approaches:

1. I interviewed potential users of wall-sized displays, who are scientists working with large data sets and need to view them all at once. This helped me to identify the tasks they need to perform with the data.

2. I observed a real task - a large conference scheduling task, performed by real users on a high-resolution wall-sized display and gained intuition about the possible benefits and drawbacks of using such displays.

3. In order to measure causal effects on the interaction, I defined an abstract classification task to operationalize the interaction part of a real task while reducing as much as possible the task-dependent cognitive part of it.

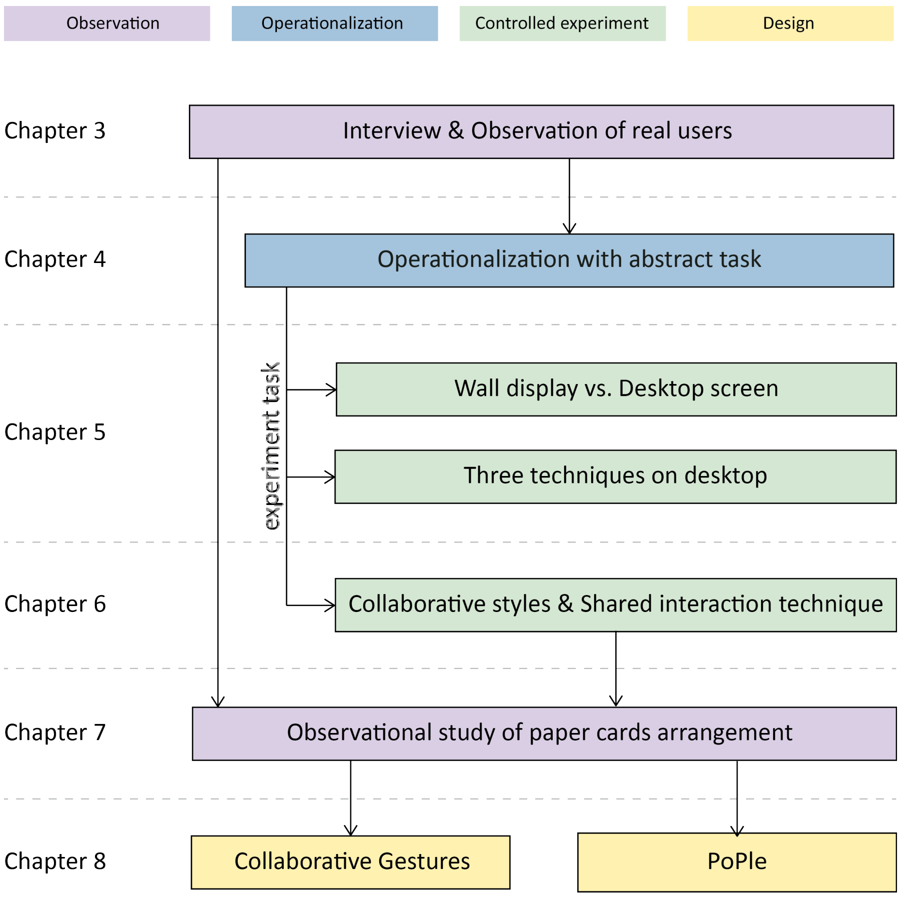


Figure 4: Overview of the thesis and the used methods. It includes formative experiments and informal observations. The experimental tasks are grounded in the real tasks from real users. The understanding from various studies informed the design of new interaction techniques.

4. I conducted several controlled experiments to understand the interaction phenomena of using a wall-sized display for manipulating large amounts of scattered data, in both single user and multiple-user cases. In the collaborative situation, the experiment uncovered some benefits of providing shared interaction techniques to support the task.

5. I conducted an observational study with a game task inspired by a real user task that requires special expertise, in order to gain insights about users' collaborative practices and help with the design of new interaction techniques.

6. I designed novel interaction techniques to assist collaborative data manipulation, based on the findings from both formative and observational studies. I also conducted an informal evaluation of one of the techniques to help improve the design of future systems.

1.3 Contributions

Through controlled experiments, I increase our understanding of interacting with a large wall-sized display, in both single-user and collaborative situations, when manipulating large data sets.

For single users, the benefit of using a wall display, compared to a desktop, comes from allowing users to physically navigate data using head and body movements, while manipulating data with hands.

For multiple users, I show the benefits of augmenting human-to-human communication with shared interaction techniques, for which each of the collaborators performs part of an action to issue a command. This takes advantage of co-located group awareness and common reference of communication, and augments existing interaction between users to empower their actions.

I then explore the design space of shared interaction techniques in the context of large data manipulation, and discuss the generalization of the concept to other environments.

1.4 Thesis Overview

Chapter 2 first summarizes related work on large wall-sized displays for overcoming user interaction challenges caused by big data. I then discuss a variety of definitions of Embodied Interaction in the literature and give my own take on it. Within this scope, I present a taxonomy of embodied interaction with example techniques and place my own work within it.

Chapter 3 describes the interviews with potential users working with large data sets and the observation of organizers scheduling a large conference with a high-resolution wall-sized display. I found that moving data items for grouping or comparison is a common essential part to make sense of large data sets. Multiple users used the spatial arrangement of data items as a tool to generate and communicate meanings, in order to find insights about a complex data set.

Chapter 4 explains the design of an abstract classification task to extract the interaction layer of a real task. It operationalizes two key factors - information density and task difficulty affecting the interaction in the observed conference scheduling task. I describe the construction of the task and the rationals.

Chapter 5 presents two controlled experiments that evaluate the benefits of using a wall-sized display for data manipulation. The first experiment compares the interaction efficiency of the wall display with a desktop setup and shows that the wall display gets increasingly more efficient with higher information density and more difficult tasks. The second experiment confirmed the finding of the first experiment by showing no significant improvement when using different navigation techniques on desktop.

Chapter 6 reports an experiment that used the same abstract task to understand interaction phenomena in collaborative data manipulation. It operationalizes different collaboration styles observed in previous literature and tests the effects of providing a shared interaction technique, with which multiple users perform part of the action to complete one elemental task. The results show advantages of the technique as well as trade-offs among different collaboration styles.

Chapter 7 explores the design of shared interaction techniques to support close and loose collaboration for a data arrangement task. It supports collaboration with different levels of coupling between collaborators. A semi-structured observational study evaluates the usability and acceptance of the techniques and gives insights for future improvement.

Chapter 8 introduces PoPle - Pointing to People, is a set of proof-of-concept techniques providing a direct way of data exchange between collaborators by leveraging human-to-human interaction in a co-located space, i.e. pointing to another user in this case. It facilitates asynchronous data exchange between collaborators by featuring a queuing mechanism.

Chapter 9 summarizes the contribution in terms of scientific understanding and design, discusses the validity and applicability of the findings in our experiments, and gives directions for future work.

"Embodiment denotes a participative status, the presence and occurrentness of a phenomenon in the world. So, physical objects are certainly embodied, but so are conversations and actions."

Paul Dourish – Seeking a foundation for context-aware computing (2001) [30].

2 CONTEXT AND BACKGROUND

This chapter begins with a summary of existing work for understanding the interaction phenomena with large wall-sized displays, in both single-user and multi-user situations. The second part discusses different levels of embodiment for interaction with virtual artifacts through a taxonomy. This chapter identifies missing points in the literature and positions my work in it.

Previous work explores interaction techniques for large displays, such as novel pointing techniques [83], mid-air interaction techniques [81, 94] as well as smart room solutions for creating an interactive workspace [16] and design requirements for interactive large public displays [79]. However, I will not give a comprehensive literature review about all the techniques for large displays.

As introduced in previous chapter, I view large displays as one way of enabling embodied interaction that uses learned physical and social skills to carry out tasks with virtual content. Therefore, my literature review in this chapter focuses on two bodies of work: studies that evaluate user interaction with large displays, and embodied interaction for interacting with virtual content.

2.1 UNDERSTANDING LARGE DISPLAYS

Most visualization techniques are designed for personal computers, where users sit in front of a desktop screen. They either improve the

Figure 5: User study of examining the productivity benefits of a Dsharp display versus a standard flat display, image reproduced from [26].

overview presentation for quicker sense-making, or provide better ways to navigate detail views. As a different approach, high-resolution large displays are capable of displaying massive amounts of data at once, thus spreading information over a large space. Users can navigate the data scene by physically moving in front of the display.

2.1.1 *Stationary Users*

Previous work has demonstrated the benefits of larger displays for traditional desktop tasks. As shown in Figure 5, Czerwinski et al. [26] conducted a user study to examine the productivity benefits of an experimental Dsharp display surface over a standard 15 inch flat panel display. The curved display is created with three projectors and has a 3072×768 resolution. Participants performed complex daily tasks requiring a large amount of task switching and multitasking. The results show significantly higher productivity and satisfaction with the larger display.

Bi and Balakrishnan [13] present a five-day study that compares a large projected wall display with single and dual desktop monitors (Figure 6). The study analyses users' behaviors in utilizing and partitioning the screen real estate and managing windows in each setup. The wall display is 16 foot wide by 6 foot high and offers a 6144×2034 resolution with self-adjusted sitting distances. The results suggest that the large display facilitates tasks with multiple windows and rich information. It also enhances users' peripheral awareness and offers a more immersive experience.

Figure 6: Diary study of users working on a large high-resolution display, image reproduced from [13].

Figure 7: Study of sense-making tasks to reveal spacial characteristics of a large display, image reproduced from [3].

More recently, Andrews et al. [3] examine the benefits of increased display size for cognitively demanding sense-making tasks (Figure 7). A sense-making task is a complex understanding process that often involves incomplete or dynamic data. Their study explores the spatial characteristics of a high-resolution large display for organization and memory. Participants were given analytic problems with a preselected data set consisting of broken information pieces. They observed and compared the task performed on a 17 inch monitor versus a tiled desktop display with 10,240×3200 resolution, consisting of a 4×2 grid of 30 inch LCD panels. They found that the large display provides a form of rapid access to external memory. It provides a space where meaning is encoded in the spatial relationships between documents on the display and distances to the users. The actual activities they observed were primarily reading, identifying information, categorizing and arranging.

There has been previous work on analyzing the effect of increased size and resolution separately. Ni et al. [85] conducted a controlled experiment that evaluates the individual and combined effects of dis-

Figure 8: Studying the effects of size and resolution separately for virtual environment tasks. From left to right: small display in either high or low resolution, projected large display in low resolution and tilted high resolution large display. Image reproduced from [85].

play size and resolution for Information-Rich Virtual Environment (Figure 8). The tasks require navigation towards a room and then search and comparison of perceptual and abstract information. They found that both size and resolution improve performance. However, larger displays reduce users' reliance on way-finding aids to acquire spatial knowledge and construct a cognitive map of the virtual environment. At the same time, higher resolution improves the legibility of textual information.

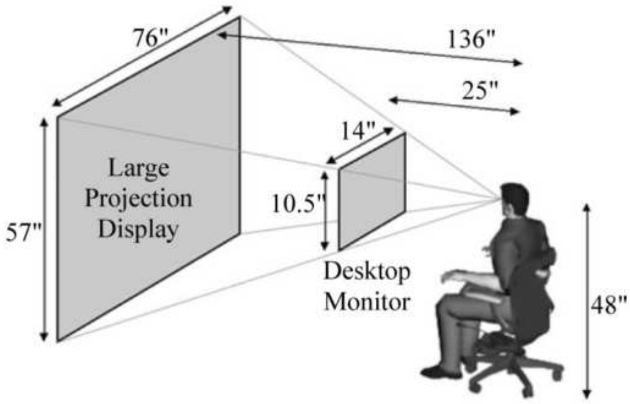


Figure 9: Setup of study comparing a desktop monitor and a large display while keeping the view angle constant, image reproduced from [107].

Tan et al. [107] conducted a set of studies to deepen our understanding of large displays for spatial tasks (Figure 9). While also comparing a large display to a desktop monitor, they configured the sitting position of the participants so that the visual angles are kept constant for both setups. They found that even with a constant view angle, a physically larger display performs better on spatial orientation tasks. Further experiments conclude that the physical size of a display implicitly affects users' choices of cognitive strategies. More immersive environments encourage users to adopt egocentric strategies, thus increase performance on tasks such as 3D navigation and mental map formation and memory.

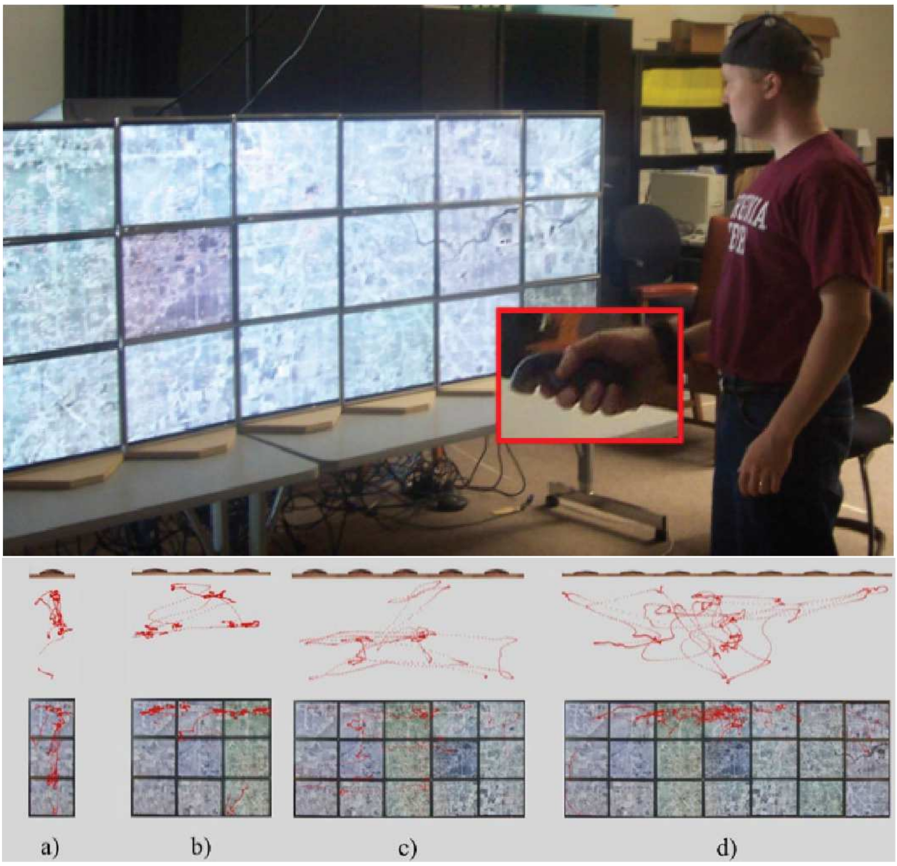


Figure 10: Study that shows the advantage of physical navigation. Top is the experiment setup. Bottom are graphs showing a user's physical navigation paths with increased display sizes from a) to d). Image reproduced from [7].

2.1.2 *User Mobility*

As display size and pixel density increase, standing and moving in front of large displays becomes necessary. This is where *Physical Navigation* is defined by Ball et al. [7] as *"bodily movement, such as walking, crouching, head rotation, etc., for the purpose of controlling the virtual camera that produces views of the information space"*. In fact, affording the mobility of users is a unique property of vertical large displays, which leads to different user behavior comparing to a desktop monitor or a tabletop surface.

Ball et al. [7] conducted an experiment to evaluate how display size affects users' choices and performance of physical or virtual navigation. As shown in Figure 10 (top), the tiled display shows the visualization of a map with housing data and provides semantic zooming. Participants use a wireless mouse to pan and zoom the view. The tasks include navigating to a target, searching for a target, pattern finding for a group of targets and open-ended insight finding. The results show that larger displays promote physical navigation and re-

Figure 11: Study investigating the effect of increased field of view and physical navigation, image reproduced from [6].

duce virtual navigation, which improves performance and user preference. In Figure 10 (bottom), the red ink in each if the four graphs visualizes the head trajectory of a user for a pattern-finding task for each screen size. The authors conclude that the key factor of a large display is to promote physical navigation, including non-tethered users, large physical space and high resolution.

Ball and North [6] investigate further the reasons for performance improvement with larger displays. Their experiment separates users' peripheral vision and physical navigation as independent variables. Figure 11 shows a participant with a blinder that limits the field of view. Participants were sitting instead of freely standing, constraining physical navigation. The tasks include navigation, comparison, search, pattern finding and estimation of values. The authors discovered that behavior and performance were most affected by physical navigation, while peripheral vision showed some improvement of performance. In addition, virtual navigation led to more cognitive load and less efficient strategies.

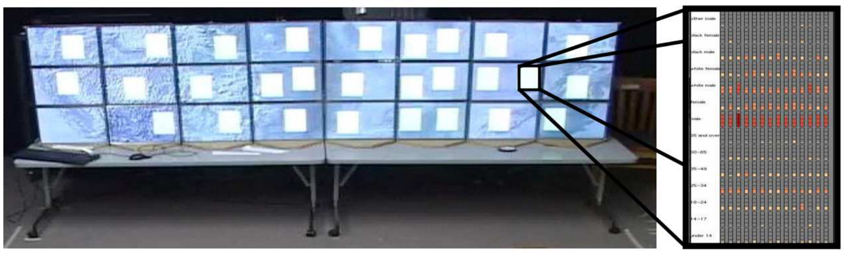


Figure 12: Study of visualization scalability on a large display, image reproduced from [115].

Yost et al. [115] investigate the perceptual scalability of visualization on a 32-megapixel tiled large display (Figure 12) compared to a 2-megapixel, 2-monitor display. Participants were presented attribute-

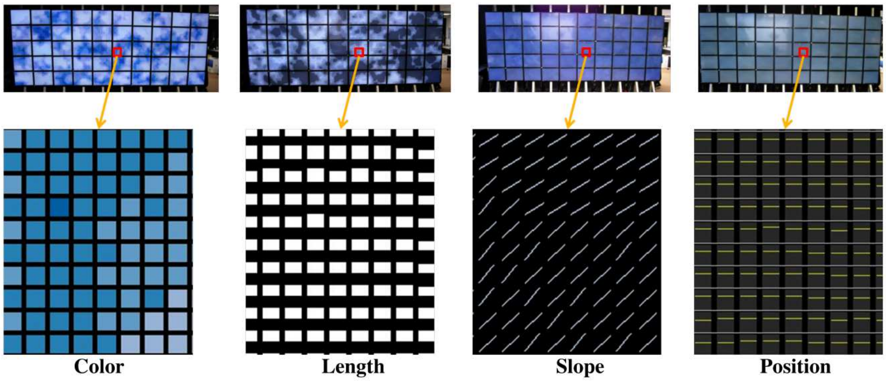


Figure 13: Effect of visual encoding on physical navigation, image reproduced from [35].

centric and space-centric visualizations to complete both overview and detail tasks involving searching and sense-making. As the information density is kept constant, the amount of data increases proportionally to the screen size. The results show that a twenty-times increase in data on large display only resulted in a three times increase in performance time, without a significant decrease in accuracy. In addition, they also suggest that spatial grouping is more important with more data on a larger display.

Recent work from Reda et al. [95] studies the effects of display size and resolution on user behavior and insight acquisition in visual exploration tasks. They find that larger displays with higher resolution can significantly increase discoveries and yield broader and more integrative insights. They show that larger displays engage cognitive activities for analytic tasks, because they promote spatial separation of information and reduce the need for visual context switches.

Researchers have tested the effect of visual encoding on physical navigation. Endert et al. [35] demonstrate that visual encodings provide visual aggregation when the user steps back from the display (Figure 13). The authors provide the viewer with meaningful patterns when the visualization is viewed at a distance. They compared four basic visual encodings (color, length, slope, position) with small and large displays. They found that color encoding promotes physical navigation more effectively, and strongly improves time and accuracy when compared with the other encodings.

Furthermore, perspective distortion on large displays has also been studied. When users look at a 2D data visualization with a large angle, their judgement of relative size and orientation is affected. Bezerianos and Isenberg [12] examined the misjudgment of data visualization elements, i.e., lenths, angle and area, caused by perspective distortion when viewing from the side of a wall display. They compared the situations with and without physical navigation and concluded that physical navigation does not help with this misjudgement effect.

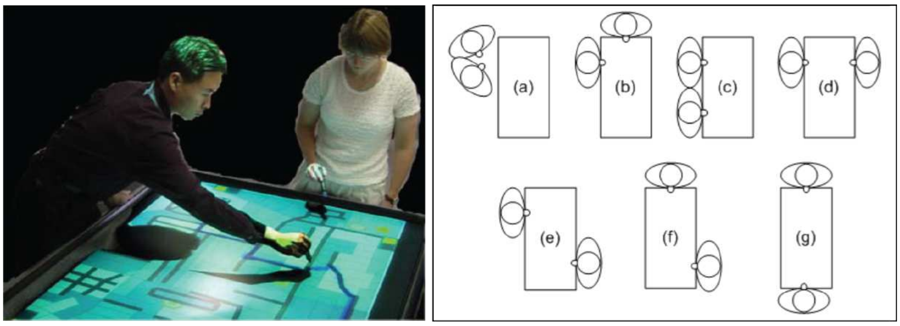


Figure 14: Users' position arrangements found in collaboration on tabletop: a) together, b) kitty corner, c) side by side, d) straight across, e) angle across, f) end side, and g) opposite ends. Image reproduced from [109].

2.1.3 *Multiple Users*

Large wall-sized displays afford multiple users interacting with the surface and with each other. Collaboration potentially also helps to cope with big data, as co-workers can combine their knowledge and inspire ideas to each other by discussing and using common references [10, 98]. Collaboration with large displays has mostly been explored through observational studies.

Scott et al. [103] conducted observational studies to understand the division of shared space and the ownership of content in natural collaboration practice on a tabletop. They found that people coordinate their interaction with three types of territories: personal, group and storage. Each of these territories have different functionalities as well as spatial properties.

Tang et al. [109] refer to *collaborative coupling* as "*the manner in which collaborators are involved and occupied with each other's work*". It relates to collaborators' dependency, coordination and awareness in group activity. They observed pairs of users working around a tabletop for a route creation task over spatially fixed data sets. They identified six typical collaborative coupling styles (Figure 14): *same problem, same area*; *one working, another viewing in an engaged manner*; *same problem, different areas*; *one working, another viewing*; *one working, another disengaged*; and *different problems*. The studies showed the relationship between coupling styles and preferred tools, physical arrangement and interferences. The authors suggest to support mixed-focus collaboration by providing tools for flexible view transitions.

In the context of visual analytics tasks, Isenberg et al. [53] studied collaboration on a tabletop and identified eight types of collaborative styles ranging from loose to close collaboration (Figure 15). They also found that groups that collaborated more closely together were more successful at the task and required less assistance.

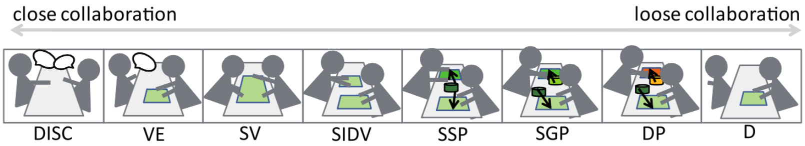
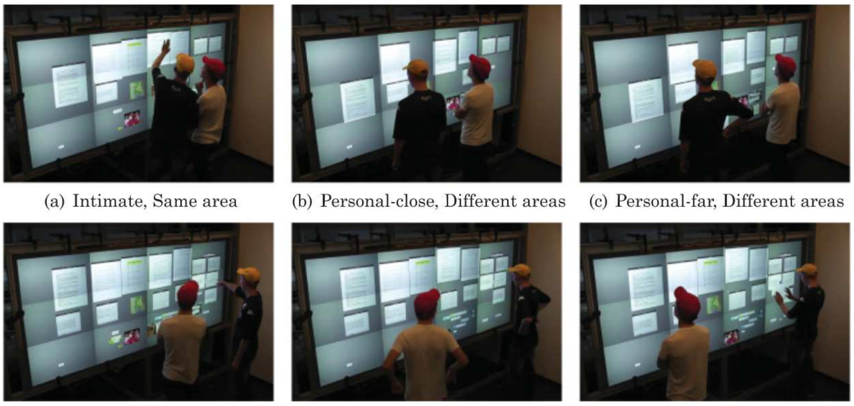


Figure 15: Eight collaboration styles ranging from loose to close collaboration. DISC: discussion; VE: view engaged; SV: sharing view; SIDV: sharing information with different view; SSP: same specific problem; SGP: same general problem; DP: different problem; D: disengaged. Image reproduced from [53].

Vertical and horizontal displays have different properties, thus they lead to different user behavior in terms of physical movement and arrangement, as well as their reach range and their position when they manipulate content. As discussed in earlier sections, user mobility is higher when working with vertical displays, which promotes physical navigation when viewing large amounts of data [7]. Therefore the results from tabletop surfaces can only be partially applied to wall-sized displays.



(a) Intimate, Same area (b) Personal-close, Different areas (c) Personal-far, Different areas

(d) Personal-far, Same areas (e) Social, Look (f) Social, Different areas

Figure 16: Six coupling styles found when pairs of users performing a problem-solving task with a multi-touch wall display, image reproduced from [57].

With a multi-touch wall-sized display, Jakobsen & Hornbaek [57] observed pairs performing a collaborative problem-solving task with a collection of documents (Figure 16). They identified six collaboration styles based on the pairs' proximity and visual attention. They found a significant association between the proximity of the pairs and the closeness of joint work. The simultaneous input afforded by the touch surface reduced the need for coordination for loosely coupled work. Moreover, collaborating groups did not divide the display into physical territories, but shared it evenly with frequent switches between joint and parallel work. Regarding physical navigation, participants passed documents to each other to reach different parts of the display. Their physical positions provided context for conversation.

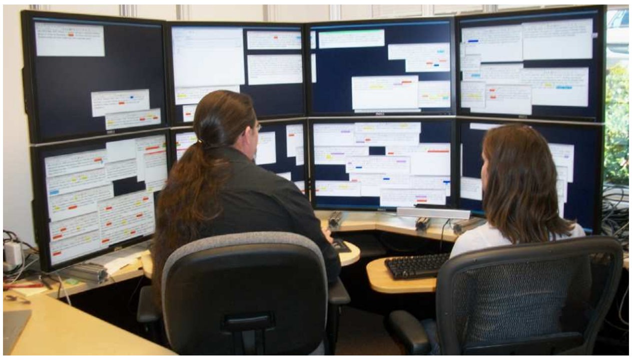


Figure 17: Study of a collaborative sense-making task with a high-resolution display in desktop setup, image reproduced from [17].

Bradel et al. [17] conducted a study of collaborative sense-making task while pairs of participants sitting in front of a curved large display and interacting with mice (Figure 17). They analyzed how users externalize and organize information in different territories and collaborative styles. It appeared that the spatial organization of documents for building common knowledge promotes close collaboration, which leads to better performance.

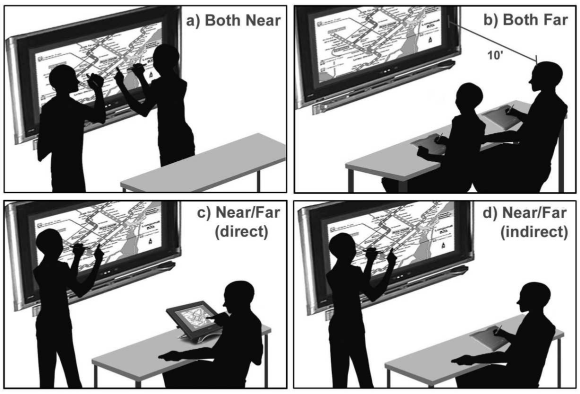


Figure 18: Experiment evaluating four proximity and arrangement conditions, image reproduced from [48].

A rare example of a controlled experiment on collaboration with large displays is Hawkey et al.'s work [48], which compares different physical configurations of pairs of users with direct input on the wall or indirect input on a fixed tablet in front of the display (Figure 18). Pairs of participants completed route planning tasks with each of the configurations and evaluated the effectiveness and enjoyment of collaboration. The results showed that participants preferred working together regardless of the distance to the display. Awareness was compromised when participants were far from the display as the gestures are less visible. In addition, the fact that the tablet was away from the wall display degraded shared understanding due to the lack of shared reference.

2.1.4 *Position of My Work*

Existing work on evaluating effects of large displays focuses on pure visualization tasks, such as sense-making and insight finding tasks. I found that there is a lack of attention to data manipulation tasks in existing work. However, as I will describe in Chapter 3, based on my interviews and observation of real users, rearranging and manipulating data items is essential in many tasks involving large data sets. Also, for collaborative tasks, the majority of literature focuses on observational studies that provide rich insights about how people freely collaborate.

I believe such studies must be complemented by controlled experiments that help us better understand the phenomena at play and uncover causal effects. Chapter 4, Chapter 5 and Chapter 6 describe how I operationalized a data manipulation task to study interaction phenomena of using large displays for such tasks.

As mentioned before, I view large displays as one of the new forms of interactive environments that enables embodied interaction, in a way that users use their whole-body capability for interacting with virtual content. Therefore the next section illustrate an overview with example of embodied interaction that capitalize different levels of interaction capabilities.

2.2 EMBODIED INTERACTION

Embodiment is a term discussed in various domains such as philosophy, cognitive science, neuroscience and psychology. The concept of embodiment has been developed in many directions over the last decade. Marshall et al. [72] summarize the issues of the theory of "Embodied Interaction", and raise the question of whether embodiment can be considered as a single concept or a set of distinct perspectives. For example, recent work in cognitive science has developed theories of seeing embodied cognition as "less abstract and less brain-based and more embodied, embedded, extended or enactive" [71].

Embodiment has been introduced and developed from an HCI perspective since the publication of Dourish's "Where The Action Is" [29]. According to this book, the embodied nature of action from Merleau-Ponty's philosophy gives three meanings of embodiment: *"The first is the physical embodiment of a human subject, with legs and arms, and of a certain size and shape; the second is the set of bodily skills and situational responses that we have developed; and the third is the cultural "skills", abilities, and understanding that we responsively gain from the cultural world in which we are embedded"*. Dourish takes a broader perspective to define embodiment as "*the property of being manifest in and of the every-day world*". He claims that embodiment includes both physical and social aspects. Not only physical objects are embodied, but also other as-

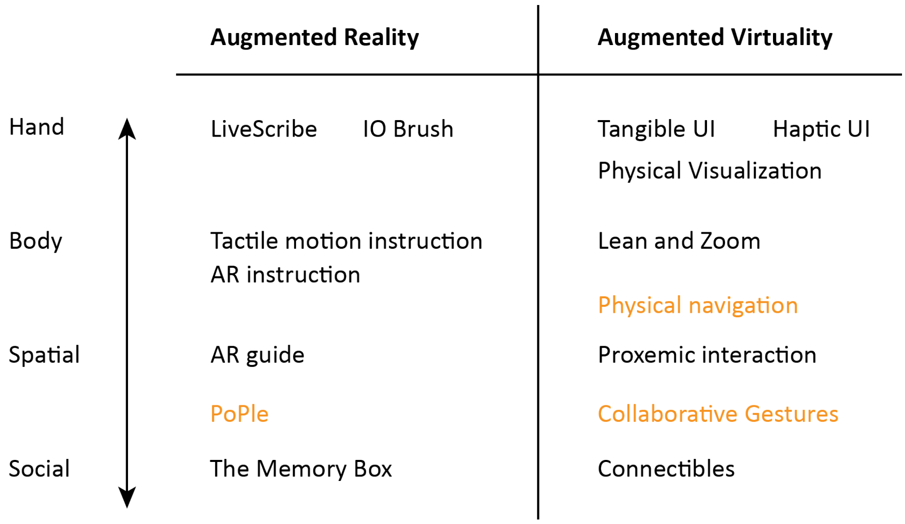


Figure 19: Examples of interaction techniques that take advantage of learned human skills, categorized by the scope as the y axis and the ways of combining physical and virtual resources. The highlighted items are studied in this dissertation.

pects of everyday life, such as conversation, is also embodied. Dourish emphasizes the situated, participatory aspect of interaction that sits in users' real world practice.

Klemmer et al. [65] present several themes for designing and evaluating interactive systems based on theories of embodiment that ranges from cognition to interaction and from single-user interaction to collaboration. They highlighted that embodied interaction facilitates learning, because people think and learn through doing. In terms of performance, users' action-centered skills and motor memory help to improve task efficiency. Moreover, the presence of physical artifacts affects cognition and collaboration. In fact the meanings of embodiment has been extended with the emerging of a large variety of terms in HCI, including "embodied conduct", "embodied cognition", "whole-body interaction", "embodied conceptual metaphors", etc [72].

Jakob et al. [56] use the notion of *Reality-Based Interaction* to describe the interaction styles that "draw strength by building on users' pre-existing knowledge of the everyday, non-digital world to a much greater extent than before". They categorizes four themes of reality, including users' understanding of naive physics, their own bodies, the surrounding environment, and other people. This notion shares a common ground with aspects of embodied interaction discussed in the literature. They both highlight the value of taking advantage of users' skills learned from the real world.

In this dissertation, I discuss embodied interaction in the context of data manipulation. I focus on the embodiment for users to perform actions or tasks, rather than to learn or understand through acting.

Based on the literature above, I use *Embodiment* as an umbrella term to represent properties of interaction that allows users to engage their body in the interaction with virtual content, as well as capitalizes on human's experiences gained through interacting with the physical world and with other people. The embodiment resources include not only physical skills for interacting with physical objects, but also social skills for interacting with people, such as the awareness of co-presence, understanding of deictic actions and coordination, etc. In the following, I present a taxonomy (Figure 19) with examples of such embodied interaction following the above definition.

The primary dimension categorizes human skills for interaction in four levels: hand, body, spatial to social. The second dimension discuss embodiment together with different forms of augmentation that empower users. Milgram and Kishino [73] illustrated a spectrum ranging from real to virtual environment, with Augmented Reality and Augmented Virtuality in the middle. The former augments interaction in the real world with computation, while with the latter, interaction with virtual content is augmented by objects and actions in the real world. These two terms can be distinguished according to whether or not the augmented activity exists even without the extra functionality. For example, a digital pen with real ink augments the inking activity by digitalizing the ink traces, while a tangible knob exists only to control parameters of videos or images with haptic feedback. I use this distinction as another dimension to describe the example techniques that enable embodied interaction.

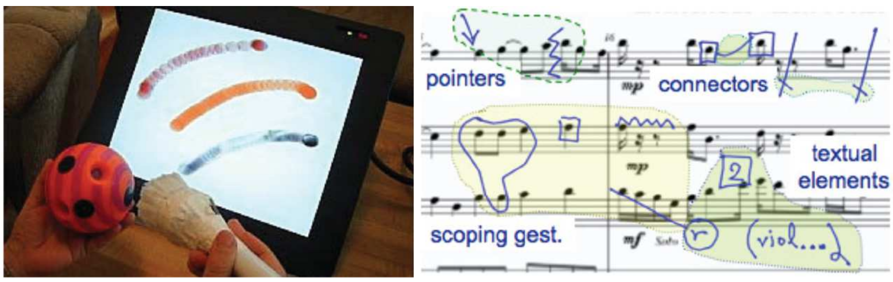


Figure 20: Example techniques augmenting inking: I/O Brush (left) and Musink (right), images reproduced from [101] and [110]

2.2.1 *Hand Level*

Augmented Reality techniques augment existing physical activities with extra functionalities. With I/O Brush [101] (Figure 20 left), users can pick up a pen and "dip" it on any real objects, then paint with the texture patterns of the real objects on a screen. The pen is equipped with a camera so that it can capture any visible optical patterns. The drawing action is augmented with the function of capturing non-digital patterns. As another example, digital pens such as LiveScribe[1] track the pen tip on paper to digitize the ink while users write and draw on paper. The writing action is augmented with digital properties. Musink [110] (Figure 20 right) allows composers to explore ideas with free-hand drawing while being able to listen to associated music components during the inking process. Therefore the drawing patterns are augmented with sound and controls. These approaches empower users while they perform their existing real world activities. From an embodiment perspective, virtual content is created on-the-fly during users' writing and drawing process.

Augmented Virtuality uses an opposite way to blend physical and virtual interaction, by helping users apply real world actions to carry out a task with virtual content. Graspable interfaces [39] and Tangible interfaces [54] let users control virtual artifacts with physical controls that provide a rich input vocabulary. Instead of selecting menus or pressing buttons, the user can grasp the object, turn it or rotate it, or deform it. Widgets such as knobs and sliders are physical objects, computerized to sense user input.

The earliest example of tangible interface is the Marble Answering Machine imagined by Durrell Bishop in 1992 [54] (Figure 21). Marbles represent voice messages that have arrived in order. A user can pick up a marble and put it in a dent to listen to the message. SLAP widgets [114] is a more recent example of tangible user interface. By recognizing physical widgets placed on a tabletop surface as well as the users' manipulation, users can associate a widget to virtual content (eg. video) to control its parameters (Figure 23 left). Figure 23 right shows The reacTable [61], a tangible music instrument that allows users to control and mix music components with physical objects. It affords multi-parametric and shared control among multiple users and provides a playful collaboration experience.

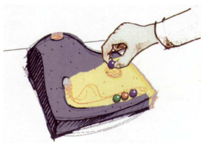


Figure 21: Marble Answering Machine

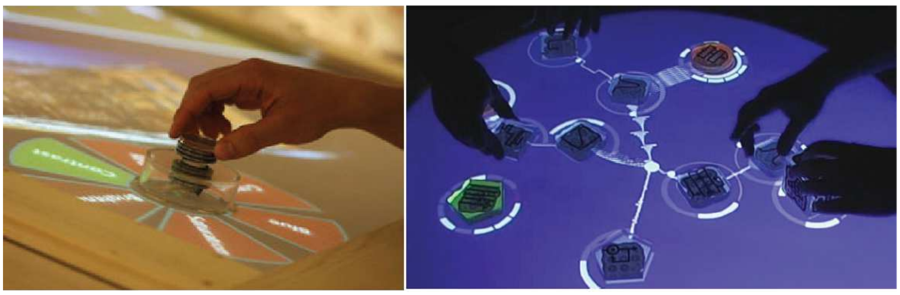


Figure 23: Example of tangible interfaces: SLAP widget (left) and The reacTable (right), images reproduced from [114] and [61]

Another approach is to provide haptic or tactile feedback to simulate a similar physical sensation - haptic or tactile feedback, when users interact with virtual content. For example, TeslaTouch [8] (Fig-

1 http://www.livescribe.com/enus/

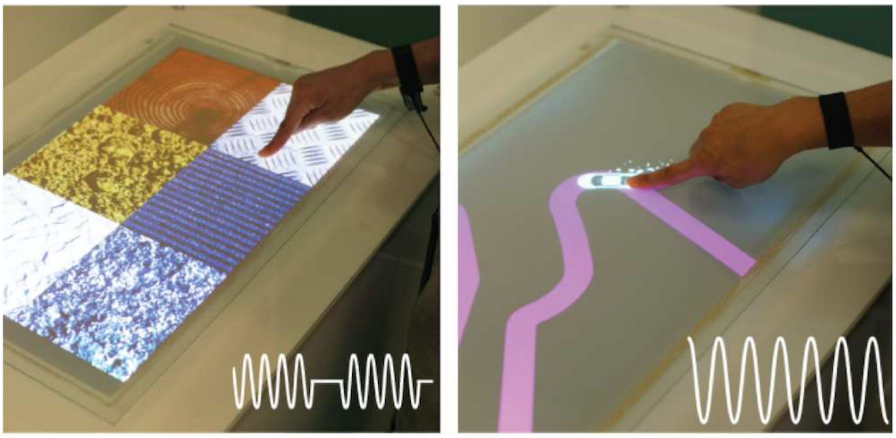


Figure 24: TeslaTouch, image reproduced from [8]

ure 24) uses electrovibration to produce tactile feedback on touch displays when users touch virtual objects. Different vibration patterns are provided to give a similar touch sensation as the real material, so that users could distinguish the objects even when they are eyes-free. This takes advantage of users' physical sensation learned from interacting with the real world.

Physical visualization[2] gives virtual content physical shapes and properties. Data visualization is depicted in 3D shapes, and the viewer can turn the physical objects around and look at it from different perspectives. Project FEELEX [55] is an early example of a shape display, displaying height information in geographic data with a haptic surface consisting of height-changing pins. Digital fabrication facilitates the fast creation of physical presentation, which is shown to be beneficial as a new form of interaction [60]. Physical visualizations vary from small physical bar charts that can be rotated in hands to large ones that fills a table, a wall or more space. The larger ones involve the users' body and spatial skills for viewing data as well.

The embodiment of above techniques lies in the fact that users can sense the property of virtual content through body sensation, and interact with virtual content with their hand manipulation skills.

2.2.2 *Body Level*

As Augmented Reality, body position and postures for performing a task can be augmented by providing virtual instructions. Tactile motion instruction [106] is a wearable instructor that provides vibration feedback to correct users' wrong body position while learning to skiing. It augments users' physical activity with haptic information.

As shown in Figure 25 [49], a see-through display augments the environment with an instruction overlay so that users can follow them

2 http://dataphys.org/list/

Figure 25: AR assistance for an assembly task. Dynamic 3D overlay shows the current manipulation step, then switch to next step upon correct completion. Image reproduced from [49].

step by step while proceeding with the manipulation task. The instruction is embedded in-place and responds to the user action so that users do not need to separate interaction from reading the manual. One of my earlier work [66] evaluates the benefit of providing feedback in mobile AR instructions. It is shown to reduce the cognitive load of switching attention between the physical task and the virtual instruction overlay. AR lets users interact with physical objects and navigate in the environment in the same way as they do in the real world: users browse virtual content with physical actions, taking advantage of humans' physical capability as well as spatial and environmental awareness.

Figure 26: Tactile Motion Instruction (left) and Lean and Zoom (right). Images reproduced from [106] and [46]

As an Augmented Virtuality approach, Lean and Zoom [46] is a technique that automatically enlarges content on the screen when the user gets closer to it (Figure 26 right). Above techniques support embodied interaction that capitalizes on users' body-scale and spatial skills. They augment users' body coordination while learning a sport or actions in assembly activities, etc.

2.2.3 *Spatial Level*

Commonly seen AR (Augmented Reality) techniques augment the physical world with overlaid virtual imagery associated to objects or location. For instance, AR tourist guides let users navigate and interact with the physical world as they would normally do while providing physically indexed information.

In a room-scale space, people are generally aware of existing entities in their surroundings, and particularly notice social activities happening nearby. Previous work also takes advantage of environmental and social awareness based on Edward Hall's Proxemics theory [33], which reveals four zones around people have different social meanings: intimate, personal, social and public.

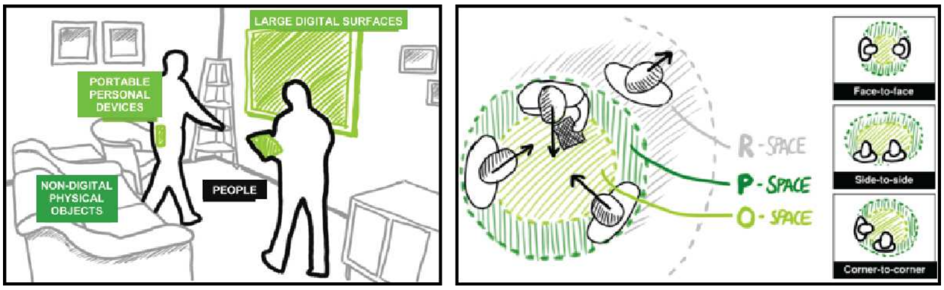


Figure 27: Proxemic interaction (left) and Cross-device interaction (right), images reproduced from [42] and [70].

Proxemic Interaction [42] takes a different approach, applies the social implication of proxemics in interaction design and explores various possibilities to use the distance between the users and the display as implicit input for interaction (Figure 27 left). Marquardt et al. [70] explore, through F-formations and Micro-mobility, how distance and relative body orientation among multiple users and the orientation of their devices towards each other can serve as input parameters for sharing information (Figure 27 right). Pass-Them-Around [68] allows users to share photos by "throwing" the photo towards another mobile phone's direction. Tilting the phone differently triggers different sharing interactions.

The embodiment of these technique is depicted through taking advantage of users' orientation and navigation skills for exploring a physical space, as well as the social resources between co-located users, such as their awareness of others' action. Going further, social and cultural conventions can also facilitate interaction.

2.2.4 *Social Level*

The Memory Box [41] allows users to place a physical object, e.g. an old photo, in a box that is associated to an audio source (Figure 28 right). Whenever this memory box is opened for viewing the photo, the music would play and set an atmosphere. Similarly with Spyn [99], users can associate video or image with positions on a scarf to share memories of a knitting process. They can be played on a see-through mobile phone lens.

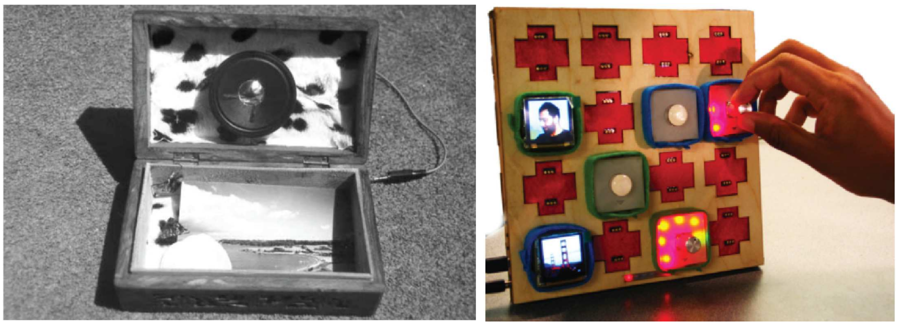


Figure 28: The Memory Box (left) and Connectibles (right), images reproduced from [41] and [62].

Connectibles [62] (Figure 28 left) leverages a gift-giving practice, with which users exchange tangible objects to automatically form communication channels between them. This physical-based social network practice is called “tangible social network”.

In my opinion, embodied interaction techniques in social situations are those that participate and blend in the social interaction between users, taking advantage of social resources while respecting the boundaries.

2.2.5 *Position of My Work*

In this dissertation, I have studied embodied interaction in both single-user and multiple-user situations. Physical navigation enabled by a high-resolutaion wall-sized display is formally evaluated in Chapter 5, which shows its benefits compared to virtual navigation for a classification task. In collaborative situations, Chapter 6 shows the benefits of providing a shared interaction technique, which lets multiple users collaboratively perform one operation, i.e. exchanging data, to support close and loose collaboration. I explore the design space of shared interaction techniques with Collaborative Gestures. Also I introduce PoPle - Pointing to People, a concept of augmenting co-located human-to-human communication to interact with virtual content.

Figure 19 shows where I place *Physical navigation*, *Shared Interaction Technique*, and *PoPle* in my taxonomy of embodiment. Phyical navigation takes advantage of kinesthetic body skills in a large space for data exploration and manipulation. Shared Interaction Techniques and PoPle take advantage of users' communication and coordination skills in a co-located environment including face-to-face communication, deictic actions and awareness, in either Augmented Reality or Augmented Virtuality.

2.3 SUMMARY

This chapter began with a review of studies for understanding phenomena when users interact with large displays, in single-user and collaborative cases. Then the second section took a broader view of embodied interaction, which takes advantage of users' learned physical and social skills for interacting with virtual content. I discussed various technologies that blend physical and virtual interaction in order to provide a smooth user experience. I provided a taxonomy of embodiment that covers single-user and multi-user interaction. The next chapter comes back to the topic of large wall-sized displays and begins with the interviews and observations I conducted with users of such displays.

3 THE NEED FOR DATA MANIPULATION

Today big data is used in various professional domains, such as scientific research, medical and financial analysis. Users work with large amounts of data both individually and collaboratively. In order to understand users' need, I interviewed several potential users of large displays, who face interaction challenges while working with large data sets on a desktop computer. I also observed a large conference scheduling task performed on a wall-sized display. This chapter describes the interviews and observations, and discusses the insights gained from investigating users' needs and problems.

3.1 INTERVIEW 1 - GRAPH THEORY RESEARCHER

Graph theory researchers work with *graphs* as mathematical structures to model relationships between objects [14]. They work with large and complex graphical representations consisting of *vertices* connected by *links*. I interviewed two graph theory researchers from the GALaC group in LRI research lab in France. They talked about their interest in using a large display and their needs as potential users.

Two graph theory researchers were interviewed. They were asked to show samples of data they work with and explain the actual tasks performed with it. Figure 29 shows a printed graph that is 5 meter long and 0.3 meter high. They printed it segment by segment on 24 A4 sheets and then they stitched them together manually with

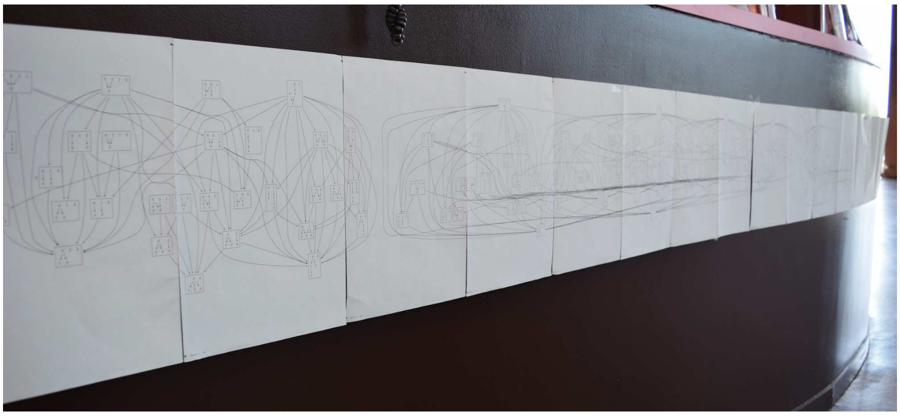


Figure 29: Long graph from graph theory researchers, printed on 24 A4 papers and stitched together

adhesive tape. Obviously, this process can be cumbersome and time-consuming. Moreover, once printed, any update on the graph would require a new print.

This was the only way they could look at the whole graph in detail. Asked how they view these graphs otherwise, a researcher answered *"We basically dropped the idea of looking at them"*. Apparently, the professional software they use lacks support for visualizing and navigating such large data.

Figure 30 shows a digital scan of a graph they worked with. The graph is printed on two sheets of A4 paper stitched together. As shown in the overview, the graph is composed of rectangles that are connected by directional colored curves that are numbered. Numbers provide meta-information and are written on top of each rectangle. Each rectangle contains two diamond-shaped graphs. Each diamond shape is a net whose links and nodes are drawn in different colors. In each rectangle, mathematical operations are annotated using numbered arrows.

The paper was annotated by the researchers during the work process. They annotated some of the rectangles with mathematical operations such as plus, minus and multiply. The pencil traces show that they tried to group rectangles by similarity. Each group was annotated by a number sequence such as 2341.

Through this interview, I found that potential users of large displays use the graphical structure to encode complex relations between data units. Some basic tasks emerged while the users were working with the large data set, such as:

- Viewing the data set in an overview and at several levels of detail in order to understand the structure;
- Comparing differences between units at various levels of detail;
- Grouping similar units at the same level.

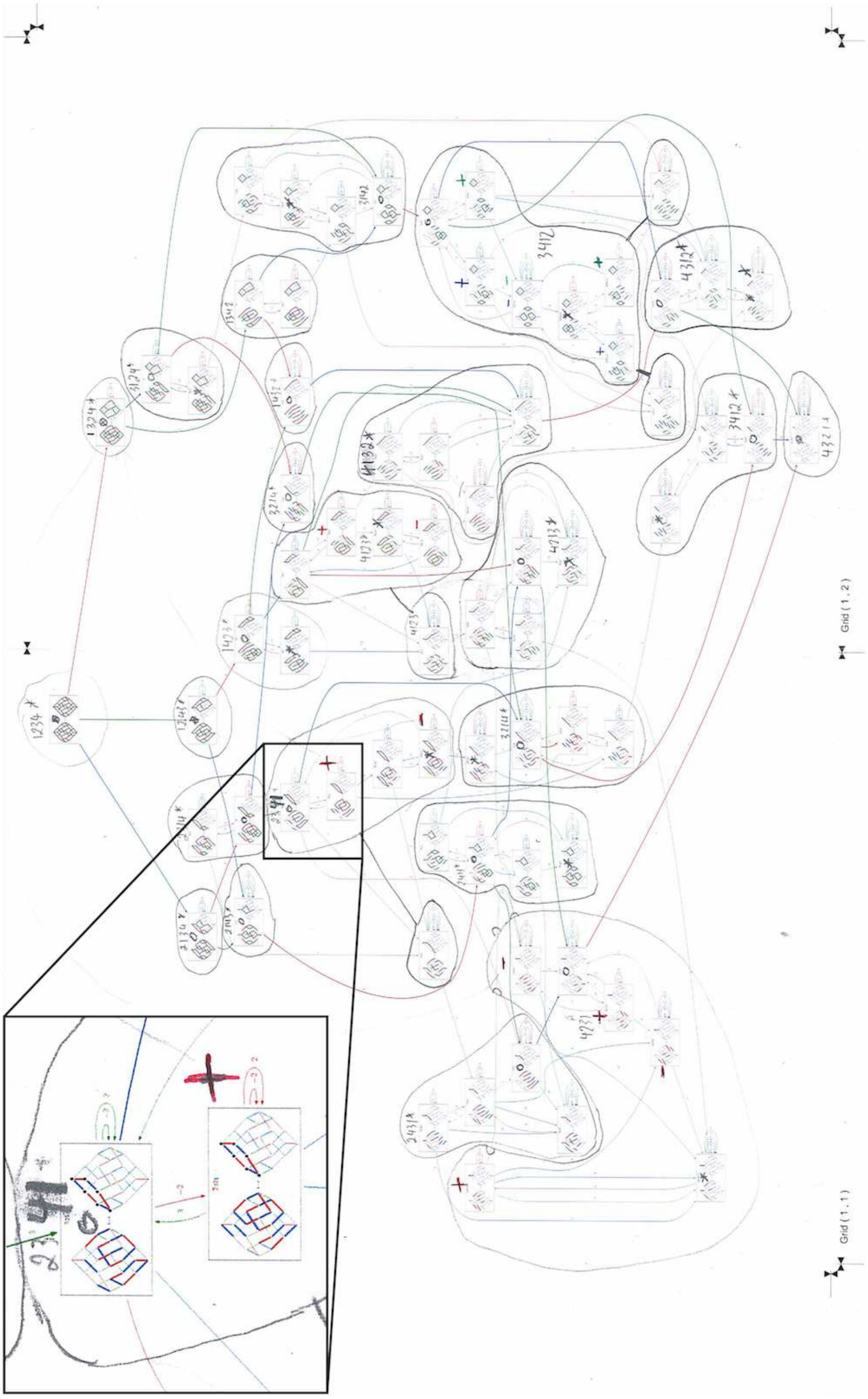


Figure 30: Scan of a draft from graph theory researcher. It is annotated with hand writing and drawing.

3.2 INTERVIEW 2 - SOCIOLOGIST

Another interview I conducted was with a sociologist, who leads a project that aims to understand the relations between personal social network behavior and eating disorder diseases. They collected several hundreds of questionnaires about the respondents' personal social network. Each respondent created a personal social network with a tool that visualizes their social connections with different levels of closeness and different kinds of subgroups (Figure 31 b). This produced hundreds of images of personal networks. Then, they invited professionals with various expertise, eg. statisticians and sociologists, to look at all the images together. A large amount of images were displayed on a touch-sensitive wall display to allow experts to see and manipulate them together. The goal was to combine the experts' knowledge to find metrics to classify the data elements, so that they might be able to design classification algorithms in later stages.

Motivation for colocated collaboration

The interviewee was asked to describe the project and the motivation of letting experts collaborate in this colocated environment. I asked Critical Incident questions [40], and he replayed some scenarios from their previous work session.

3.2.1 *Provide Common Vocabulary For Discussion*

Figure 31 (b) shows an example of the visualized social network of a questionnaire respondent. Each colored dot represents a social contact of the respondent. They are linked if they know each other, and grouped if they belong to a group. Yellow indicates close friendship, blue means the contact is emotionally connected to the respondent, and red dots are acquaintances. Filled dots mean they are connected online and empty circle that they are connected offline.

The interviewee explained that this visualization was designed and agreed upon by experts in different domain, so as to show factors that are important to all of them. When they view the images together, they actually see them from different perspectives. Meanwhile, they can discuss the classification with a common "vocabulary", whereas otherwise they would face communication problems with different terms used in their respective domains.

3.2.2 *Encode Meaning In Space*

The interviewee emphasized that the purpose of letting the experts collaborate in a co-located way was to spark ideas through communication. The expected outcome should be creative ways of grouping the images. Images can be classified by certain types of similarities or differences. By arranging them in a certain way, they can also tell a story. The experts communicate with each other by moving the im-

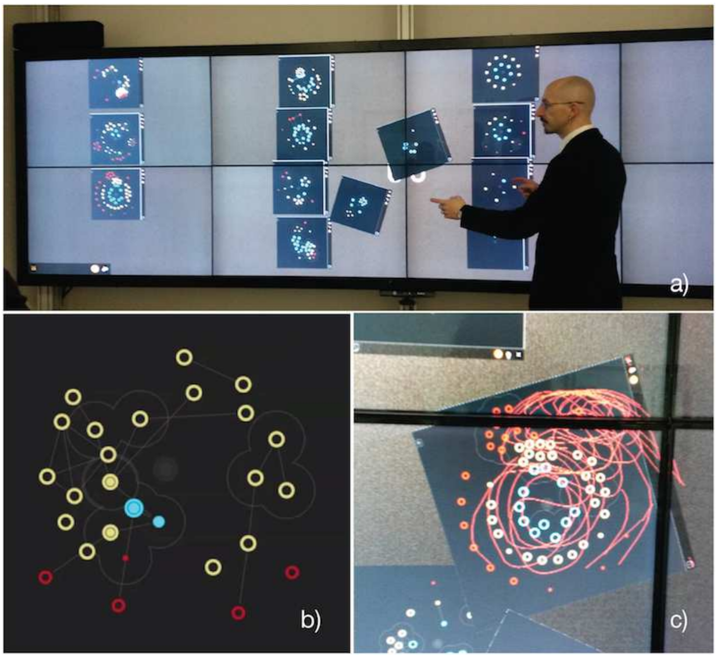


Figure 31: Sociologists' work on a touch-sensitive wall display. a) a sociologist (the interviewee) showing his work with example images; b) one example of a personal social network visualization; c) example of an expert's annotations on the image.

ages and showing the arrangement to another expert, or by manipulating them together to get ideas. Space is thus encoded with rich meaning.

In Figure 31 (a), the images are arranged by the interviewee to demonstrate one example. As we can see, the complexity of the images increases from right to left. They are roughly classified in three groups by column according to complexity. Across groups, he thought it could be telling a story of different phases of the disease. One can inspect the graphs to see if there is any distinguishable trend when it evolves from phase to phase.

Figure 31 (c) shows an interesting annotation from one user in their previous work session. This user was trying to find a circular pattern and went extreme with it. This reveals the rich meaning of the topology in such a task. Annotations are also important to support thinking and communication.

3.2.3 *Fatigue Problem*

It is a well known problem that performing mid-air interaction without arm support leads to fatigue, which is referred as "gorilla-arm" effect [50]. The interviewee complained about this problem during the interview.

I noticed the lack of interaction support for the user to move items easily over a large distance. Using touch input to move items on a large surface required ample body movements, including walking and large arm movements. The amount of time and physical effort taken for simply arranging items seemed to be too high. The interviewee mentioned that the collaborators often passed the images around during the collaborative session.

3.3 INTERVIEW 3 - ARMY SOLDIER

Paper walls are used in the military to support collaboration among soldiers. As another possible use case of a wall display, I interviewed a soldier about the use of their paper wall, in order to understand how the collaboration is supported by such a setup. As an interview technique, I asked the interviewee about the life cycle of the paper wall.

A paper wall consists of a grid of transparencies holding paper documents. It is divided into several areas, including a map in the center, a timetable, todo lists and availability lists around. They are updated by different troops that are responsible for different subtasks. In the case of the interviewee, they had five groups of soldiers in charge of: personal, intelligence, operations, logistics and communication. Each of them have clear division of labor and responsibility, as well as hierarchical order. This divide and conquer strategy maximizes the parallelization of tasks. The ultimate goal is to improve executive efficiency.

Divide and conquer to improve efficiency

The paper wall is updated by each troop once the soldiers finish a subtask or observe new events. It is designed this way so that information about task progress is always available to all of them. Verbal communication between the troops is minimized. Furthermore, the information is collected and visualized altogether to help them make decisions.

Paper wall's role - show task progress

This interview reminds a common sense that work parallelization among multiple people can effectively improve efficiency. It also show an interesting use of placing information over space, which is to provide task progress information for asynchronous communication.

3.4 OBSERVATION OF A CONFERENCE SCHEDULING TASK

The previous interviews provided insights about interaction challenges users face with big data and their motivation for using a large wall-sized display to complete individual or collaborative work. Given the limited availability of high-resolution wall-sized displays in professional settings as well as their technical constrains, there are few opportunity to observe the use of such a setup by real users for real tasks.

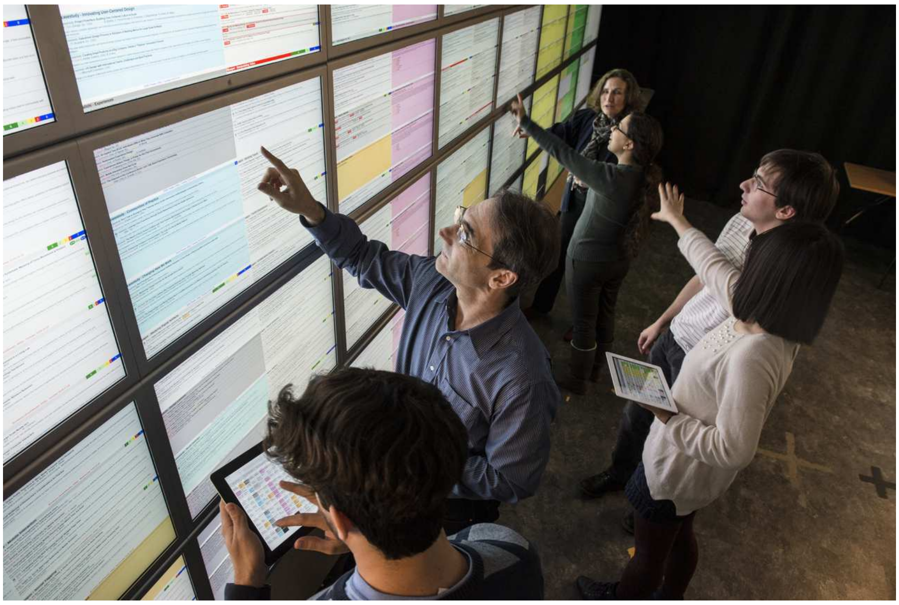


Figure 32: Scheduling a large conference on the WILD wall-sized display: Teams and individuals move close for detail or stand back for an overview. Photo copyright © 2013 Inria - H. Raguet.

I observed part of the process of scheduling a large international conference (CHI 2013). More than five hundred presentations of papers were to be scheduled into two hundred sessions, and organized into thirteen parallel tracks that lasted four days. The full program was displayed on a high-resolution wall-sized display (5.5×1.8 meter large with 20480×6400 pixels). A twelve point font size was used so that the full program could be displayed with paper titles and author names as a grid of sessions. Each paper occupied a 20×3 cm square, and each session contained four to six papers.

This scheduling task is constrained by many factors: papers with related topics should go in the same session so that the audience does not have to switch rooms in the middle of sessions; if strongly related papers are not in the same session, they shoud not be in the same time slot; presenters with multiple papers should not have their talks at the same time slot in different sessions; events with possibly large amounts of audience should be in large rooms; papers with awards should be distributed properly, etc.

A group of researchers performed this task on the WILD high-resolution wall-sized display at our lab (Figure 32). The complete conference program was displayed at once with the titles and the authors of the papers. Various colors and labels were used to visualize conflicts at different levels: hard conflicts, softer constraints and how "good" a session is according to the inter-event affinities.

Due to some technical and administrative constrains, the data was displayed on the wall-sized display but data manipulation, i.e., moving a paper to a different session or moving a session to a different

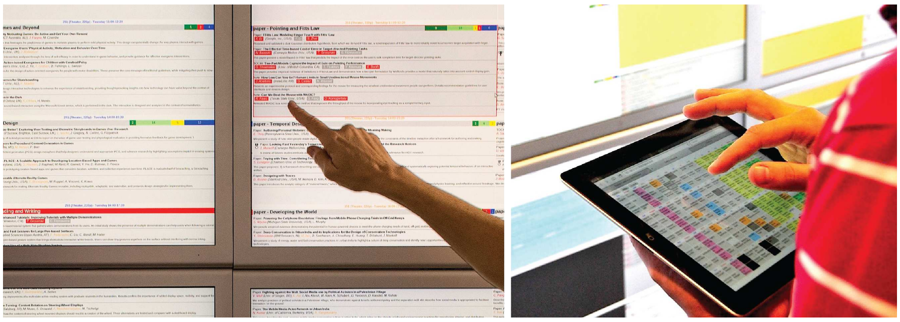


Figure 33: Left: a closeup view of the schedule program displayed on the wall. Right: the tablet interface for interacting with the wall display.

time slot, was performed on a separate desktop computer located in the same room. A program running on this computer allowed one user to move a presentation from one session to another. At the same time, another program [64] detected conflicts caused by each move and suggested alternative moves based on affinity matching. Due to the relatively small screen of the desktop computer (30”), the program could only display an overview table without any detail about the sessions.

3.4.1 *The Scheduling Process*

The most frequently performed task by the organizers was to identify an in-conflict or misplaced presentation, find a better session for it and move the presentation to that session. Various visual representations are used to highlight important information on the wall display.

When presentations with common authors are scheduled at the same time slots, the in-conflict authors’ names were highlighted in red or orange colors (Figure 33). In the header of each session, a rectangular bar consisting of colored segments indicates the affinity between presentations in the session. The colors range from red, yellow to blue and green, representing the paper authors’ questionnaire responses about how well or poorly their paper and the other papers fit in the same session. The sizes of the color segments reflect the overall count of each score for all the papers within the session. For example, the larger the red segment is for a session, the worse the papers fit together.

Colored representation showing conflicts and session cohesion

To move a presentation from one session to another, a user - typically one of the conference organizers, went through the following steps:

1. The user looked at the display from a distance to get an overview, in order to spot a relatively “bad” session or an in-conflict author.

2. The user got close to the display to read the text and spotted an orphan paper. Due to the limited space on the wall display, she could read the paper abstract on a tablet by interacting with a zoomable interface on the tablet (Figure 33).

3. The user selected the orphan paper with the tablet interface, then the wall display was updated to show fitting scores given between this paper and other papers, as colored labels besides each of these papers.

4. Then the user searched on the wall for a session with the most positive responses, read the title of each paper and decided if the selected paper should be moved into one of these sessions. Decisions were made based on various factors, such as whether the session was full or if there was a poorly fit paper to be taken out.

5. If the decision was to move the paper, the user went to the desktop computer to edit the program of the conference, due to technical constraints. If the user could not find a proper session to move the paper, she either left it in the original session or moved it to a separate area of the wall display collecting papers to be rescheduled later.

6. If another paper was taken out from the destination session because of this move, a new scheduling task would start with this paper.

3.4.2 *Observed Issues*

3.4.2.1 *Overview and Detail*

A large high-resolution display offers the possibility to display a massive amount of data without needing digital zooming. However, detailed information can be too small to be read when users step back to see the overview. In the case of this conference program, it was impossible to have an overview (e.g. days, rooms, session types, full/empty sessions, etc.) and see all the details (paper title, authors and abstract) at the same time. Therefore this requires users to step closer to the display in order to see details and step back to see an overview. Part of the overview is lost when users get close to the display. Ball et al. [7] defined this as *Physical Navigation*.

lack of overview while being close to the display

This can also make it challenging for users to notice changes on the wall display. Visual feedback of changes cannot be perceived if it is outside of a user's view.

In this scheduling task, I frequently observed that one organizer would help another by standing close to the display and telling the paper title to another organizer when it was not convenient for him to come closer. Users coped with this difficulty through collaboration.

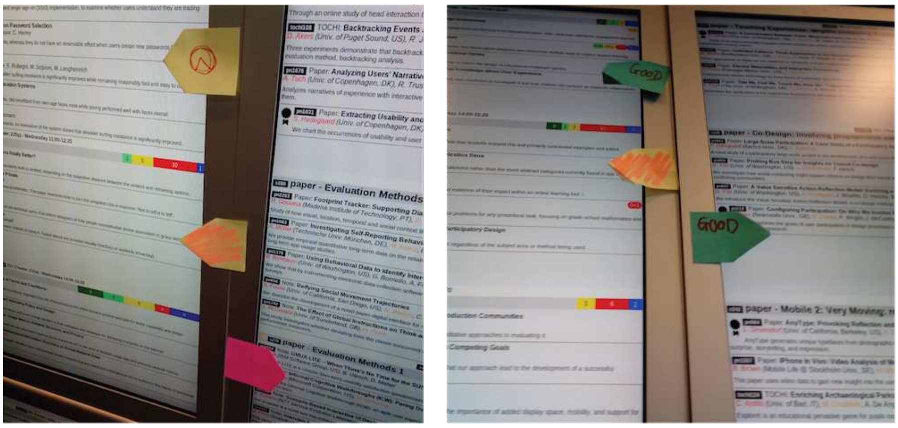


Figure 34: organizers left post-it annotations to keep track of positions of data items

3.4.2.2 *Loss Of Reference*

As physical navigation was needed, I observed the organizers' difficulty of keeping track of the position of a previously considered paper or session. This happened frequently, especially when an organizer needed to consider a sequence of moves or multiple sessions to decide if a paper should be moved to a different session.

Interestingly, the organizers used physical post-it notes to cope with this problem (Figure 34). Before walking away from a relevant session to look for another item, an organizer would leave a note to easily get back to it later. These physical references functioned surprisingly well, possibly because they can be easily personalized and are easy to distinguish from the digital data layer.

The bezels of the wall display also helped the organizers to find information. A huge spread-sheet data view like this does not have a geographical shape, contrary to maps. Therefore users need certain sorts of landmarks to find particular pieces of information or a previously viewed position. In the observed task, the colored highlights and backgrounds all helped the navigation to a certain extent.

It was also difficult to interact with the wall and desktop simultaneously. On the tablet, users are given an overview of the whole conference schedule. When a user found an interesting paper on the wall and wanted to select it on the tablet, she had to find the session first via the relative geographical position on the wall mapped to a grid cell on the tablet, which could lead to a high visual search effort.

3.4.2.3 *Collaboration And Conflicts*

A range of collaboration patterns were observed, from independent, parallel work to pairs working closely together. The organizers tended to group together, often in pairs, to perform a task together. Judging the fitness of a paper for a session is a complex task requiring advanced knowledge about topics. They frequently asked another person: "Have you seen something related to xxx?" or "Where should I put this one?" They shared their knowledge about the topic as well as memory about sessions and papers. When they worked in pairs, one tablet was shared between a pair. In this case, one person usually performed the interaction with the tablet interface to select paper. After the display is updated for showing information related to the selected paper, both of them searched at different areas of the display to find a better session for it. Above observed behaviors showed that users did not only frequently share their knowledge and memory, but also shared their interaction with the system.

collaborate to discuss difficult cases

Since the display is only updated to show information related to one selected paper at a time, conflicts did occur when multiple groups tried to update the display for the paper they were working on. The organizers verbally solved the conflicts by coordinating between groups.

3.5 DISCUSSION

The above sections describe interviews and observations that provide insights about users' needs and challenges when working with a large high-resolution display. This section discusses a few questions and answers derived from my understanding of real users.

3.5.1 *Why Use A High-resolution Wall-sized Display?*

The complexity of a data set or a task is the primary reason for people to be willing to view it all at once with an easy transition between overview and details. Facing a complex problem with massive data sets, researchers can in theory develop advanced algorithms to tackle the problems. However, human judgment is always needed at some point. For example, to find out important parameters for algorithms, users need initial understanding about the data set. This is indeed the goal of the sociologists I interviewed, which is to find various matrices to categorize the data so that they can later design algorithms. The task was difficult enough to require the combined opinions from different experts. The conference organizers mentioned that they came up with the idea of having themes for each day during the actual process of scheduling it. This would not have happened with a pure automatic process.

Why Working with Big Data Manually?

Such an exploration phase can be supported by viewing the data on a high-resolution large display. Users can step back to see the whole data set at once, yet step towards the display to access details. By doing so, they take advantage of spatial memory to associate content with positions that are coupled with their physical movement in space. In contrast, on desktop computers, virtual space can be disorienting as the mappings between positions and the physical move-

Physical navigation and spatial memory ments of the users depend on the input method used for the navigation. Physical navigation is more embodied than virtual navigation in this case.

Enable group work

Another reason to use a large display is to enable group work. Both examples of the sociologists and the conference organizers suggest that the primary motivation for collaborating in such an environment is to be able to discuss easily with a shared reference. In the case of the sociologists, they specially designed the visualization to allow different experts to communicate using the spatial arrangement of items. Users can take advantage of their awareness of the environment and between each other to smoothly transition between different collaboration styles. The spatial relations between the users and the data set as well as among the users have subtle implications, which can be subconsciously perceived and used in the workflow.

3.5.2 *Need Of Manipulation*

All the use cases I introduced above involve some kind of manipulation of the data items. Grouping is an elementary task for the graph theory researchers. In the case of the sociologists, moving images around is essential as the experts communicate through the spatial arrangement of the images and make stories based on their relative positions. Taking one data item and putting it in another container is the primary atomic task in the conference scheduling process.

In addition to the use cases I investigated, previous experience of using the wall display in our lab revealed a similar phenomenon. For instance, the display was once used to assign 145 submitted papers to 13 associated chairs (ACs) by two program committee chairs for an academic conference reviewing process. The display was divided into columns, one per AC, plus an area for papers to be assigned. The chairs could pick up a paper and assign to an AC by moving it to the corresponding column. The spatial layout of each column made it easy to determine the relative workload of each AC and adjust it.

Another application was developed for neuroanatomists to display several hundred 3D brain scans. Their goal was to compare, contrast and classify healthy and diseased brains. 64 high-resolution brain scans were displayed simultaneously on the wall display. The application provided drag-and-drop interaction for each brain image so that the users could place similar ones together to compare them.

Although the intellectual tasks in all these examples are different, moving items around and arranging them in a certain way appears to be a common interaction, integral to the task. The purpose is either to make sense of the data, to form an opinion, or to enact a decision.

In addition, physical fatigue is a real issue when users directly manipulate items with touch input on a large display, as mentioned by the sociologist. Interaction design should take this into account to support data manipulation on wall-sized displays.

4 ABSTRACTION OF REAL TASKS

By observing users engaged in real tasks, we can gain insights into their needs and the various challenges they face when interacting with wall-sized displays. However, to better understand the phenomena at play and trace the causal relationships, we need to conduct experiments where we can better control the conditions, in particular the task that users perform. This chapter describes the construction of an abstract task, based on the observation of the conference scheduling task. It eliminates the task-dependent cognitive part of a task and allows us to focus on interaction and collect time performance measures. The task is used in several experiments described in this dissertation.

Chapter 2 presented previous work on evaluating the benefits of using a large display to help users deal with large data sets: large displays encourage physical navigation, which improves performance in tasks such as searching, sense-making and pattern finding. However, such tasks do not involve manipulations to move data on the display. Tasks involving more manipulation, such as daily window management tasks, were evaluated in desktop setups where users sit in front of the display with limited locomotion. But there is a lack of investigation of data manipulation tasks with physical navigation, where users stand and move in front of a large wall-sized display.

Yet Chapter 3 showed that users need to navigate in large data sets and manipulate data items simultaneously to accompany a complex decision-making process. The potential benefits of wall-sized displays extend well beyond the passive visualization of extremely large images.

Therefore the question is how to gain a deeper understanding of interaction with large displays, not just visualizing or navigating content. A large body of existing research on large displays, especially for collaborative situations, applies ethnographic methodologies. Such studies provide insights by observing people using these setups and from their subjective assessment. While ethnographic approaches preserve the richness and complexity of real-world phenomena, they often lack quantitative measures that might help better identify causal effects.

This dissertation explores a different methodology that allows controlling factors and quantifying effects, by creating artificial tasks and instructions for conducting formative experiments. In particular, I developed an experimental task, which was constructed by abstracting an observed real task, to operationalize several important factors for data manipulation tasks on wall-sized displays. This chapter describes how this task was elaborated and its variants.

4.1 QUANTITATIVE EVALUATION METHODS

4.1.1 *Classic Operationalizations*

Conducting controlled experiments using an abstract task and operationalized factors is a standard evaluation method in psychology and HCI. I introduce a few widely used examples in the following.

A classical example is pointing experiments [37], which aim to test a theory of human motor system embodied by Fitts' Law. In the original experiment, participants perform a reciprocal tapping task, where they tap two rectangular metal plates alternately with a stylus. The movement tolerance and amplitude are controlled by the width of the plates and distance between them. This results in the widely used Fitts' Law Model, which reveals the correlation between movement time and the size and distance of pointed target. Fitts' Law has also been tested with everyday user interfaces in various conditions, and has been shown to be strongly robust and predictive [20].

Several variants of this task have been considered in HCI. For instance, MacKenzie et al. [69] designed an alternative pointing experimental task using circularly arranged targets. Besides traditional time and error measures, this task allows to evaluate pointing accuracy by measuring various aspects of movement such as target re-entry, task axis crossing, movement direction change, orthogonal direction change, movement variability, error and offset.

Complementary to Fitts' Law, Accot and Zhai [1] developed the Steering Law as an evaluation paradigm for input devices. In their experiment the participants performed two types of steering tasks with straight and circular tunnels with five different input devices. With this quantitative model, they were able to classify these devices into three categories based on statistical differences in performance.

Card et al. [19] proposed the Keystroke-Level Model for predicting user performance time for a given task. This simple model evaluates time by counting keystrokes and other low-level operations such as the user's mental preparations and the system's responses. It has been tested with several systems and shown to be accurate and flexible enough to help practical design and evaluation.

4.1.2 *Quantifying Collaborative Behaviors*

Most of the work evaluating collaboration with large displays consists of observational studies identifying users' collaboration strategies [31, 111, 109, 57, 53]. Within the scope of studying collaboration with large displays, quantitative approaches have drawn less attention in the literature.

Pinelle et al. [90] describe a task modeling scheme to evaluate groupware. Quantitative measures are collected by counting actions according to the articulatory aspects of the coordination. Other methods attempt to measure the equity of contribution from group members, for instance by counting the interactions made by individuals [76] or by considering the Gini Coefficient [38]. Tan et al. [108] designed a job-scheduling task for evaluating coordination in group work. This task enforces collaboration and coordination and makes it possible to measure performance against an optimal solution.

In the case of wall-sized displays, a rare example is the work of Hawkey et al. [48], who conducted an experiment that varies proximity among participants and with the shared display. They showed that participants preferred to work closely together.

In the field of remote collaboration and crowd sourcing, André et al. [2] compared simultaneous and sequential group work when performing a creative task. Their results are quantified by ratings along several dimensions. They suggest that sequential work is more effective for large-size groups, which is likely due to the cost of coordination and communication involved in simultaneous work.

Qualitative and quantitative approaches are complementary and both are needed to better understand collaborative work and, ultimately to inform the design of more effective computer support.

4.2 CONSTRUCTION OF THE TASK

A real task typically consists of a high-level cognitive part including understanding and making judgments and decisions, and a low-level part focusing on the interaction of acting the task. These two levels are entangled and occurring simultaneously. The intellectual part of a real task takes place in the user's head and involves a combination of planning and decision-making that relies on the users' domain expertise (e.g. the thematic similarity between conference talks or similitude into two brains). The interaction part of the task involves physical actions and depends on users' ability to quickly navigate and access content. The time cost of the intellectual part is largely dependent on each user's knowledge and is unrelated to interaction performance. Hence, comparing real tasks may make it difficult or even impossible to evaluate the performance of interaction alone.

Need to eliminate the task-dependent cognitive part of a real task

Here the ultimate goal is to understand interaction with wall-sized displays. Therefore we need to separate interaction from the task-dependent cognitive task. This way, the gained understanding can benefit a wide range of real-world tasks. Moreover, to gain a deep understanding of causal effects, I constructed an experimental task where both the performance and elemental interactions can be quantitatively measured and compared. I started by extracting the interaction level of the real tasks, then I operationalized the factors that are important for the interaction.

4.2.1 *Extracting the Interaction Layer*

An abstract task is designed to capture the essential interaction elements that commonly occur in some real tasks (Chapter 3). The conference scheduling task, the insight finding task of the sociologists and the brain image classification task of the neuroanatomists involve three common components: a decision-making task, a display in which information is logically organized, and the need to move items from one position to another.

A simple classification task was chosen to operationalize the factors that affect interaction. Typically, a classification task requires users to put together items that belong to the same category. In our examples, items from the same class ought to be grouped together into containers. In the neuroanatomists' use case of comparing brain images, the grouping was rather free, while, for conference scheduling task, it was both constrained within a session (corresponding to a container) and across sessions. The idea was to take a middle ground between the simpler and more complex examples.

Take common elements of the real tasks

Our classification task is slightly more complex than just grouping items because containers have a limited capacity of containers. As in the conference scheduling task, the abstract classification task involves more containers than classes. For example, no more than four talks may be allowed in a single conference session, but several sessions may have the same theme. The advantage of designing the task in this way is that the difficulty of searching for the class can be controlled and varied. A user must place items having the same properties into the matching containers but, as containers have a limited

Classification task with controllable difficulty

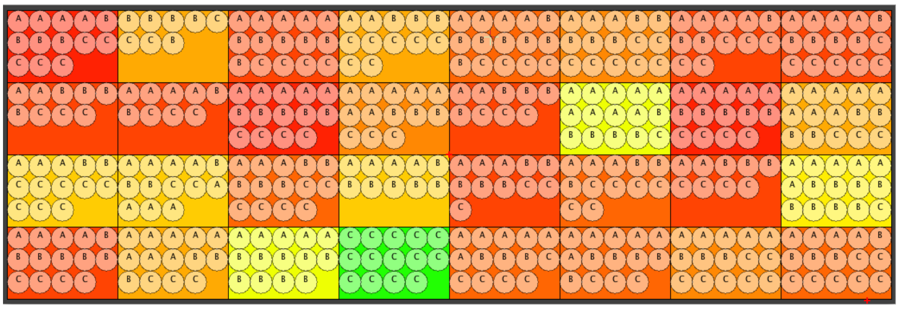


Figure 35: First version of the abstract task. Color shows the degree of misclassification of the container. The more categories of discs one container has, the more red it becomes. Green discs share the same category.

capacity, he must not put too many items in the same container. This adds a constraint that can be controlled.

A typical real-life example is when we must arrange books, CDs or DVDs by category in a bookcase. The bookcase has a given number of shelves, each of them with a limited capacity. If the items of a given category cannot fit on one shelf, they must be spread over several shelves. Scheduling courses for a university is another similar example. Courses are to be allocated into "boxes" of time slots multiplied by the number of classrooms. The constraints are that some courses can appear in the same time slot and some cannot, depending on the educational program. All these tasks include resource allocation as an essential element, especially from the interaction point of view.

4.2.2 *Eliminating Task-dependent Judgment*

The common part of the decision making process is that users need a way to judge if an item belongs to a class. As described before, this judgment is domain-specific and depends on the users' expertise. For example, grouping brain images depends on the similarity on many details, while determining if a talk belongs to a session involves many constraints. This process incurs heavy cognitive load and requires unpredictable time.

Letters represent classes

In order to properly control this aspect, a simple relationship should be determined and it should be well-known to all of the participants. Hence, we choose to represent each class by a different letter, which is labeled on the item. This makes the similarity criterion simple and unambigous: items with the same label belong to the same class. This variable moderates the difficulty of the cognitive task that would occur in a real task for estimating item similarity.

Figure 35 shows an initial version of the abstract task. The whole scene consists of 32 containers, each of which contains a maximum of 15 discs. Each disc has a letter in the center that indicates its category. The layout of the discs is generated randomly. Each container's color

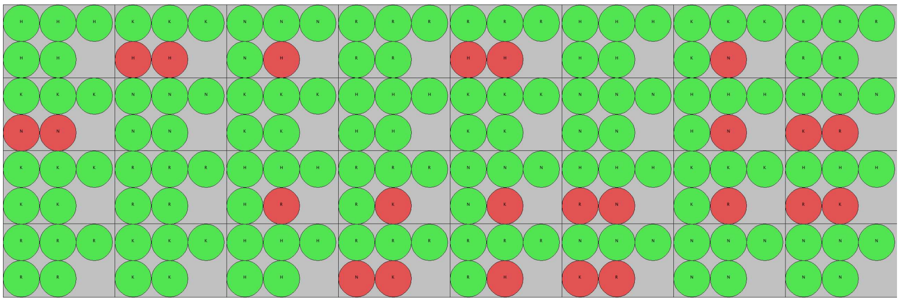


Figure 36: Experiment task. Display layout at the beginning of a trial in the *Large-Easy* condition.

reflects its degree of misclassification, which depends on the number of discs of the most dominant category compared to the number of discs of other categories that are located in this container. If all the discs located in a given container have the same category, the container is displayed in green. Otherwise, it is displayed in red and the intensity of the red color increases with misclassification. The user must move the discs from one container to another to reduce misclassification, thus to get as much green and as little intense red containers as possible.

4.3 MEET THE EXPERIMENTS

The goal is to use this task as a basic paradigm to design experiments that can be replicated with different factors. It features a simple data manipulation task with a controllable amount and distribution of data. It focuses on the interaction layer, so that we can evaluate and observe various phenomena affected by interaction, i.e., how users move discs across containers.

The experiments I conducted in this thesis include a single user experiment (Chapter 5) that evaluates the performance of manipulating scattered data on a wall-sized display vs. on a desktop computer, and a collaborative study with pairs of users (Chapter 6) that compares the performance and user behaviors across different collaboration styles.

The initial abstract task was modified to meet the requirements of these two experiments (Figure 36). The final task consists of moving discs between containers so that all containers hold discs of the same class (so that everything should be green at the end). Moreover, discs were individually colored instead of the containers. The discs of the most dominant category in a container are colored in green and the other discs in red. If there is no dominant category in the container, all of the discs would be red. Each time a disc is moved, the discs in both the original container and the destination container may change, depending on above rules. This makes it easier to spot the misclassified discs and the correctly placed ones.

Simplification

The task is therefore very straightforward. Participants must move a red disc into a container that contains green discs with the same label. Once put in a correct container, the red disc becomes green. While the initial abstract task may not have a definite correct answer for the whole scene, the experiment task has a clear and absolute goal: "make everything green". This way the performance time of each trial can be measured and compared. Errors need to be corrected to complete the trial, thus the cost of errors is included in the performance time.

Definite correct answer

The number of containers is 32 and they are organized in a 8×4 matrix to match the tiles on the wall and reduce any possible visual effect caused by the bezels. The maximum number of items in each container is 6. The diameter of each disc is such that 6 of them fill up one container. The overall number of discs is set to $32 \times 5 = 160$.

A pick-and-drop interaction is used for moving discs on both wall and desktop conditions. A participant first moves a cursor into the area of a disc, then one click picks it up and attaches it to the cursor. The disc follows while the participant moves the cursor. A second click inside a container drops the disc into it. The destination container automatically organizes the positions of its discs. Thus, the participant does not need to drop with high precision to keep the layout tidy. If the destination container is full, the disc snaps back to its original container. To ensure similar operation across participants, this interaction does not allow multiple selections at the same time.

Pick-and-drop

4.3.1 *Information Density*

The density of data shown on the display may have a large influence on interaction. Users may need to navigate more to see detail when there is more information per unit square. Therefore it is interesting to compare how users interact with data at different levels of density. Information density is directly affected by the amount of visible information on the display. However, if we display more data to increase the density, both information density and amount of information are increased. This would create two confounding factors that prevent us from attributing effects to one of them.

When a larger amount of data is displayed over the same screen space, the details have to be smaller, forcing users to navigate to, e.g. read text in smaller font. Therefore in order to operationalize information density while keeping the amount of displayed data constant, we chose to vary font size while using a single letter to show the category of a disc. This allows us to compare interaction phenomena at different levels of information density, while independently controlling the amount of information, e.g. by the number of discs.

Font size operationalizes information density

4.3.2 *Task Difficulty*

For our task, the difficulty is affected by several parameters:

- number of items,
- number of categories,
- number of containers,
- layout of items.

I chose to operationalize task difficulty with the number of categories, while reducing the variance of the rest parameters as much as possible.

Control the task difficulty

The number of red discs was chosen based on pilot tests considering the length of the experiments. Each task begins with a configuration of 24 misclassified discs (Figure 36). The red discs are evenly distributed across classes while being randomly distributed among the containers. 8 containers initially contain 2 red discs and the other 8 containers hold 1 red disc.

A completely randomly generated layout can take a long time to sort. An overly long experiment can cause a fatigue effect, thus should be avoided. Therefore an initial configuration is generated such that part of the items are already classified. Since one trial consists of moving all the misclassified discs into correct containers, the initial layout, meaning the number of red discs and the distances to their correct containers, influences the time needed for each task. Particular layouts are generated under constraints for each trial and were used for all the participants. This design has ecological validity, since in the real world task I observed that users indeed built upon an initial classification generated by the computer or made by collaborators. It is also one of the cases of using a large display for a classification task, where the users' task is to correct the errors in the computer-generated results.

Start with partially correct layouts

In order to minimize any effect caused by the differences between layouts, a minimum number of layouts are generated and selected. More layouts for some conditions and replications are created by permuting the letters and flipping this layout. Horizontal, vertical symmetry flipping and central rotation are applied to generate different layouts. This ensures that layouts with the same information density and task difficulty have similar structure while being visually different.

Minimize layout differences

4.3.3 *For Collaboration*

This task consists of atomic tasks that can be done either by a single user or by multiple users. It affords flexible levels of parallelization of task execution. This allows us to operationalize different styles of collaboration, ranging from purely parallel work with no coupling between users to closely coupled collaboration where users perform the atomic tasks together. The details of the operationalization of collaboration will be described in Chapter 6.

4.4 SUMMARY

This chapter introduced an experimental task to evaluate and observe interaction phenomena for data manipulation tasks on large wall-sized displays. The experimental task is based on observations of real tasks performed on a large display. The task was inspired by my observation of a real conference scheduling task performed on the wall-sized display (Chapter 3). It eliminats the domain-specific human judgment and focuses on the interaction level, so that the performance of setups can be measured and compared across experimental conditions. The task is rich yet controllable with several factors, thus allowing to quantitatively compare not only the results but also the procedure and user behaviors while performing the task.

This task is used in several experiments described in the following sections of this dissertation. Each experiment uses different factors and control parameters to suit the research questions and practical considerations. After this work was published [67], Jakobsen and Hornbæk [58] used this task to test the effects of locomotion on interaction with a wall-sized display.

"Evaluation is the worst form of HCI research except all those other forms that have been tried."

Shumin Zhai – CHI Place (2003).

5 BENEFITS OF PHYSICAL NAVIGATION

This chapter describes two controlled experiments that evaluate the benefits of a high-resolution wall-sized display for a data manipulation task. It compares single users interacting with a wall display versus with a desktop computer, for completing an abstract classification task. We found that the data manipulation is much faster with a wall display than a desktop computer for scattered and dense data sets. I will introduce the experiment design, procedure and insights gained from the results. [1]

While physical navigation in front of a wall-sized display may allow users to better use spatial memory and easily switch between views at different levels of detail, the virtual navigation imposed by a desktop interface may be disorienting and require additional effort for navigation, eg. pan-and-zoom, besides data manipulation.

However, physical locomotion and walking can be more time-consuming and tiring than virtual navigation. Manipulating with standard desktop input devices might be more efficient than using mid-air techniques on a wall-sized display [81]. So it is unclear which approach could be more efficient in terms of the interaction.

Therefore, the goal is to build upon previous work and gain more understanding of the trade-offs between these two types of navigation. The interest is to understand the advantages and disadvantages of physical navigation enforced by a high-resolution wall-sized dis-

1 This chapter is a revised and extended version of our publication at CHI [67].

play, comparing to virtual navigation enforced by a desktop computer.

This chapter delves deeper into the questions raised by the observations of real users and systematically evaluates the advantages and disadvantages of manipulating data on a wall-sized display. A formal experiment is designed to test and compare the efficiency of interacting with a wall-sized display versus with a desktop setup, for a data manipulation task.

5.1 COMPARING WALL DISPLAY AND DESKTOP

The first experiment investigates the trade-offs between physical navigation enforced by a wall display and virtual navigation required by a desktop screen. The abstract task described in Chapter 4 is used for this experiment to compare the *Wall* condition and *Desktop* conditions.

5.1.1 *Choices Of Input*

In front of the wall display, participants do not virtually zoom the scene, instead they physically approach the display to read the labels and find the target container while moving a disc (Figure 37). The cursor is controlled with a 13×13cm Apple Magic Trackpad[2]. A single tap triggers a click on the scene. It weights 165 gram with batteries.

The experiment software runs on a front-end computer, which runs the rendering part on the wall display. The trackpad is connected to the computer via Bluetooth. The computer displays a scaled-down representation of the scene on a full screen window with the same size ratio as the wall. The position of the cursor on the window (which is 2560 pixel wide) is linearly mapped to the corresponding position on the wall. This approach takes advantage of the CD-gain transfer function provided by the operating system and allows participants to easily move the cursor over large distances on the wall. As the sizes of the discs and containers are relatively large, this technique also provides sufficient precision for pointing.

Control a cluster display with a trackpad

On the desktop, participants need to pan and zoom to read the labels and to navigate in the scene while performing pick-and-drop operations (Figure 37). The screen is a 30 inch Apple Cinema Display[3] (2560×1600, 100 dpi) of the same model as the ones used in the wall display. An Apple Mighty Mouse[4] with default acceleration is used for input. The mouse wheel controls the zooming level.

The reason for choosing different input devices for the *Wall* and *Desktop* was to preserve external validity. Standard mouse input was

Rationals of using different input

2 https://www.apple.com/magictrackpad/
3 http://www.everymac.com/monitors/apple/studio_cinema/specs/apple_cinema_display_30.html
4 https://en.wikipedia.org/wiki/Apple_Mighty_Mouse

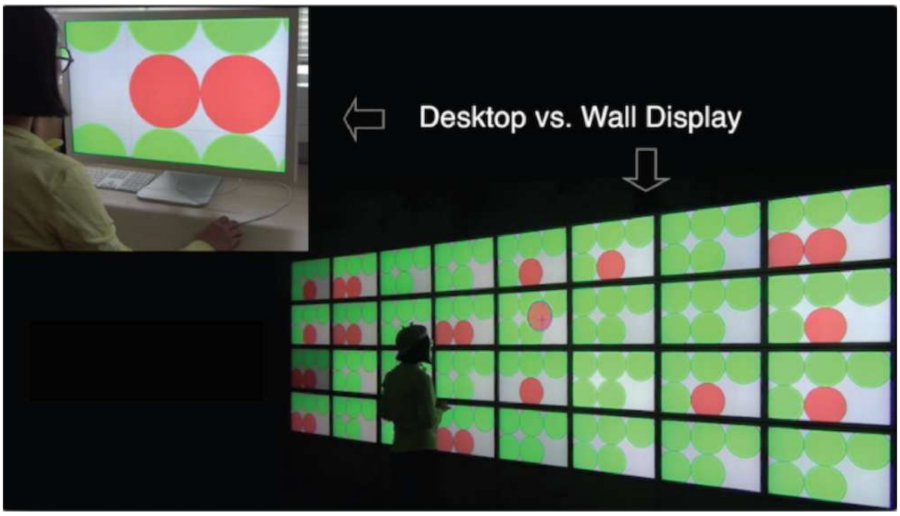


Figure 37: Wall vs Desktop conditions. A participant controls a cursor with a trackpad in front of the wall display and a mouse on the desktop.

chosen for the desktop setup to ensure comfort and efficiency and that any superior performance of the wall display could not be caused by using a non-standard or suboptimal input device for controlling the desktop. I am not aware of any literature suggesting a more efficient input device than a mouse for daily desktop tasks. In addition, an informal pilot study also suggested that using the mouse was slightly faster on *Desktop* than using the same trackpad as in the *Wall* condition.

Since research on wall displays is rather recent, there is not yet an acknowledged standard input device for this setup. Previous work [82] shows that using a handheld device for relative pointing is as efficient as known mid-air pointing techniques for remote pointing tasks. A trackpad was thus chosen for this experiment. This solution also avoids data noises and other technical problems that might be caused by a motion-tracked pointer or a Wiimote when participants move. The light weight and small size of the Magic Trackpad, compared to commercial tablets at that time, minimizes the effect of physical fatigue from carrying input devices.

5.1.2 *Factors*

As described in Section 4.3, the experiment task is an abstract classification task, for which participants need to pick a red disc and move it to a container of green discs with the same category. The category is annotated in the center of each disc as a letter label. The two factors chosen for this experiment is the *Label Size* and the *Number of Categories*.

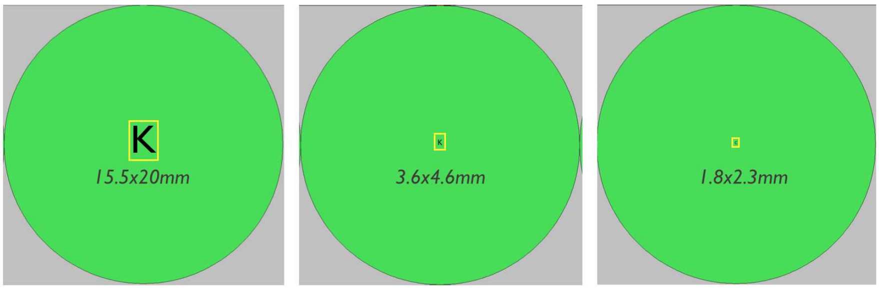


Figure 38: Label sizes in the experiment for *Large*, *Medium* and *Small* on the wall display

5.1.2.1 *Label Size*

Three levels (Figure 38) were chosen for the size of the labels (LABEL-SIZE factor). The letter sizes on the wall display and at the maximal zoomed-in level on the desktop were as follows:

- *Small*: standard computer font 12 point, letter size 1.8×2.3 mm;
- *Medium*: twice of the size of *Small*, letter size 3.6×4.6 mm;
- *Large*: 100 point computer font, letter size about 15.5×20 mm.

Rationale of the choices of label sizes

The *Large* size was chosen such that characters had the same size as the *Small* size when the whole scene was scaled down to fill the desktop screen (a 30-inch Apple Cinema Display, as said above). This would ensure that neither physical walking in front of the wall nor virtual zooming on the desktop was required in this condition. The *Medium* size was chosen so that the labels could be read when the zoom level allowed to display 9 containers on the desktop screen. Meanwhile on the wall display, users with normal or corrected vision could read the labels of 9 containers without walking. The rationale of such a choice is to maximumly cover a range of situations. The underlying hypothesis here is that the situation requiring walking might lead to a different trend compared to the situation where only locomotion (head rotation and body leaning) is required for performing the task.

5.1.2.2 *Number of Categories*

The difficulty of the task was controlled by the number of categories. Two levels of the DIFFICULTY factor were chosen:

- *Easy*: 2 categories labeled "C" and "D";
- *Hard*: 4 categories labeled "H", "K", "N" and "R".

These letters were chosen according to the BS 4274-1:2003 vision test standard [18] to guarantee equal legibility.

To ensure an equal difficulty of the tasks, the layouts were selected among randomly generated layouts under the following constraint. The distance between two containers was the Euclidean distance, and the unit was defined as the size of a container. Thus the distance between two adjacent containers was 1. For *Easy* tasks, the average distance between a red disc and the closest suitable container was between 1.25 and 1.46. For *Hard* tasks, this average distance was between 2.5 and 2.7. The goal was to create different situations where a participant could perform the task with or without necessary walking in *Medium* conditions. They were used in counter-balanced conditions within and across participants. All participants repeated the same set of layouts in different orders.

Generate and choose the layouts

5.1.3 *Experiment Design*

Existing literature suggests benefits of physical navigation[7, 107, 115, 6]. Based on this literature and the observation of real tasks (Chapter 3), three hypotheses were formulated:

- H1: *Wall* performs better than *Desktop* for smaller labels;
- H2: *Wall* performs better than *Desktop* for harder tasks;
- H3: *Desktop* performs better than *Wall* for larger labels and simpler tasks.

5.1.3.1 *Participants*

Twelve volunteers aged between 20 to 30 were recruited. Five of them were female. Four of them had normal vision and eight of them had corrected near-sighted vision. Seven of them used a trackpad daily; two had never used one.

5.1.3.2 *Apparatus*

In the *Wall* condition, the large display, sized 5.5m × 1.8m, is made of an 8×4 matrix of 30 inch Apple Cinema Display[5]. Its total resolution is 20480×6400 pixels (Figure 37). A cluster of 16 Apple Mac Pro computers running Mac OS X, each with two graphics cards, communicate via a dedicated high-speed network. The cluster is controlled by a front-end computer connected to the network. To track the participants' position movements, infrared retro-reflective markers are attached to a hat worn by the participants. A VICON[6] motion-capture system tracks the 3D positions of these markers with 1mm accuracy. In the *Desktop* condition, one 30 inch Apple Cinema Display is provided.

5 https://support.apple.com/kb/SP79?locale=en_US
6 http://www.vicon.com/

As explained in the previous section, a cursor for pick-and-drop interaction is controlled by an Apple Magic Trackpad in the *Wall* condition, and by an Apple Mighty Mouse in the *Desktop* condition. On the wall, participants normally hold the trackpad with one hand and tap it with one finger of the other hand to pick or drop discs.

The experiment software is implemented using jBricks [88], a Java toolkit that supports applications running both on a cluster-driven wall display and on a regular desktop. It uses an open source ZVTM toolkit[7] to render graphical interfaces by issuing commands from the front-end computer to the cluster computers.

5.1.3.3 *Procedure*

The experiment has a [2×3×2] within-subject design with three factors:

- DISPLAY: display type, *Wall* or *Desktop*;
- LABELSIZE: label size, *Large*, *Medium* or *Small*;
- DIFFICULTY: number of categories (letters), *Easy* has two and *Hard* has four.

Prior to the study, a vision test and color-blindness test are taken to ensure all participants have normal or corrected vision. Participants start the experiment after reading a standard textual explanation of the task. The experiment is blocked into two sessions, one for *Wall* and the other for *Desktop*. Half of the participants start with *Wall* and the other half start with *Desktop*. Each DISPLAY block begins with four training trials to get used to the task and input device. Participants are instructed to complete the tasks as quickly as possible while trying to avoid dropping items into wrong containers. This is to ensure that participants drop the item after reading the label, instead of using a trial-and-error strategy.

The order of the DIFFICULTY and LABELSIZE conditions are counterbalanced across participants using Latin Squares. For each participant, the same sequence of trials and layouts is used for both the *Wall* and *Desktop* conditions. This is to minimize the potential effect caused by different trial orders between the DISPLAY conditions. The experiment lasts about one hour.

5.1.3.4 *Data collection*

288 measured trials (2 DISPLAY × 3 LABELSIZE × 2 DIFFICULTY × 2 replications x 12 participants) were collected in the experiment. The measures include Task Completion Time, *TCT*, and the number of pick-and-drop actions performed to complete each trial. Since mistakes

7 http://zvtm.sourceforge.net/

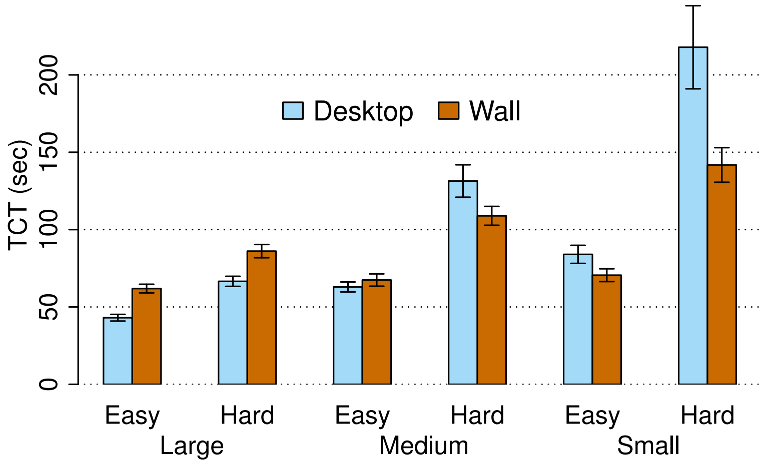


Figure 39: Task completion time (*TCT*) for each condition. Note that all bar graphs display the mean of each condition, with the error bars showing the corresponding confidence intervals.

must be corrected by participants to finish a trial, the cost of errors is included in the task performance time.

Physical navigation in the *Wall* condition is recorded using the motion-tracking system tracking a participant's head. Kinematic data of cursor movements, panning and zooming are logged on the *Desktop*. At the end of each DISPLAY block, participants fill out a questionnaire for assessing their subjective physical and mental load as well as frustration level, with 5-point Likert scales. A separate questionnaire is filled out at the end of the experiment to collect their preferences regarding different conditions. To avoid biasing in participants' rating, the DIFFICULTY is displayed as number of letters (2-letters or 4-letters) instead of *Easy* and *Hard* whenever it is shown in the instructions.

5.1.4 *Performance Results*

Task Completion Time (*TCT*) is the main performance measure in this experiment. Statistical analysis was performed on this measure.

Prior to performing the comparisons, the collected data was checked for outliers and normality on Task Completion Time. The measured trials were compared to the mean of the replications per condition and for each participant. 95% of the trials had within 15% differences to its mean. Three trials were more than 20% slower, namely 21%, 23% and 29%. No trials were removed as outliers.

No outliers removed

Shapiro-Wilk normality test was performed for each condition over the mean time of all participants. It did not show evidence of non-normality for all other conditions except in the *Desktop-Large-Easy* condition, with two participants being very slow. This does not affect the findings described below.

Effect	n, d	$F_{n,d}$	p	η^2_G
DISPLAY	1, 11	6.95	0.0231	0.07
LABELSIZE	2, 22	117	< 0.0001	0.62
DIFFICULTY	1, 11	229	< 0.0001	0.68
DISPLAY×LABELSIZE	2, 22	33.7	< 0.0001	0.29
DISPLAY×DIFFICULTY	1, 11	28.4	0.0002	0.11
LABELSIZE×DIFFICULTY	2, 22	62.9	< 0.0001	0.38
DISPLAY×LABELSIZE×DIFFICULTY	2, 22	12.1	0.0003	0.09

Table 1: Full factorial ANOVA of Task Completion Time with participant as random factor.

Figure 39 shows the mean task completion time for each condition with highlight on the differences between *Wall* and *Desktop*. Correspondingly, Table 1 presents the results of the full factorial ANOVA for the model:

$$TCT \sim \text{DISPLAY}\times\text{LABELSIZE}\times\text{DIFFICULTY}\times\text{Rand(Participant)}.$$

As we can see, all main and interaction effects are significant ($p < 0.05$). The η^2_G statistic measures effect size. 0.02 is considered as a small effect size, 0.13 as medium and 0.26 as large [24], although Bakeman [5] encourages each field to develop its own guidelines. Apparently in our case the effect sizes vary across conditions. Most of the effect sizes can be considered large, except for DISPLAY and the triple interaction where it is rather moderate.

Our analysis focuses on the differences between the DISPLAY conditions. Given the significant interaction effects, we compare *Wall* and *Desktop* with t-tests on *TCT* with Bonferroni correction [32] ($n = 6$). They are performed for each LABELSIZE×DIFFICULTY condition. The findings are listed in the following:

1. For *Large* labels, *Desktop* is faster than *Wall* for both *Easy* ($p < 0.0001$, 30.5% faster) and *Hard* ($p = 0.0001$, 17.1% faster);

2. For *Medium* labels, there is no significant difference for *Easy* ($p = 1$), whereas *Wall* is faster than *Desktop* for *Hard* ($p = 0.0222$, 22.7% faster);

3. For *Small* labels, *Wall* is faster than *Desktop* for both *Easy* ($p = 0.0315$, 16.0% faster) and *Hard* ($p = 0.0059$, 34.9% faster).

Our hypotheses are supported by these results. *Desktop* is faster for *Large* labels (H3) and *Wall* is faster for *Small* labels (H1), with both *Easy* and *Hard* tasks. This is a source of the DISPLAY×LABELSIZE interaction. DIFFICULTY seems to affect the magnitude of the differences between *Wall* and *Desktop*. For instance, the *Wall* shows a larger advantage for *Hard* tasks than for *Easy* tasks in the *Small* label condition (H2). In the *Medium* label condition, the *Wall* and *Desktop* perform closely in the *Easy* condition, while the *Wall* is faster in the *Hard* condition. These are sources of the DISPLAY×DIFFICULTY interaction.

We can see in Figure 39 that *TCT* increases with higher DIFFICULTY and smaller LABELSIZE. At the same time, the effect of LABELSIZE accelerates with increasing DIFFICULTY. This could explain the LABELSIZE ×DIFFICULTY interaction. Especially in the *Small-Hard* condition, the *Wall* performed 35% faster than the *Desktop*. This means that complex tasks with a high information density become exponentially hard to perform on a desktop, while a wall-sized display eases the effort of managing them.

However for the *Easy* tasks on the wall, the *TCT* on different LABELSIZE seems to be very close. To confirm this observation, T-tests (Bonferroni correction, $n = 12$) comparing *TCT* of the three LABELSIZE were performed for each DISPLAY×DIFFICULTY condition. On the *Wall-Easy* condition, it showed no significance between *Large* vs *Medium* labels and *Medium* vs *Small* labels. Despite this, the differences between label sizes are significant ($p < 0.005$) for all the other conditions. This suggests that, on the *Wall*, the label size does not affect the performance as much for easy tasks. But this is not the case for *Desktop*.

Label sizes have less effect on the wall for easy tasks

5.2 UNDERSTANDING THE CAUSES

The previous section shows solid results suggesting superior performance of a wall display comparing to a desktop for complex tasks with a large amount of data. This section analyses other measures to understand the causes of this finding. We seek differences among equivalent or comparable measures between the physical and virtual navigation, which include the number of pick-and-drop actions, angular sizes of labels, physical move distances, the reach range and the trajectories.

5.2.1 *Number of Pick-and-Drop Actions*

Differences in the number of actions for completing a trial can cause major performance differences. Therefore we compared the number of pick-and-drop actions used in each condition.

The results show no significant difference between *Wall* and *Desktop* overall. Since there are 24 red discs in each trial and multiple selection is not supported, 24 is the optimal number of steps. The data shows that all participants were able to solve the task in a more or less optimal way - 25.3 ± 0.20 pick-and-drop actions per trial on average.

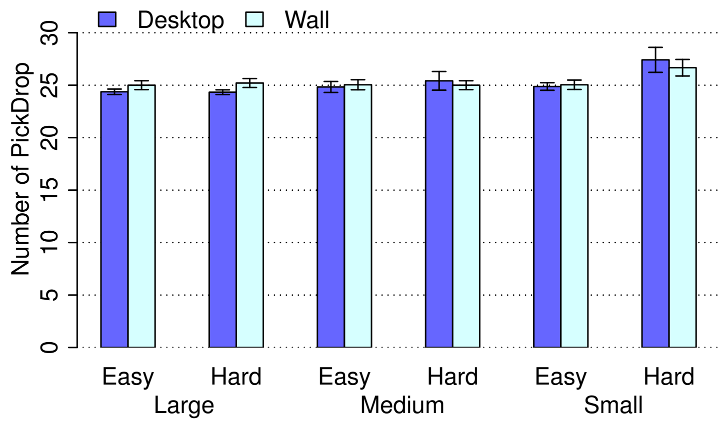


Figure 40: Number of pick-and-drop for all conditions.

Regarding the other conditions, the only significant difference is between the *Small-Hard* condition (27.0±1.43) and all other LABELSIZE× DIFFICULTY conditions (24.91±0.15) (Figure 40).

5.2.2 *Angular Size of the Labels*

This experiment operationalizes information density with LABELSIZE. Smaller sizes enforce participants to get closer to the data scene through either physical navigation on the *Wall* or virtual navigation on the *Desktop*.

Physical navigation in front of the wall would causes shrinkage and distortion of the perceived content, due to the changing position and view angle of the participants. On the desktop, the labels at a virtual zoom level may be rendered badly by the graphical rendering engine, which can cause them to be perceived differently from the equivalent of physical "zoom" level. This might degrade the performance on one of the setup, as participants might need to "zoom" closer to the scene on one than necessary.

In order to assess this effect of physical vs. virtual navigation, we computed the angular sizes of the labels when a disc is picked, taking into account the distortion effect of viewing the labels with an angle. Correspondingly for the desktop, we took a similar measure based on the assumption that the user's eyes are 60 cm away from the screen.

Table 2 shows the average angular width of the labels at pick time in arc-minute for the *Medium* and *Small* labels on each DISPLAY setup: These values are around 5 arc-minute, corresponding to 20/20 vision acuity. The differences between *Wall* and *Desktop* are small. There is no considerable differences in terms of visual acuity between physical and virtual navigation. This suggests that in both display conditions, participants optimized their navigation according to their visual acuity.

DISPLAY	*Easy-Medium*	*Hard-Medium*	*Easy-Small*	*Hard-Small*
Wall	6.78 ± 0.37	6.74 ± 0.32	3.90 ± 0.16	4.20 ± 0.17
Desktop	5.78 ± 0.26	6.27 ± 0.25	4.22 ± 0.17	4.76 ± 0.24

Table 2: Average angular width of the labels at pick time in arc-minute

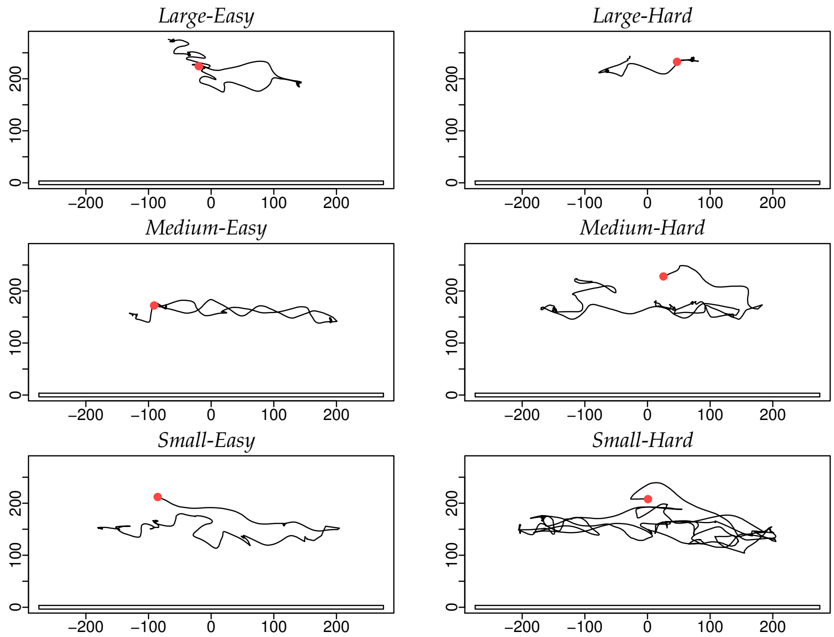


Figure 41: Head trajectories of participant P06 in front of the wall for the first measured block for each LABELSIZE×DIFFICULTY condition. This is a bird's eye view of the room with the wall display at the bottom along the x axis. The y axis represents the orthogonal distance of the head to the wall in centimeters.

5.2.3 *Movements of the Virtual Viewpoint and of the Participants*

The experiment is designed such that the exact same virtual scene is rendered on the wall and desktop display. The desktop screen is of the same type as the screens composing the wall display, so that the same pixel density is provided in both conditions. This ensures that the exact same pixels are displayed on the desktop at the maximum zoomed-in scale and on the wall. Moreover, the physical navigation of a participant is recored by tracking the head position and orientation, while the virtual navigation is recorded as viewpoint movement by using the ZVTM toolkit Therefore the physical and virtual navigation paths of a participant can be compared.

While the length of a participant's physical path for the *Wall* condition is computed (Figure 41), two length measures are computed for the viewpoint movements in the *Desktop* condition:

- *Screen space*: the physical screen space; a path is rendered to show the actual cursor movements while panning and zooming (Figure 42 top).
- *Scene space*: the virtual scene space that is rendered at the maximum level of detail; panning movements are scaled by the current zoom factor in order to render a virtual path in front of the scene. This path matches the equivalent physical navigation in front of the wall (Figure 42 bottom).

While Figure 41 and Figure 42 illustrate example paths of physical and virtual navigation, Figure 43 compares the average distances traveled by the participants according to these measures. Now we can compare the navigation trajectories of this participant on different conditions. Note that the physical navigation trajectories that we found are consistent with those reported in Ball et al. [7].

Overall, the amount of movement increases significantly both with smaller LABELSIZE and higher DIFFICULTY. These differences correlate with the differences in task completion time. Especially in the *Desktop* conditions, viewpoint movements in scene space increase acceleratingly for *Small-Hard*.

With large labels, there is no need for either physical nor virtual navigation to perform the task. Indeed in Figure 42 we can see almost no viewpoint movement in the *Large* condition. However, participants did move in front of the wall (about 482 cm per trial on average). There is no evidence of a dependency between this movement and DIFFICULTY. In addition, the large size of the *Wall* probably requires more head movements. These might explain why the *Desktop* is faster with large labels.

smoother and shorter paths of physical navigation due to the flexibility of body and head movements

For *Small* and *Medium* sizes, the lengths of the virtual navigation in *scene space* are longer than that of physical navigation (Figure 43). Moreover, the physical and virtual navigation paths draw different navigation patterns if we look at them from a bird's eye view (Figure 41 and Figure 42 odd rows). The head movements are smoother than the view point movements, which have hard angles and larger amplitude. For example, the position of the head has very low variability in the y dimension, between 5.4±2.0cm for the *Medium-Easy* condition and 15±11cm for the *Small-Hard* condition, while the viewpoint has a larger amplitude (Fig. 42, bottom row): from 101±33cm for the *Medium-Easy* condition to 164±33cm for the *Small-Hard* condition. These can be attributed to the participants' flexibility of leaning their bodies and moving their heads in physical navigation [7].

However, the movement distance calculated in *screen space* (Figure 43) for the desktop is shorter or close to that of the wall. Therefore in screen space, participants' physical movements on *Desktop* are not longer than on the *Wall* even in the case of *Small-Hard* condition. Thus the time differences between *Wall* and *Desktop* in smaller labels and harder tasks cannot be attributed to physical movements.

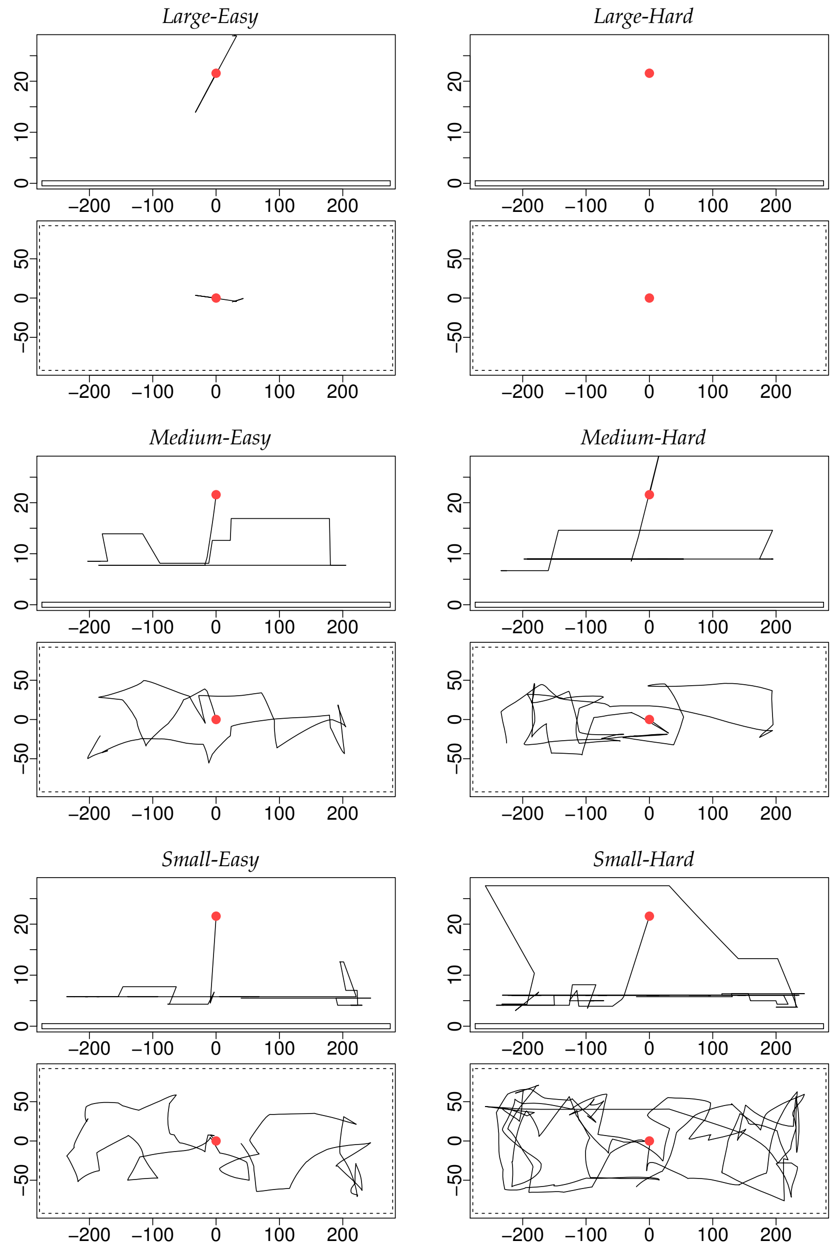


Figure 42: Movements of the viewpoint of participant P06 in the *scene space* on *Desktop* conditions. For each condition, the top graph plots the trajectory of the virtual viewpoint from a bird's eye view with the display on the bottom along x axis (zoom factor is converted to a distance on the y axis). The bottom graph plots the trajectory from a front view of the display (orthogonal projection of the viewpoint on the scene). All distances are in centimeters.

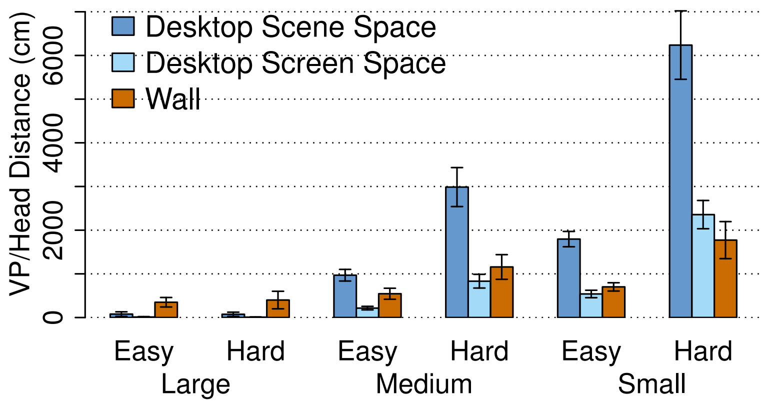


Figure 43: Average distance traveled by participants for each DISPLAY condition. (See the text for the definition of scene vs. screen space).

5.2.4 *Physical vs Virtual Reach*

The wall display creates an "immersive" interaction environment, which could affect users' body coordination for picking and dropping data items. In contrast on the desktop, users need to clutch for navigation and manipulating data items. Therefore we look at participants' ability to interact with distant targets and see if there is any difference between *Wall* and *Desktop*.

Although no eye tracker was used to collect the data of participants' orientation and viewing points, it is reasonable to assume that a participant did look at the item when it was picked up. For the *Wall* condition, we compute the relative position of a participant's head to the cursor position whenever an item is picked up, given the tracked head positions and logged cursor positions (Figure 44 left column). An equivalent measure is collected for the *Desktop* by computing the relative position of the view center to the cursor in the virtual scene (Figure 44 right column).

reach range reduces with smaller labels

Figure 44 visualizes participants' reach ranges in each condition. Note that the data points for each condition shown here includes all participants and both DIFFICULTY. We can see that the cursor points are more condensed for smaller LABELSIZE for both *Wall* and *Desktop*. This is reasonable since the smaller label sizes, operationalizing higher information density, draw participants closer to the wall or force them to zoom more into the scene on *Desktop* to be able to read the letters.

smaller reach ranges on desktop

Interestingly, the points are more closely clustered for the *Desktop* than for the *Wall* with *Medium* and *Small* labels. This indicates that the participants' reach range is larger on the wall than on the desktop, which might be a major reason for the reduced navigation distance in scene space. Indeed, the larger size of the wall-size display enables participants to reach targets at a distance by turning their heads and

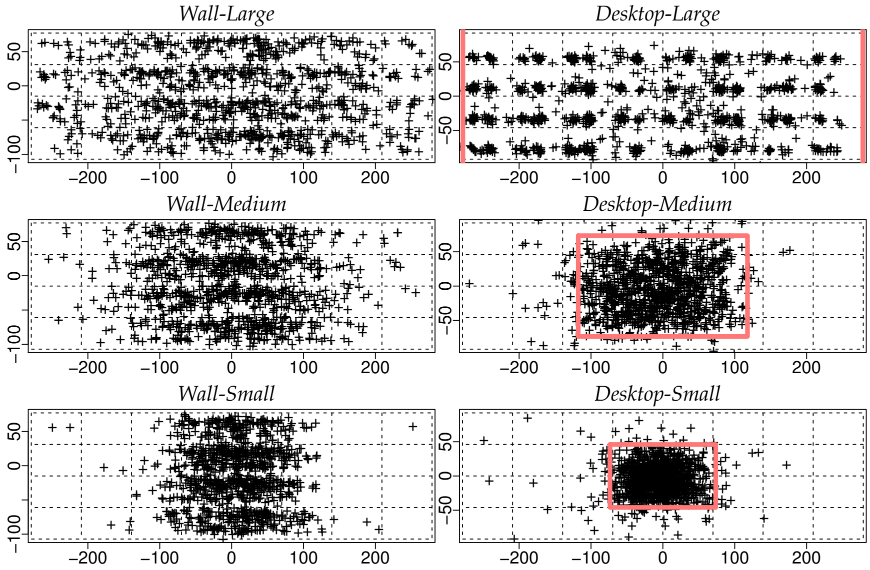


Figure 44: Positions of the cursor at pick time relative to the orthogonal projection of all the participants' heads or view points, which are translated to the center of the displays, for each LABELSIZE condition on wall and desktop. The dotted lines show the wall or the full scene and the containers for reference. The red rectangles in *Desktop* show the sizes of the physical screen relative to the virtual scene in average on each condition at pick time. Results at drop time are similar.

leaning their bodies. On the desktop participants must bring the target into view with pan-and-zoom.

Note that, in the *Large* condition, about 20 points are outside the wall boundaries and are thus not shown in the figure. These correspond to pick actions where the distance between the projection of the head and the cursor was greater than half the wall width, e.g. picking a disc on the left side of the wall while the user is on the right side.

At each pick time, we compute the area of the scene displayed on the desktop screen with the logged zoom factor. The red rectangles in Figure 44 for *Desktop* show this measure relative to the whole virtual scene size. The reach ranges on desktop almost fits in the view of the screen window, suggesting a constraint of window boundaries.

local versus distant operation

If we look closely, for *Small* labels, participants were picking items with about 4 containers (2×2) filling the screen, and about 9 containers (3×3) with *Medium* labels. This partly explains some performance differences in time. If we recall in Figure 39, there are large performance gaps between the *Easy* and *Hard* tasks for the *Desktop* conditions, especially with *Small* labels. In the *Easy* condition, most misplaced items can be moved to an adjacent container while in the *Hard* condition, the destination containers are further away, requiring the participant to pan-and-zoom during the pick-and-drop action. This distinction is very strong in *Small* label condition, while being less extreme with *Medium* label. Nevertheless, having 9 containers in the view still reduced the chances that the destination container was out of sight, thus reducing virtual navigation.

This finding is confirmed by Table 3: the average number of pan and zoom actions during pick-and-drop more than doubles between the *Easy* and *Hard* conditions, albeit with large variability, probably due to different participant strategies.

	Easy-Medium	*Hard-Medium*	*Easy-Small*	*Hard-Small*
number of pan	14.4±5.2	42.5±18	30.6±11	97.8±29
number of pan in pick-drop	9.71±3.4	18.8±7.6	21.2±5.8	42.0±19
number of zoom	5.46±2.4	12.2±7.3	17.0±6.2	47.2±23
number of zoom in pick-drop	4.92±2.2	8.54±4.3	14.0±4.7	30.9±20

Table 3: The average number of pan and zoom actions during pick-and-drop

5.2.5 *Subjective Assessment*

After each DISPLAY session, participants filled out a questionnaire with Likert scales to rate their degree of fatigue, mental load and frustration while performing the tasks. Their preferences between wall and desktop for each condition was collected with a separate questionnaire at the end of the experiment.

5.2.5.1 *Workload*

Figure 45 shows participants' physical and mental workload. We use pairwise Wilcoxon rank sum tests with Bonferroni corrections to test

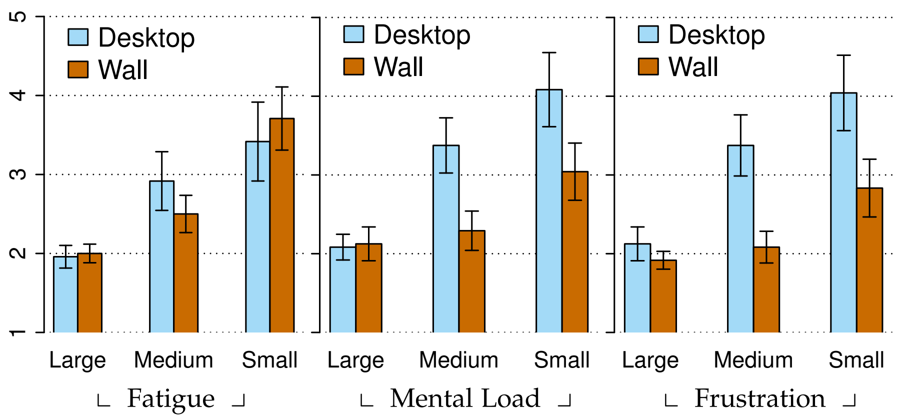


Figure 45: Physical fatigue, mental load and frustration on a five-point Likert item (1 is best, 5 is worst) for each LABELSIZE condition.

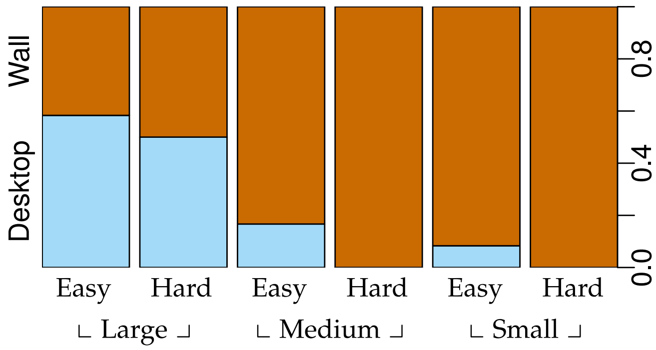


Figure 46: Proportion of the participants that preferred *Desktop* (■) or *Wall* (■) for all LABELSIZE×DIFFICULTY conditions.

significant differences among DISPLAY×LABELSIZE conditions and focus on the differences between *Desktop* and *Wall*.

The results on fatigue are not significant ($p \geq 0.5$), which is consistent with the lack of significant difference between physical travel on wall versus the movement on desktop screen space (Figure 43). This is in contrast to previous work that found physical navigation to be more tiring than virtual navigation [7]. The small labels with four letters (hard task) seemed to be tiring on both wall and desktop. One participant said: "*Desktop's repetitive work was somehow tiring. However the wall was very tiring after a while. If I could perform the wall task by resting my hands on a desk it would be ideal.*"

No difference in physical fatigue on wall and desktop

There is no significant differences for *Large* labels ($p \geq 0.9$) on mental load and frustration. However, for medium and large labels, the *Desktop* causes significantly higher subjective mental load ($p = 0.0007$ for *Medium*, $p = 0.02$ for *Small*) and frustration ($p < 0.0001$ for *Medium*, $p = 0.01$ for *Small*). A few participants mentioned mental load in memory: "*with the small labels, it was more difficult to get a mental map of the layout.*"

Heavier mental load and frustration on desktop

5.2.5.2 *Preference and Explanation*

Figure 46 summarizes participants' preferences between the desktop and the wall. Except for *Large* labels, almost all participants preferred the wall. The subjective preferences are more strongly in favor of the wall than the quantitative measure of time performance, where the *Medium* label sizes performed about the same in both environments. This may be due to the higher mental load and frustration on desktop, or novelty effect of using a wall-sized display.

Other reasons for participants' preference for the wall are yet to be identified. For example the participants' feedback suggests possible benefits of the wall for facilitating the use of spatial memory. 11 out of 12 participants tried to remember the positions of the items and / or containers, and 7 of them commented that it was easier to remember the positions when they were in front of the wall: "*...because*

Possible benefit of wall on spatial memory

I remembered the spatial location (in the room) of some particular rectangles"; "*I have better vision with the wall. It was more fun standing up and walking. It was also easier to remember where to go because of the movement memory.*" This is consistent with the participants' description of their strategies: in the *Hard* conditions, many of them tried to remember the positions of the misplaced items or the containers layout to reduce necessary navigation. Whereas in the *Easy* conditions most participants performed pick and drop between adjacent containers.

Some participants' comments help to explain the interaction effect on time performance between LABELSIZE and DISPLAY (Figure 39): "*The desktop with large labels is very fast, but exploring small and medium labels is painful*"; "*With the small and four letters in the wall, I didn't have to pan and zoom all the time, which was tiring. I just had to move a little bit, which was fine.*"

Other comments reflect the different sense of engagement between the desktop and the wall, e.g., "*For the desktop, I use the mouse [...] I feel I am under control. For the wall, I can move around, I feel I am a part of the interaction, and I feel I am controlling everything.*"

5.2.6 *Summary*

In this experiment, our quantitative results show a robust interaction effect between display type and information density (label size) with task difficulty (number of categories). The wall is up to 35% faster than the desktop with the highest information density in the harder condition.

As quantitative results, we analyzed and compared the wall and desktop for several measures, including the number of pick-and-drop actions, the angular size of the labels, the travel path and physical movement for navigation, and the reach range in the scene for manipulation. The major differences lie in the reach range and the navigation path. Users have larger reach ranges and smoother navigation paths on the wall than on the desktop with higher information densities. This suggests that the benefit of the wall can be attributed to users' ability of moving their heads and bodies flexibly to physically navigat in the data, while on the desktop users have to pan and zoom while performing pick-and-drop actions at the same time.

The qualitative results support and explain some of the quantitative results by showing no major difference in physical fatigue, but higher mental load and frustration on desktop and possible spatial memory support on the wall.

5.3 COMPARING TECHNIQUES ON THE DESKTOP

Multiscale navigation techniques were designed for small displays to visualize large data sets [22]. Presumably users could use an efficient navigation technique instead of a large display to perform the same task. Experiment 1 showed a strong performance advantage of physical navigation on a wall-size display, compared with pan-and-zoom navigation on a desktop setup for classifying a large scattered data set. Could these results be different with other types of multiscale navigation techniques?

5.3.1 *Multiscale Navigation and Display Size*

With a few exceptions [81], existing techniques for multiscale navigation have been mainly studied and deployed on desktop computers. A few studies investigate the effect of display size on multiscale navigation. Guiard et al. [43] compare small to medium display sizes for a target acquisition task with pan-and-zoom navigation. The larger sized displays show a minor performance improvement. Jakobsen and Hornbæk [59] evaluate the effect of three display sizes for three classic interactive visualization techniques, including overview+detail, focus+context and pan-and-zoom. They find similar performance between medium and large displays, indicating that larger is not always faster. One suggested reason is that focus+context and pan-and-zoom require more target searching time on larger displays. In addition, focus+context is found to be difficult on small displays.

This previous work suggests only a small or no benefit of larger displays using multiscale navigation techniques. However, these studies were conducted in desktop setups where users sit in front of the display. Moreover, the studied tasks involve only visualization or target acquisition, not data manipulation. In the context of this thesis work, we ask a different research question: if the performance advantage of the large display in Experiment 1 would remain if a different navigation technique is used on desktop.

5.3.2 *Comparing Three Desktop Techniques*

To answer this research question, a second experiment compares three desktop techniques with the same task in Experiment 1. We choose the state-or-art techniques based on the review of Cockburn et al. [22].

- the pure pan-and-zoom technique in Experiment 1 as baseline - *PanZoom*,
- an overview+detail technique - *PZ+OV* (Figure 47 left),
- a focus+context technique - *Fisheye* (Figure 47 right).

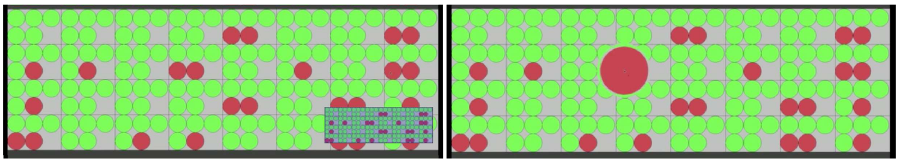


Figure 47: The *PZ+OV* technique (left) and the *Fisheye* technique (right) used in experiment to compare with pure pan-and-zoom

5.3.2.1 *Overview + Detail*

As shown in Figure 47, the *PZ+OV* technique adds a miniature view (the overview) of the virtual scene in the right bottom corner of the screen view (the detail view). While zooming in, the overview shows a rectangle that highlights the area of the detail view in the overview.

Many existing implementations of this technique allow the user to interact with the rectangle in the overview, eg. by dragging the rectangle or clicking at another area, to navigate the detail view. The literature suggests that adding an overview to a pan-and-zoom interface increases user satisfaction [51, 84]. Pietriga et al. [87] show that an interactive overview can be very efficient for search tasks.

Why not an interactive overview?

We informally tested an interactive overview, but found that it did slow down the performance. We noticed that with a data manipulation task, switching between picking and dropping on the detail view and navigating by clicking on the overview was too time-consuming. This suggests an interesting effect: the usability of navigation techniques can be different in data manipulation tasks than pure search or visualization tasks. Moreover, deciding to navigate with the detail view or the overview might add mental load to the participants. Therefore, the *PZ+OV* technique we use in the experiment has a passive overview fixed in a screen corner. Participants can only navigate in the scene by pan-and-zoom of the detail view.

5.3.2.2 *Focus + Context*

Fisheye lens [102] is another way to combine overview (context) and detail view (focus) in a single view. For this experiment, a fisheye lens is implemented as a permanent attachment of the cursor.

The cursor is at the center of the lens, which has the same radius as the discs (Figure 47 right). With a magnifying factor of 6, the lens makes the small labels readable when moved on top of a disc. The entire virtual scene is scaled down to fit the display. Pan and zoom are not provided in this case. When a disc is picked up, it is attached to the bottom of the lens, instead of on the cursor like in other conditions, in order to avoid occlusion.

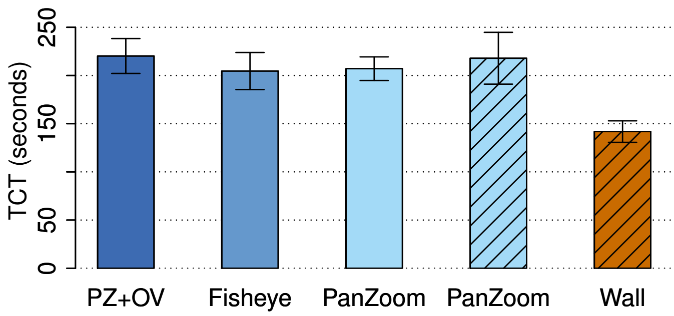


Figure 48: Task completion time (*TCT*) for the three techniques in Experiment 2. The hashed bars on the right show the results from the corresponding conditions of Experiment 1 for reference.

5.3.3 *Method*

We recruited 12 volunteers (6 female), aged 22 to 38, all with normal or corrected vision. Half had participated in Experiment 1. We use the same desktop apparatus and the same task as in Experiment 1.

The experiment is a within-subjects design with one factor (TECHNIQUE): *PanZoom*, *PZ+OV*, *Fisheye*. Trials are grouped by technique. The 6 possible orders are used once for each participant from Experiment 1 and once for each new participant.

In Experiment 1, the wall-sized display has the maximum performance gain than desktop on the *Small–Hard* condition. Since the goal of this second experiment is to test if the desktop setup can beat the wall with a different navigation technique, participants only performed tasks on the *Small–Hard* condition. The same layout configurations as in Experiment 1 are used.

Participants start with a training trial with *PanZoom* to learn or recall the task. Then, they perform one training trial and two measured trails for each TECHNIQUE. At the end of the experiment, participants rate their preferences and give subjective assessment about each technique. The experiment lasts about 35 minutes. The same measures are collected as in Experiment 1. Overall there were 72 trials: 3 TECHNIQUE x 2 REPLICATION x 12 participants.

5.3.4 *Results*

We analyzed the variance on the task completion time (*TCT*). The result reveals NO significant effect of TECHNIQUE: $F_{2,22} = 0.68$, $p = 0.5149$, $\eta^2_G = 0.03$.

Figure 48 shows that the three desktop techniques are very close. Compared to Experiment 1 on the corresponding condition, none of the three techniques reaches a performance close to the wall. Therefore, existing navigation techniques do not significantly improve the performance for our task.

The average numbers of pick-and-drop per trial are very close across the techniques and similar to Experiment 1. The *PanZoom* technique here has on average 25.4 pick-and-drops per trial, while *PZ+OV* has 26.0 and *Fisheye* has 25.6.

Regarding preferences, nine participants preferred the *Fisheye* technique, while three preferred *PZ+OV*. The noted reasons for preferring the lens include that they did not have to zoom and pan all the time. Although the lens was preferred by the majority, some participants complained that it was hard to focus on the labels with the lens, despite the size of the lens being large enough to cover a disc. This is probably due to the distortion caused by the high magnification factor, which was needed to make the labels readable while keeping the overview fit in the desktop screen. This suggests the limitation of lenses when a large amount of data is displayed.

As a side note, the abstract task we used is in favor of lens techniques, as there is only one letter to read in the center. Generally, lens techniques could be more problematic with a real task. For example reading a text area with distortion can be difficult. Variations of lens techniques have been shown in recent research, such as JellyLens [89] and SchemeLens [23]. They need to be adapted to various type of data.

Eight participants stated that the overview in *PZ+OV* was not very helpful. However some mentioned that it helped to locate the red discs and empty slots in containers. But it was also mentioned that the overview was not needed, as they could do so by zooming out.

In summary, this follow-up experiment confirmed that the performance benefit of a wall-sized display for manipulating scattered data sets. Although new techniques could be devised to improve the desktop condition, e.g., using multiple or adaptive lenses, the interaction for data manipulation adds complexity and needs to be considered in this case.

5.4 SUMMARY

This chapter described two experiments for evaluating the benefits of physical navigation compared to virtual navigation for a classification task. The first experiment compared using a wall-sized display to perform the task versus on a desktop setup. The results show a robust interaction effect, where the desktop is more efficient for easy tasks with low information density, while the wall display outperforms the desktop up to 35% for tasks with high information density. The second experiment further confirms this finding by comparing three navigation techniques on the desktop in the same task and exhibiting no significant improvement in performance compared to the pure pan-and-zoom used in the first experiment.

The data analysis of the first experiment attributes the benefits of the wall-sized display for such tasks to a form of embodied interaction that takes advantage of whole body movements and the ability to walk around for navigating data while manipulating the items with their hands (through an input device). The large display embeds the data scene in the physical space, making the navigation paths more smooth and natural. Users do not need to explicitely think about it when they "pan and zoom" by physically moving in front of the wall display, while on the desktop they need to take explicit actions with the mouse to navigate data.

This work is but a first step in understanding the interaction environment provided by wall-size displays. Besides evaluating their benefits in some situations, it triggers research questions that can be answered with future replications and new studies. For example, can we design a more efficient navigation technique on the desktop setup and take into account of the data manipulation aspect?

The next chapter will extend this series of work to collaborative situations. I will describe an experiment that studies multiple users performing the same classification task with various strategies and interaction support.

6 COLLABORATIVE INTERACTION WITH A WALL-SIZED DISPLAY

As wall-sized displays afford multiple users working collaboratively in a co-located space, the next step of my research attempts to understand collaborative interaction phenomena, with a focus on the interaction with the data and between the users. This chapter describes the methodology, procedure and findings of an experiment, which operationalises loosely and closely coupled collaboration and evaluates the effects of a shared interaction technique for collaborative data manipulation. The data analysis deepens our understanding of interaction efficiency and behaviour differences between loose and close collaboration for data manipulation.

Data manipulation is an integral element of a wide range of tasks, including many cognitive tasks such as sense-making [17] or problem-solving [57], in which users need to reorganize data items to solve the task. Chapter 5 showed the benefit of using a ultra-high resolution wall-sized display vs. a classical desktop setup for a single user manipulating a large amount of scattered data.

Wall-sized displays are well-suited to small groups of users working together on large amounts of data: multiple users can move freely in front of the display to get an overview or see details of different parts or perspectives. This could lead to a variety of collaboration styles, from purely parallel work with little interaction between users to closely coupled work in pairs or triads [57].

Physical navigation is one type of embodied interaction enabled by a wall-sized display for single users. How does the situation change in collaborative situations? What else can we take into account to design embodied interaction for collaboration in such a co-located space? Previous literature shows that co-located collaboration is facilitated by rich social resources such as group awareness [45] and face-to-face communication with direct deictic actions. I believe that the spatial and social awareness and skills can be taken advantage of, and need to be considered for designing embodied interaction techniques to support collaboration in such an environment.

I define a term *Shared Interaction Technique* to describe interaction techniques that combine multiple users' input actions to perform one operation. Such interaction techniques can be embodied when designed properly, as they could blend with the operation and communication flow between users while augmenting their capabilities. They should take advantage of the co-located resources while respecting users' social convention and collaboration strategies.

This chapter presents an experiment that aims to deepen our understanding of collaborative interaction on a wall-sized display in data manipulation tasks. The same experimental task (Chapter 4) is used to focus on the interaction level. We use instructions and layouts to operationalize collaboration styles and information distribution. The goal is to build on existing qualitative findings in the literature and go one step further, to separate the collaborative strategies and compare interaction phenomena among them. As a first stepping stone, this experiment studies one example of shared interaction techniques - drop-for-partner, in both loose and close collaborative situations.

6.1 BACKGROUND

The interview with sociologists revealed users' needs and problems when manipulating scattered data collaboratively (Section 3.2). Moving data over large distances is both tiring and time-consuming. Collaborative manipulation in large physical spaces can be made more efficient by providing shared interaction techniques, in which multiple users each perform a part of the action [77, 92]. These techniques may reduce users' effort of performing large-scaled movement, thus also improving performance and user experience. On the other hand, they may support new forms of division of labor. Nevertheless, the application of such techniques to wall-sized displays have received little attention. I am not aware of any formal study identifying their effects on collaborative behavior and productivity measures.

Designing effective interaction techniques for collaboration requires a deep understanding of situations involving complex phenomena when multiple users interact with the artifacts as well as communicate with each other [116]. However, little research has studied the interaction efficiency in co-located collaborative work with such settings. Previous work has studied how input techniques affect collaboration, such as touch and mouse input [52], or single-user interaction techniques like drag-and-drop, radar views and laser beam [80]. I am not aware of formative studies evaluating the effects of techniques enabling shared interaction for collocated collaborative work. The study reported in this chapter compares pairs manipulating data with and without a shared interaction technique in different collaborative styles. The goal is to gain a deeper understanding of their respective costs and benefits and to inspire the design of new interaction techniques to support collaboration.

6.2 OPERATIONALIZATION

The goal is to understand the phenomena at the interaction level when multiple users manipulate data in collaborative situations. Previous observational studies show that users collaborate with different coupling levels and switch among them fluidly [57, 109, 53]. Here we intentionally separate different collaboration coupling levels, so that we can compare phenomena in each of them. Our experimental task makes this possible, as it can be performed either by single-users alone or collaboratively.

The abstract task described previously (Chapter 4) is used in this experiment. It is a classification task where participants pick and drop items from one container to another after searching and simple judgement. Each trial consists of pick-and-drop subtasks, which can be performed sequentially or in parallel. Each subtask can be executed either with or without a co-worker's explicit help, making it possible to enforce and control different levels of collaborative coupling. In this experiment we give instructions to participants to enforce their strategies.

6.2.1 *Collaboration Styles*

Four collaboration styles are operationalized by crossing two dimensions(Table 4). One dimension is collaborative coupling, which is operationalized by task parallelization. By enforcing the subtasks (moving a disk) to be executed in parallel vs. sequentially by the pair of participants, we obtain loose vs. close collaboration. The other dimension is the availability of a shared interaction technique, called drop-for-partner, which is a simple technique designed for this task. We observe its effects in loose and close collaboration respectively: how does it affect collaboration, efficiency and subjective experience.

Collaborative Coupling \ Shared Interaction	Not provided	Provided
Loose collaboration	*LooseComm*	*LooseTech*
Close collaboration	*CloseComm*	*CloseTech*

Table 4: Four collaboration styles generated by crossing two dimensions. Loose and close collaborative couplings are operationalized by task parallelization. Shared interaction is enabled by the drop-for-partner technique, which lets a partner drop the disc picked by the other partner. When not provided, partners can help each other only by communicating with each other.

6.2.1.1 *Task Parallelization*

The overall task of classifying the red discs is performed by a set of individual pick-and-drops. These pick-and-drop tasks can be performed by a pair of users either in parallel or sequentially, representing two basic patterns of division of labor.

In the parallel execution cases, participants perform tasks independently of each other: each user picks one of the remaining red discs and puts it in an appropriate container. This operationalizes loose collaboration. In the sequential execution cases, the pair performs each pick-and-drop in tight collaboration. This style is enforced by allowing only one disc to be picked at any one time. This operationalizes close collaboration. While the parallel cases result in each participant solving a harder task concurrently, participants can benefit from the partner's help in the sequential cases.

6.2.1.2 *Shared Interaction Support*

Beyond the coupling level, collaboration in the present task is also affected by how users help each other. From the interview with the sociologist about their collaborative work, I found that the users frequently discussed about data items and passed them among collaborators. Inspired by this, I designed a simple technique to support collaborative interaction. It is called *drop-for-partner*. This shared interaction technique is provided for each coupling level and can thus be compared to situations without it.

Each participant can spell out the label of the disc they have picked. The collaborator may be able to help him if she sees or remembers one suitable destination container. Without additional technical support, she can indicate the destination container through verbal and gestural communication, for instance: "It's here!" With the drop-for-partner technique (Figure 49), instead of indicating the destination container, the collaborator can finish the pick-and-drop action on behalf of the first user by dropping his disc, e.g. with a dedicated button on her device.

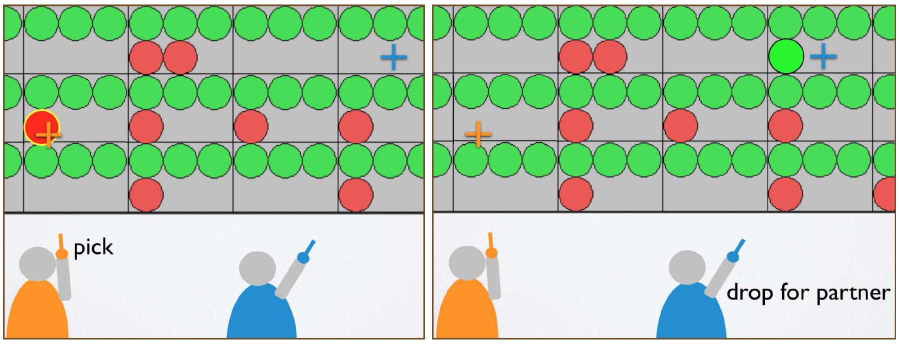


Figure 49: Drop-for-partner technique. The user in orange picks a disc (left), which can then be dropped by the coworker in blue (right) by clicking a button on his input device.

Drop-for-partner is an example of a shared interaction technique that provides minimal technical support for collaborative data manipulation. It is tailored to the pick-and-drop interaction. The purpose is to compare collaboration with shared interaction support and without it, i.e. with verbal or gestural communication only.

Both types of help require the participants to coordinate and synchronize their actions, so the time saved by using the co-worker's knowledge or action may be offset by the coordination overhead. We are interested in observing this trade-off with quantitative measures.

6.2.1.3 *Five Collaboration Styles*

Crossing the above two dimensions, *Loose* vs. *Close* Collaboration and *Communication Only* vs. *Shared Interaction*, leads to four collaboration styles. One additional baseline condition is added, *Divide&Conquer*, to contrast with the other four explicit collaborative styles. In order to avoid the influence of the condition names to the participants' behavior, different names are used during the experiment.

- *Divide&Conquer*: the pick-and-drops are performed in parallel and pairs are not allowed to communicate nor help each other (Figure 50-1). No spatial division is required, participants move freely in the entire space. The condition name shown to the participants is "*Separate*".
- *LooseComm*: the pick-and-drops are performed in parallel, but pairs are allowed to communicate and encouraged to help each other by indicating to their partner a correct container if they can (Figure 50-2). The condition name shown to the participants is "*Only Communicate*".
- *LooseTech*: the pick-and-drops are performed in parallel, but the pair can also help each other by using the drop-for-partner technique (Figure 50-3). The condition name shown to the participants is "*Drop For Partner*".

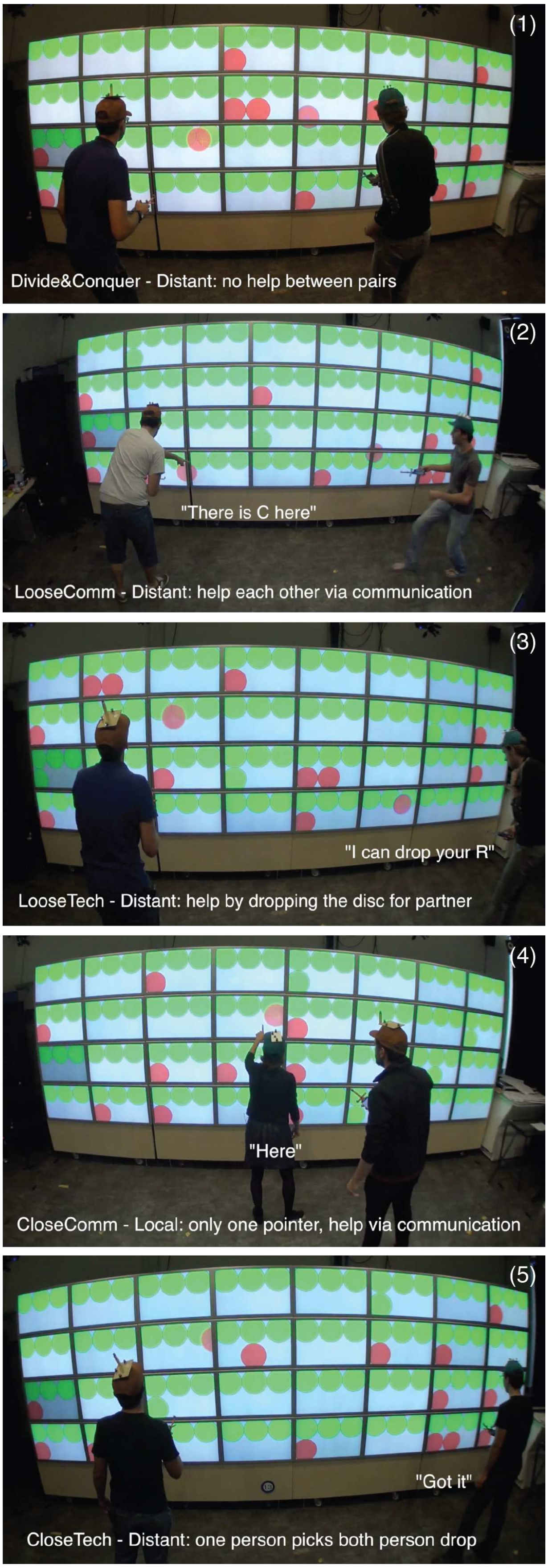


Figure 50: Representative pair behaviors with five collaboration styles. Text in quote is their conversation. Pairs help each other via verbal and gestural communication without shared interaction support, while using drop-for-partner instead when it is provided.

- *CloseComm*: the pick-and-drops are performed sequentially. This is enforced by giving each pair only one input device (Figure 50-4). Only one partner can perform the pick-and-drops, and the partner is encouraged to help, e.g. by searching for containers or items. The condition is shown as "*One Pointer*".
- *CloseTech*: the pick-and-drops are performed sequentially, while each participant has an input device. The sequential execution is enforced by allowing only one person, the "picker", to pick for each trial. The other partner can help with the task by dropping the disc picked by the picker using drop-for-partner (Figure 50-5). The partners switch their roles so each is the picker for half of the trials. The condition is shown as "*One Pick Both Drop*".

As discussed before, these conditions will let us to compare basic interaction phenomena among collaborative coupling levels: no coupling (baseline), loose coupling and close coupling. Beyond this, it allows to compare the effects of shared interaction on loose and close collaboration respectively.

6.2.2 *Layout Locality*

We used the smallest size (12 pt) from the previous experiment (Chapter 5), as it corresponds to the highest information density. Participants must be close to the display to read the labels and must move around to read the labels of discs that are further away than the adjacent columns. While 4 categories are used for hard tasks in previous experiment, 8 categories are chosen for all conditions in this experiment. Combined with the small label size, this results in a task hard enough so that collaboration is likely to be needed or beneficial.

LAYOUT is chosen as a factor to test the effects of information distribution on interaction. Two types of layouts, *Local* and *Distant*, operationalize the distribution of information on the wall by controlling the distance of misclassified discs to the closest correct container. With *Local* layouts, all pick-and-drops can be done between adjacent containers: The distance between a pick and a drop is short enough that users only need to search the area in front of them. With *Distant* layouts, all pick-and-drops involve containers in non-adjacent columns: users must move and search a larger part of the display to find a proper container.

These two types of layouts are likely to encourage or favor different collaboration styles. *Distant* layouts simulate real-world situations where a large amount of data is scattered over the whole space, forcing users to manipulate data across large distances. It likely benefits from the sharing of knowledge and reaching capability of co-workers. *Local* layouts simulate the opposite situation where work can be done locally. It may discourage collaboration, or encourage different division of labor.

6.3 EXPERIMENT DESIGN

The goal of this experiment is to understand how collaboration affects interaction efficiency and the effects of using a shared interaction technique in loose and close collaboration . We compare five collaboration styles (STYLE = *Divide&Conquer*, *LooseComm*, *CloseComm*, *CloseTech*, *LooseTech*) for two layouts (LAYOUT = *Local*, *Distant*) in terms of the amount of collaboration and coordination, time performance, physical navigation and subjective assessment.

We ask the following research questions:

- What are the efficiency gains and costs at the interaction level, when pairs of users manipulate data on a wall-sized display in loose vs close collaboration?
- What are the effects of providing a shared interaction technique on collaboration behavior, efficiency, physical navigation and subjective assessment?
- Do the effects depend on how data is distributed?

6.3.1 *Participants*

We recruited 10 pairs of volunteers, aged 21 to 40, all with normal or corrected vision. 5 pairs were male-only, 1 was female-only, 4 were mixed. 7 pairs were acquaintances.

6.3.2 *Apparatus*

The same wall-sized display as in the previous experiment was used (Section 5.1.3.2). It is again controlled by a front-end computer running the experiment software, which is implemented using the jBricks Java toolkit [88].

The input device is a motion-tracked pointer. Each pointer controls a cursor on the wall by raycasting. The pointer is mounted on a mobile phone. The orientations and positions of the participants' heads and their pointers are tracked by a VICON motion-capture system. The mobile phone runs an Android application that communicates with the front-end computer via OSC[1] messages. The interface features one pick-drop button and an additional button for the drop-for-partner technique, according to the condition. The buttons are large and provide standard vibration feedback when tapped, so that participants can click them without shifting their visual attention away from the wall.

1 http://opensoundcontrol.org/introduction-osc

Raycasting is used instead of the mobile trackpad from the previous experiment for two reasons. One is that raycasting provides more awareness [45] of the partner's actions. Another is that it needs only one hand to interact, leaving one hand free to perform communicative gestures. Direct touch input is not preferred in this experiment, as direct touch requires close proximity with the display, which makes some areas hard to reach and prevents users from reaching distant items.

6.3.3 *Task Details*

As in the previous experiment, each screen of the wall display serves as a container and can hold up to 6 discs. Each trial starts with a layout containing 16 misclassified discs colored in red while all other discs are in green. Once a disc is moved to a correct container, it becomes green. 8 letters with similar shapes are chosen to label the categories: C, D, H, N, K, R, X, Z.

To ensure equal difficulty, one layout is randomly generated with constraints for each LAYOUT condition, and ten additional layouts are derived from it by symmetric transformation and label permutation.

6.3.4 *Procedure*

The experiment is a [5×2] within-participants design with factors STYLE and LAYOUT. Before starting, participants read a standard explanation of the task and perform an initial four-trial training session, two for each LAYOUT. They are told to get familiar with the task and can practice the drop-for-partner technique. Free form collaboration is encouraged and we recommend participants to find an efficient and comfortable way to communicate and coordinate with each other.

They are specifically asked to establish efficient oral protocols (except, of course, for *Divide&Conquer*) and are given the following examples:

1. "I have [letter]" AFTER picking a disc;
2. "I drop [letter]" when dropping a disc;
3. "I drop [letter] for you" BEFORE dropping a disc for the partner.

After training, participants are instructed to follow the rules of each collaboration style while performing the tasks. To study the effect of locality, participants are told about the differences between *Local* and *Distant* layouts. For all conditions, participants are asked to complete the task as quickly as possible while avoiding dropping discs to the wrong containers.

The experiment is blocked by STYLE, then by LAYOUT. There is one training trial and two replications for each STYLE×LAYOUT condition. The same 20 layouts (5 STYLE × 2 LAYOUT × 2 replications) are used for each pair, distributed so that each layout is used for each STYLE condition the same number of times in the whole experiment. The order of STYLE blocks is counterbalanced across pairs using a Latin Square. The order of LAYOUT is swapped for each STYLE.

The experiment lasts about 70 minutes.

6.3.5 *Data collection*

200 measured trials were collected (5 STYLE × 2 LAYOUT × 2 replications x 10 groups). Task Completion Time (*TCT*) was measured, and misplaced discs were counted as errors. The experiment was recorded on video. The participants' interactions with the display and within pairs, as well as their physical navigation were logged as kinematic data.

Cursor movements on the display, the number of pick-and-drop and drop-for-partner actions were recorded through the experiment software. The participants' physical navigation were tracked by the VICON motion tracking system. Two moderators manually recorded the communication and interaction within pairs to quantify the effectiveness of the collaboration. Each moderator monitored one participant of a pair using a mobile device. The mobile device features three buttons that record the following actions in the kinematic log:

Quantify the effectiveness of collaboration

1. "Asked for Help", when the participant told the letter of his disc to his partner, e.g. "I have D";
2. "Effective Help", when the participant found a correct container for his partner and the disc is dropped into this container, this help can be either through communication, eg. "K is here", or by drop-for-partner when available;
3. "Pick Conflict", when the partners tried to pick the same circle.

At the end of the experiment, participants filled out a questionnaire about their assessment of each collaboration style.

6.4 DATA ANALYSIS

This section explains the data analysis with statistical methods to uncover effects and explain possible reasons. Paired t-tests (by pairs) with Bonferroni correction were used as a main method. In all barplots, error bars show the 95% confidence interval.

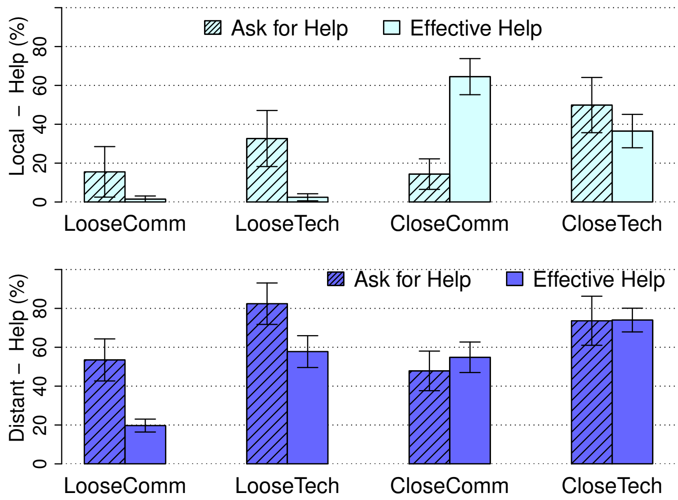


Figure 51: Percentage of pick-and-drops by STYLE and LAYOUT(top: *Local*, bottom: *Distant*) where a participant asked for help (hashed) and effectively helped their partner (solid).

6.4.1 *Effective Collaboration*

Communication within the pairs may lead to beneficial help but also takes time. In order to quantify the collaboration effort and effectiveness, for each trial we counted the frequency of participants asking for help (by spelling out the letter on the disc they picked) as well as the number of pick-and-drops completed with effective help from the partner. Figure 51 shows the proportion of pick-and-drops that involve explicit help by the other partner for each collaborative condition. *Divide&Conquer* is excluded because explicit collaboration is forbidden in this condition.

extensive use of drop-for-partner

Although participants were free to use verbal and gestural communication or drop-for-partner to collaborate, they used drop-for-partner whenever it is provided. 5% effective help in *CloseTech-Local* was via verbal instruction, in the other three conditions with drop-for-partner, all the effective help was through this technique.

6.4.1.1 *Loose Collaboration*

For loose collaboration, we observe that participants help each other significantly more often for *Distant* layouts and when drop-for-partner is available (*LooseTech* and *CloseTech*).

With *Distant* layouts, participants helped each other when working in parallel: in 20% of the cases for *LooseComm*, and in 60% of the cases for *LooseTech* (significantly more than for *LooseComm*, $p < 0.001$). The effective help increased to a large extent due to the use of drop-for-partner. This means that the drop-for-partner technique boosted collaboration in a loose collaboration context.

drop-for-partner boosts collaboration in loose collaboration

Confirming the above effect, participants did not help each other for *Local* layouts while working in parallel (*LooseComm* and *LooseTech*). The moderators noted that the participants performed the task similarly to *Divide&Conquer* in these two conditions, except for spelling out the disc they picked form time to time.

6.4.1.2 *Close collaboration*

For close collaboration, participants helped each other in both layout conditions. Interestingly, the results also exhibit different division of labor.

For *Distant* layouts, about 75% of the pick-and-drops were performed with effective help in *CloseTech*, showing indeed a very effective collaboration. Without shared interaction support (*CloseComm*), this number dropped significantly to 57% ($p < 0.001$).

drop-for-partner increased collaboration with distant layouts but decreased it with local layouts

For the *Local* layouts the situation was reverted: there was significantly ($p < 0.001$) less effective help with drop-for-partner (35% for *CloseTech*) than without (63% for *CloseComm*). In fact for *CloseComm*, the help is provided spontaneously, with few instances of "Ask for Help". The moderators observed that the participant without the pointer was very close to her partner and often became very proactive and planned ahead the next pick-and-drops for the partner (e.g., "put H there"), see Figure 50-4. This created an effective division of labor, where one partner focused on manipulation with the device and the other on planning next actions. Whereas when both of them had a device but only one could help via drop-for-partner (*CloseTech*), the "helper", who was not allowed to pick, waited for her partner to tell the label and then started to search.

Effective division of labor

Overall, these results demonstrate that both the spatial distribution of information (*Distant* vs. *Local* layout) and the availability of shared interaction support (drop-for-partner) had a large influence on collaborative behaviors.

6.4.2 *Errors & Conflicts*

On average, participants made 0.88 erroneous pick-and-drop per trial, from 0.5 for *CloseTech-Distant* to 1.45 for *LooseComm-Distant*, but the differences are not significant.

Only 6 pick conflicts are logged overall. Apparently in such a colocated environment, the pairs are generally aware of the partner's position or action and avoid conflicts. However, 5 out of the 6 conflicts occurred in the same condition (with different groups): *Divide&Conquer-Distant*. This suggests that performing the task independently may lead to slightly more conflicts.

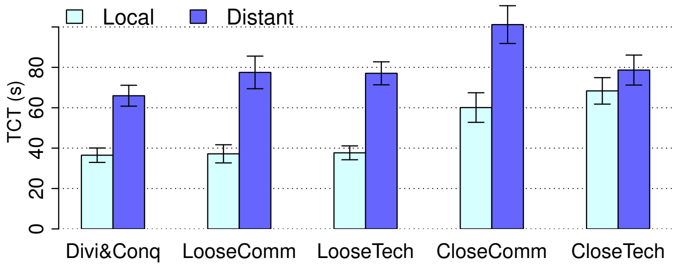


Figure 52: Mean Task Completion Time by STYLE × LAYOUT conditions.

6.4.3 *Performance Cost and Gain*

No data point is removed as outlier, since all trials are within 30% of the mean completion time per STYLE × LAYOUT condition and participant group. Shapiro-Wilky normality tests for the means per groups for each condition show no evidence of non-normality, except for the *LooseComm-Distant* condition where one group is clearly slower than the others. However, this does not affect the analyses below.

6.4.3.1 *Task Completion Time*

An ANOVA in the model *TCT* ~ STYLE × LAYOUT × Rand(PARTGROUP) shows significant effects of STYLE ($F_{4,36} = 33.0$, $p < 0.0001$, $\eta^2_G = 0.46$) and LAYOUT ($F_{1,9} = 244$, $p < 0.0001$, $\eta^2_G = 0.64$), and a significant STYLE × LAYOUT interaction ($F_{4,36} = 13.2$, $p < 0.0001$, $\eta^2_G = 0.19$). The data exhibits no sphericity and the effect sizes are medium to large.

Figure 52 shows the Task Completion Time (*TCT*) for each condition. Obviously, the pairs are significantly faster when performing *Local* layouts than *Distant* layouts (48.0 ± 7.1 s vs 80.1 ± 7.8 s). Several T-tests with Bonferroni correction for comparing *Local* and *Distant* layouts by STYLE show that *Local* and *Distant* are significantly different for each STYLE (all $p < 0.001$), except for *CloseTech* ($p = 0.0983$).

For *Local* layouts, *TCT* for Divide&Conquer and the loose collaboration styles are significantly faster than the close collaboration styles ($p < 0.001$). Apparently the sequential execution in close collaboration hinders overall performance compared to parallel execution.

sequential execution is time-consuming

There is no significant difference between *Divide&Conquer*, *LooseComm* and *LooseTech* for *Local* layouts ($p = 1.0$). This is expected since in these cases the pairs performed the task similarly using a divide & conquer strategy. Moreover, there is also no significant difference between *CloseComm* and *CloseTech* ($p = 0.15$). Therefore we can see that drop-for-partner does not improve the efficiency when the data manipulation is local. In fact in close collaboration, verbal and gestural help (in *CloseComm*) seems to be more effective, due a better division of labor, as described before.

no significant difference on TCT with local layouts

For *Distant* layouts, *Divide&Conquer* is marginally faster than *LooseTech* ($p = 0.072$) and significantly faster than all the other styles (p's < 0.031).

Divide&Conquer was the most efficient

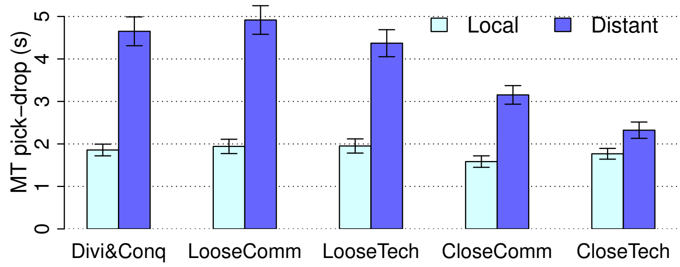


Figure 53: Mean time to perform one pick-and-drop by STYLE × LAYOUT.

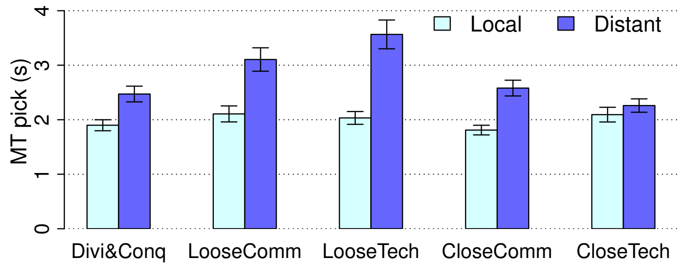


Figure 54: Mean time of the gap between a drop (or trial start) and the next pick by STYLE × LAYOUT.

The explicit collaboration did not improve overall task performance. However, drop-for-partner improves performance in close collaboration (*CloseTech* < *CloseComm*, $p < 0.001$). In this case close collaboration (sequential pick-and-drops) reaches similar performance to loose collaboration (parallel pick-and-drops), suggesting there is a benefit. Nonetheless, this benefit is not visible in loose collaboration (*LooseComm* and *LooseTech* very close, $p = 1.0$). This indicates a possible cost that outweighs the benefit of the collaboration.

In addition, although *CloseComm* is slower than the other styles for *Distant* layouts (p's < 0.001), it is 1.3 times faster than twice the *TCT* of *Divide&Conquer* which, according to my pilot studies, is a close approximation of the time for a single user to perform the task. Thus, help from the partner did improve performance.

possible cost in collaboration

While we notice different collaborative behaviors across the collaborative styles (see Section 6.4.1), especially for *Distant* layouts, overall performance is very similar for some of them. Overall the above analysis indicates that collaboration incurs both gain and cost in terms of time performance for interaction. To better understand the underlying trade-offs, we dive deeper into the data by analyzing individual pick-and-drops to find the possible sources of the cost and gain.

6.4.3.2 *Individual Pick-and-Drops*

To take a closer look at the interaction, we analyzed the movement time (MT_{PD}) of 3200 pick-and-drops (160 per condition after removing the erroneous ones). Figure 53 shows the average MT_{PD} per STYLE × LAYOUT.

For *Distant* layouts, both close collaboration styles are significantly faster than the loose collaboration styles (p's < 0.006). This shows that close collaboration increases the efficiency of individual pick-and-drops. In particular, individual pick-and-drops are twice as fast with *CloseTech* than with *Divide&Conquer*. But this advantage is offset by the fact that the 16 pick-and-drops have to be performed sequentially vs. each participant having to perform 8 pick-and-drops on average in parallel. This explains the similar *TCT* between *CloseTech* and loose collaboration styles as well as Divide&Conquer.

collaboration leads to faster pick-and-drops

On the other hand, there is no significant difference between *Divide&Conquer*, *LooseComm* and *LooseTech* ($p = 1.0$), even though the mean MT_{PD} for *LooseTech* is 7% lower than for *Divide&Conquer*. This suggests that the benefit gained by the help from the partner is offset by the cost of collaboration, whether or not if drop-for-partner is available. This might be due to multitasking: the overhead of performing one's own task while receiving the partner's request or searching for the partner's container.

gained efficiency might be offset by the cost of multi-tasking

No significant difference is found between STYLE with *Local* layouts, which again confirms the effect of the spatial distribution of data on pairs' behavior. This is consistent with the moderators' observation that pairs basically used *Divide&Conquer* in loose collaboration. Even though they tried to help in close collaboration, there is little improvement in MT_{PD} with *Local* layouts.

6.4.3.3 *Multitasking Cost*

In order to further understand the source of collaboration cost, we analyzed the time spent between pick-and-drop actions, i.e., after a drop and before the next pick. We noticed that the participants normally picked a nearby red disc immediately after dropping the previous one, if they were not distracted or interrupted by their partner. This measure thus reflects, to some extent the level of distraction or multitasking cost. As we can see in Figure 54, this measure is almost constant for *Local* layouts.

For *Distant* layouts, there are large significant differences between loose collaboration styles (*LooseComm* and *LooseTech*) and the others (p's < 0.001 for *LooseTech* and p's < 0.03 for *LooseComm*). This shows that helping each other while working in parallel has a cost, especially for *LooseTech* where participants helped each other a lot (Figure 51). This seems to cancel the small advantage in pick-and-drop for *LooseTech* (Figure 53). This is consistent with the moderators' observation that a participant would sometimes get interrupted and stop her own task to help her partner. Participants' subjective assessment also confirms this finding (see Section 6.4.5).

the measure of time spent between pick-and-drops revealed the cost

To confirm this finding, we measured the operational parallelization of the pick-and-drops in Divide&Conquer and the loose collaboration styles. The percentage of pick-and-drop overlap is calculated

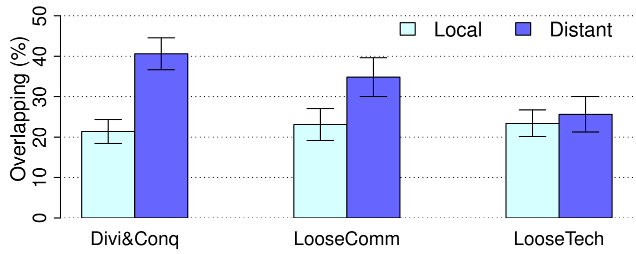


Figure 55: Percentage of time during a trial when both participants have a disc on their cursor, by STYLE × LAYOUT.

as the percentage of time during a trial when both participants have a disc picked.

As shown in Figure 55, for *Distant* layouts, this percentage is significantly lower for *LooseTech* (26% of *TCT*) than for *Divide&Conquer* ($p < 0.001$, 41% of *TCT*) and *LooseComm* ($p = 0.008$, 35% of *TCT*). Referring back to the amount of collaboration (Figure 51), this shows that more collaboration leads to less parallelization.

6.4.4 *Physical Navigation*

To understand the participants' physical navigation, we first analyze the distance traveled by each pair for each condition (Figure 56).

With *Distant* layouts, it is significantly shorter for *LooseTech* than for *Divide&Conquer* ($p = 0.0013$, 26% shorter), *LooseComm* ($p = 0.0003$, 33% shorter) and *CloseComm* ($p < 0.001$, 52% shorter). The same trend holds for *CloseTech* versus *Divide&Conquer* ($p = 0.1$, 17% shorter but not significant) and versus *LooseComm* ($p = 0.0076$, 26% shorter) and *CloseComm* ($p < 0.001$, 45% shorter). This shows that the drop-for-partner feature can reduce travel distances with scattered data.

drop-for-partner reduced travel distances

In addition, *CloseComm* leads to much longer distances than all other STYLES for *Distant* layouts ($p < 0.001$). For *Local* layouts, the distances for sequential styles are twice as long as for parallel styles ($p < 0.001$).

To understand the coupling patterns for each condition, we compute *Separation*, the mean distance between the partners at the time when discs are dropped (Figure 57), and plot the movements of the participants during a trial in each condition (see Figure 58 for typical examples). By analyzing the participants' trajectories together with the amount of collaboration between partners (Figure 51), we find evidence of different territoriality and trajectory patterns according to the styles and layouts.

6.4.4.1 *Side-by-side vs. At-a-distance*

In Figure 58 we recognize a trajectory pattern where one partner follows the other in most close collaboration conditions (both *CloseComm*

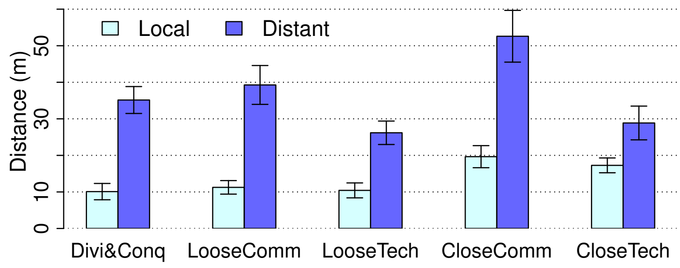


Figure 56: Mean total distance traveled by each pair, by STYLE × LAYOUT.

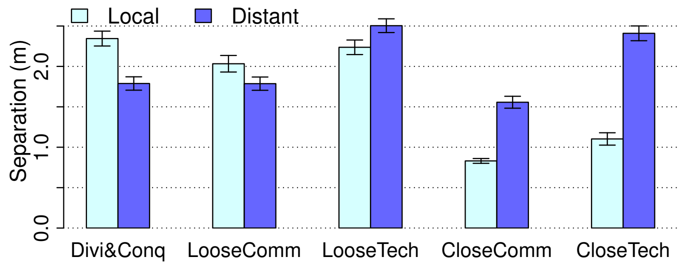


Figure 57: Mean distance between the two partners at drop time, by STYLE × LAYOUT.

and *CloseTech-Local*). In these conditions, one participant typically follows the other and the task is solved together. This corresponds to high levels of effective collaboration (Figure 51).

Interestingly, *CloseTech-Distant* does not exhibit this "following" pattern, nonetheless reaches a high rate of effective collaboration (74%). This means that the shared interaction technique enables tight collaboration at a distance when the data to be manipulated is scattered (Figure 58 bottom-right).

However with *CloseTech-Local*, the helper was sometimes further away in the hope of dropping a disc at distance. This is not optimal for *Local* layouts as the proper container is always nearby, thus leads to less effective collaboration and higher *TCT*. Therefore for non-scattered data, verbal and communication help in close collaboration is more effective than using drop-for-partner.

6.4.4.2 *Territoriality*

We can see very similar trajectory patterns for *Divide&Conquer* and loose collaboration styles for *Local* layouts (Figure 58 left top 3). Participants subconsciously split the wall display into two sides and perform the tasks independently. This is consistent with the very small amount of help in *LooseComm* and *LooseTech* for *Local* layouts (Figure 51). The mean *Separation* is 2.2m in these conditions – more than one third of the display width (Figure 57).

With *Divide&Conquer* and *Distant* layouts, both trajectories cross the entire space (Figure 58 top-right), showing that participants work independently of each other. No territoriality appears in this condition.

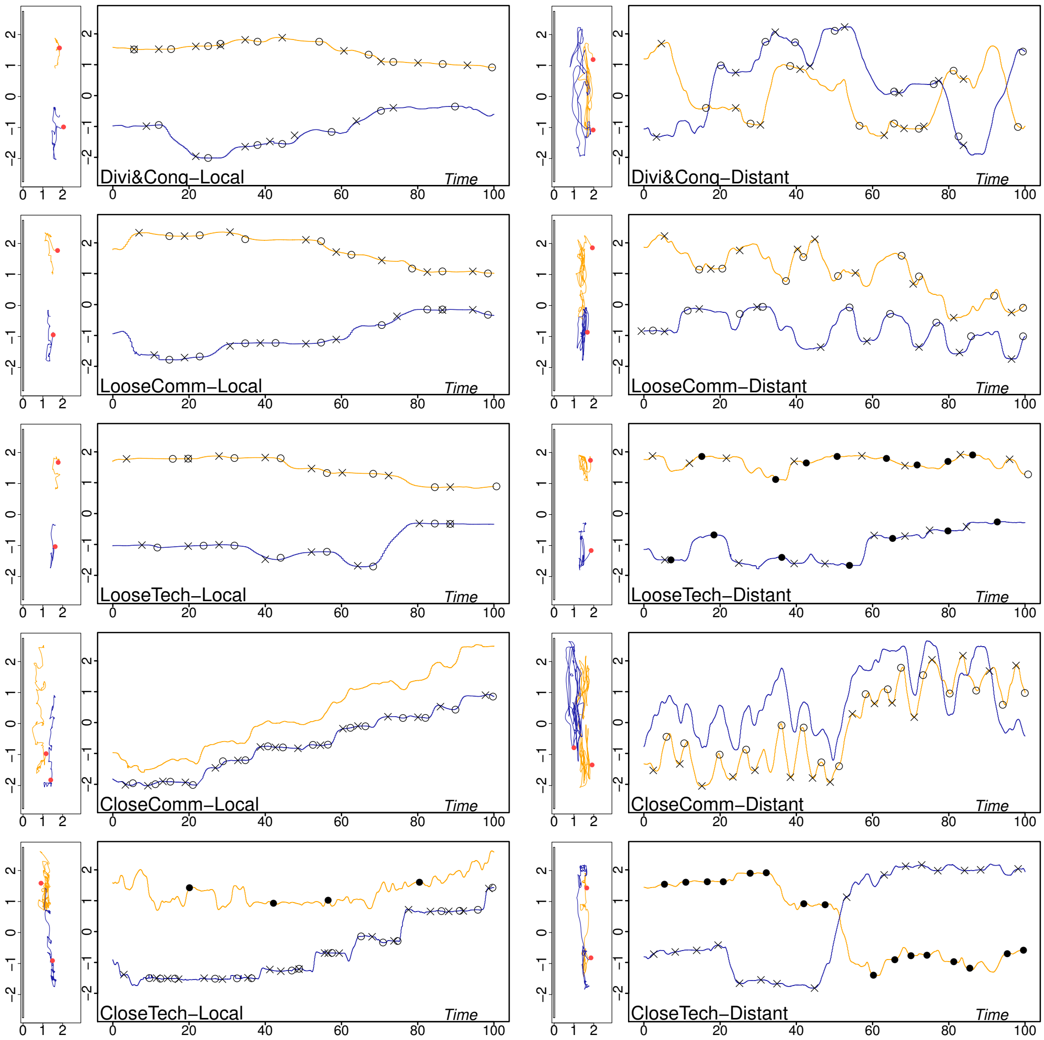


Figure 58: Ten example graphs of pairs' navigation paths, one for each STYLE × LAYOUT condition. Each graph shows: (i) on the left, the movement of both participants in a bird's eye view of the wall room with the wall on the left (unit is meter); and (ii) on the right, the same paths stretched over a normalized timeline (x-axis) to help understand the pairs' navigation patterns. In addition, picks (×), drops (∘) and drop-for-partner (•) actions are plotted on the paths.

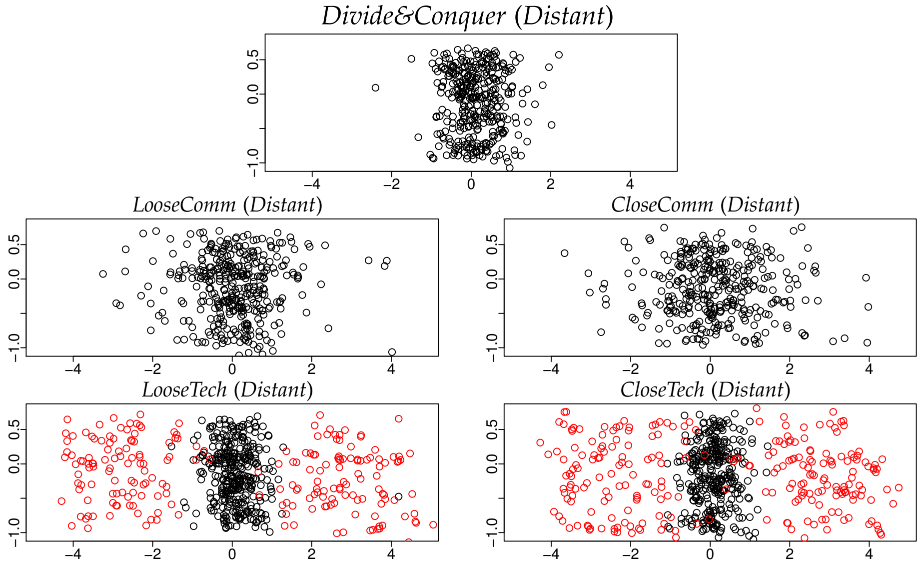


Figure 59: Positions (in meter) of the drops by the picker (∘) and by the partner through drop-for-partner (∘), relative to the head of the picker (projected at the graph center(0,0)).

With *LooseComm-Distant*, pairs help each other via communication while performing parallel pick-and-drops. They tend to work on different areas of the wall-sized display, showing certain territorial behaviors.

Moreover, *LooseTech* and *CloseTech* also lead to user territories with *Distant* layouts because the partners could share the search at a distance and use drop-for-partner to help each other (see the high level of effective help in Figure 51). The mean distance between pairs is 2.5 meter, significantly larger than all other styles (p's < 0.0056). In *CloseTech* collaboration, participants often swapped the part of the wall they were working on so that the "picker" was closer to the misclassified discs (Figure 58 bottom-right).

6.4.4.3 *Combination of Power*

The provided shared interaction technique – drop-for-partner – facilitates sharing of both mental and physical resources of the pairs over a distance. It augments the ability of single users by supporting new forms of collaboration.

Interesting evidence can be seen in Figure 59, which plots all dropping-for-self positions (in black) and dropping via drop-for-partner positions (in red) relative to the head position, which is translated to the graph center, of the person who picked the discs. Only *Distant* layout conditions are shown here. The reach range of a participant's manip-

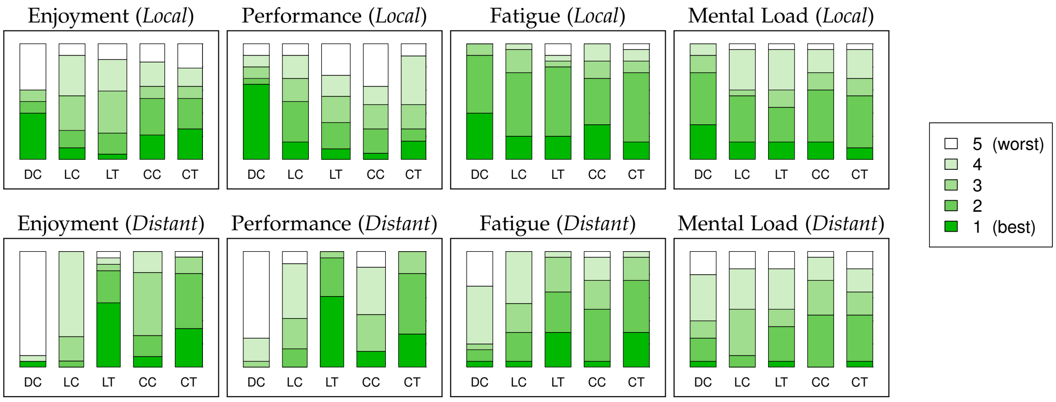


Figure 60: Participant ratings by STYLE for *Distant* layouts for Enjoyment and Perceived Performance (ranked from 1st to 5th) and for Fatigue and Mental Load on a five-point Likert item (1 is best, 5 is worst). *DC* denotes *Divide&Conquer*, *LC* denotes *LooseComm*..., *CT* denotes *CloseTech*.

ulation is the smallest with *Divide&Conquer*. This is because without help from their partner, participants can only read the labels in a small area in front of them. When communicating with their partner, some drop positions are further away from the participants (*LooseComm* and *CloseComm*). This is because they can drop into a container pointed by their partner without reading the labels in it. Furthermore, the participants' reach range extends to the entire display with the support of drop-for-partner (*LooseTech* and *CloseTech*), as their discs can be directly dropped by their partner.

This demonstrates that a shared interaction technique empowers the users for interacting in a large space, by facilitating collaborators to share their knowledge and physical capability. It encourages close collaboration and enables new forms of division of labor.

6.4.5 *Subjective Assessment*

Figure 60 illustrates the ratios of rankings for perceived efficiency and enjoyment by the 20 participants, as well as their rating of physical and mental fatigue. The analysis uses Wilcoxon rank sum tests with Bonferroni correction.

For *Local* layouts, *Divide&Conquer* was perceived to be faster than *LooseTech* ($p = 0.0234$) and *CloseComm* ($p = 0.0092$), and marginally faster than *CloseTech* ($p = 0.0801$). Other comparisons do not show significant differences.

The surprising result lies in *Distant* layouts: *Divide&Conquer* was perceived as the slowest (p's < 0.005) while it is actually the fastest in Task Completion Time (Figure 52). 15 participants ranked *Divide&Conquer* as the least efficient and 4 as the second least efficient. *CloseComm* is ranked higher than *Divide&Conquer* while it is in fact the slowest. The ranking of both perceived efficiency and enjoyment for *LooseTech* and *CloseTech* are significantly higher than for the other styles (p's < 0.01).

It seems that perceived efficiency was largely influenced by how effective the collaboration was. "I was the most satisfied by the conditions where I felt we were working efficiently as a team and making the most progress quickly", "I would say that the ones I enjoyed the most were the one where we collaborated the most".

Enjoyment for *CloseComm* is ranked significantly higher than *Divide&Conquer and LooseComm* (p's < 0.01). One participant said, "*CloseComm* requires less synchronization and we can anticipate actions of our partner by focusing on his behavior (pointer and spatial location)".

In terms of physical fatigue, *Divide&Conquer* was the most tiring style ($p = 0.0002$ compared to *LooseTech* and *CloseTech*, $p = 0.0134$ compared to *CloseComm* and $p = 0.0455$ compared to *LooseComm*). drop-for-partner reduces physical fatigue, as *LooseTech* was less tiring than *LooseComm* ($p = 0.019$) and *CloseTech* was less tiring than *CloseComm* ($p = 0.09$). This is consistent with the measured travel distances (Figure 56), which showed reduced physical navigation when providing the drop-for-partner feature.

18 out of 20 participants confirmed that they were multi-tasking in *LooseComm* and *LooseTech* conditions: they were searching for the containers for themselves and their partner concurrently. 6 of them thought it did not work well.

There is no significant difference between STYLE for the ratings of mental load in either layout (p's > 0.2). Obviously *Distant* layouts incur a significantly higher mental load than *Local* layouts ($p < 0.001$). Indeed, 17 participants tried to remember the *Distant* layouts with spatial memory, and 2 gave up due to the heavy mental demand. In additional comments, 11 participants mentioned that multi-tasking is mentally demanding, while a few others complained about the effort of verbal coordination and memorization of the layouts.

6.5 DISCUSSION

The above data analysis reveals a number of interesting phenomena. First, the provided shared interaction technique - drop-for-partner encourages collaboration . Second, the shared interaction technique reduces physical navigation, improves operation efficiency and provides a more enjoyable experience. Third, *Divide&Conquer* is faster than the other styles, but is perceived as the least efficient, and is preferred the least by the participants. Finally, the physical navigation of each pair exhibits distinctive patterns according to the collaboration styles.

6.5.1 *Implication*

The following implications can be drawn based on the findings from data analysis. They provide valuable input to improve interaction design for collaboration.

6.5.1.1 *Collaboration and Interaction Efficiency*

My interviews and observation of real users who performed co-located collaborative tasks show that their main motivation for co-located collaboration was to combine multiple users' knowledge and spark ideas through discussions. Hence, although our study suggests that communication may hinder interaction efficiency, this does not mean that communicating is not beneficial to the users. For instance in this experiment, more collaboration was perceived as more efficient and more enjoyable, probably leading to a higher level of engagement. Since collaborative tasks can have different evaluation criteria, it is the designers' choice to encourage or discourage different forms of collaboration depending on the needs of each specific task.

The point of comparing the interaction efficiency of different collaboration styles is not to tell users which collaboration strategy is "better". Here we provide a deeper understanding by identifying the gains and costs associated to the collaboration situations. Understanding the effects of the provided interaction and instruction on collaboration as well as users' behavior pattern helps design effective techniques for various goals. Moreover, appropriate interaction support facilitates users working with an ultra-large surface in tasks requiring significant amounts of time and effort, allowing them to allocate more mental resources to the intellectual part of their tasks.

We show evidence that explicit collaboration in loosely collaborative conditions adds a cognitive cost because of multi-tasking and disruption by partners. One implication for design is to provide proper interaction techniques to take advantage of collaboration while minimizing the cost of disruption.

6.5.1.2 *Shared Interaction*

Providing a shared interaction technique shows several benefits in supporting collaboration. First, close collaboration can be encouraged by such a technique. In this experiment the participants collaborated much more when the drop-for-partner technique was available.

Shared interaction techniques can also improve efficiency: drop-for-partner improved the efficiency of individual pick-and-drops and reduced travel distances and physical fatigue. This is especially beneficial for situations where close collaboration is needed. Existing collaboration techniques are rarely designed for this purpose, or used in this context. This opens up a new design space to explore techniques to support collaboration.

Drop-for-partner led to different trajectory patterns and territoriality of collaborators. The existing literature on large displays often states that users collaborate more when they are physically close. In this case the provided shared interaction technique enables tightly-coupled collaboration at a distance, thus allowing different forms of division of labor.

In fact, I see drop-for-partner as one example of embodied interaction technique in this context. It takes advantage of users' verbal communication, blends in with their collaborative behavior while making it easier for users to help each other. Certainly, drop-for-partner is only one specific instance of such shared interaction techniques. It is tailored for pairs performing a classification task with pick and drop interaction. Similar techniques can be developed for other collaborative tasks according to the task characteristics. One take-away inspiration for designers could be to observe verbal or gestural communication between users and design shared interaction techniques that augment them.

6.6 SUMMARY

In this chapter I started with the operationalization of five collaboration styles crossing two dimensions – collaborative coupling and shared interaction support. Then I described an experiment conducted with the abstract data manipulation task I designed in Chapter 4. It features two factors: collaboration style and the spatial distribution of data. The task again allowed us to focus on the interaction level and compare quantitative measures for both results and elemental operations (pick-and-drops). It quantifies the effectiveness of communication and amount of collaboration, and illustrates typical behavior patterns.

This methodology separates different collaborative interaction situations, whereas more ecological approaches such as observational studies must deal with occurrences of different collaboration styles within a single task. While this one experiment necessarily picked a particular task and interactions, the methodology can be applied with other factors and interaction techniques in different contexts.

The results show that providing a shared interaction technique can encourage collaboration and allow co-workers to collaborate tightly even when not in close proximity, which facilitates sharing of knowledge and physical ability. The technique also improves the interaction efficiency, reduces physical travel effort and improves subjective experience. Furthermore, we showed evidence of the cost of coordination and multitasking incurred by collaboration while co-workers work in parallel.

Last but not least, the mismatch between the actual and perceived efficiency from participants is an interesting finding. *Divide&Conquer* is faster than the other styles, but is perceived as the least efficient, and is preferred the least by the participants. This raises questions about the metrics to be used to assess collaborative activities. Should engagement and enjoyment be used instead of task completion time? Or can we reconcile the social qualities of collaborative interaction with raw efficiency?

"The learnability of literalism makes it a good thing. However, the designer can always provide the user with enhanced capabilities at the price of breaking out of the metaphor. These features might allow the user to do wonderful things that are far beyond the capabilities of literal features. ... There is a tradeoff between the learnability of literalism and the power of magic."

Randall B. Smith – The alternate reality kit: an example of the tension between literalism and magic (1987) [105].

7 SHARED INTERACTION TECHNIQUES FOR COLLABORATIVE MANIPULATION

Based on the findings of the previous study, this chapter explores shared interaction techniques that support collaborative data manipulation and data exchange between co-workers. I introduce the design, implementation and informal evaluation of a novel interaction technique - Collaborative Gestures. It supports collaboration by reducing the effort of manipulation and facilitating data exchange between users. It encourages close collaboration while allowing smooth transfer to loose collaboration.

Wall-sized displays provide an interactive space that enables co-located collaboration but also presents some of the problems of remote collaboration: moving content across the surface may be tiring, users may or may not see each other's actions depending on their focus of attention and on how far they are from each other. The experiment described in the previous chapter explored interaction of pairs manipulating data on a wall-sized display, and revealed various advantages of providing a shared interaction technique. This chapter explores the design space of such techniques.

To solve problems caused by distance, existing research has focused on solutions allowing single users to reach remote areas of large displays or on remote communication methods using icons or avatars for representing users and allowing them to exchange notifications. The former approach is not optimized for collaboration, and the latter takes little advantage of the co-located situation. While shared interaction techniques have been explored in the literature, for instance for multi-surface interaction, this concept has not been explored in the context of wall-sized displays. Before introducing my work on this topic, I first give an overview of shared interaction techniques and techniques for solving distance problems on large displays.

7.1 BACKGROUND

7.1.1 *Large Scale Interaction*

Drag-and-drop is a standard direct manipulation technique for moving data. It allows to pick an object using the cursor, then to move it along with the cursor to the desired location. Hascoët [47] introduced Drag-and-Throw and Push-and-Throw to improve conventional Drag-and-Drop. This technique allows throwing objects to a remote position of a large display. It aims at providing good user control and low error rates. Pick-and-drop [96] extends Drag-and-drop by avoiding the continuous contact of input while moving an object. It is a classic technique for moving content over a large distance or across surfaces.

Collomb et al. [25] evaluated several remote techniques to extend drag-and-drop for reaching distant targets on a large display (Figure 61-3). The proposed approaches include generating temporary proxies of the remote targets, or creating a virtual miniature of the whole display around the hand position after picking up an icon. They also introduced a technique that relies on a rubber band metaphor. Using this technique, the gain of the dragged distance of a target is controlled by bending the band in the opposite direction of the destination.

Various techniques such as remote control widgets or view portals were also proposed to make it easier to move or share data on large displays. Remote control widgets have been presented for reaching remote targets or transferring data to remote areas. For instance, Frisbee [63] consists of a local "telescope" and a remote "target" (Figure 61-1). Proxies of remote data around the "target" area are shown in the "telescope". Users can interact with the proxies instead of walking over to the target area. The evaluation showed that Frisbee is preferred over walking back and forth when the target is at a distance of more than 4.5 feets. The Vaccum [11] (Figure 61-2) is a circular graphical widget that can be activated to provide proxies of remote targets over a fan-shaped area near the cursor position. The target area can be adjusted by crossing over regions of the widget arc with pen strokes.

Remote control widgets

Distant pointing

TractorBeam [86] combines pen-based interaction and raycasting to provide a seamless experience for reaching distant targets on tabletop surfaces. Users point to distant positions with the same pen they use to touch the surface and can select using an additional button on the

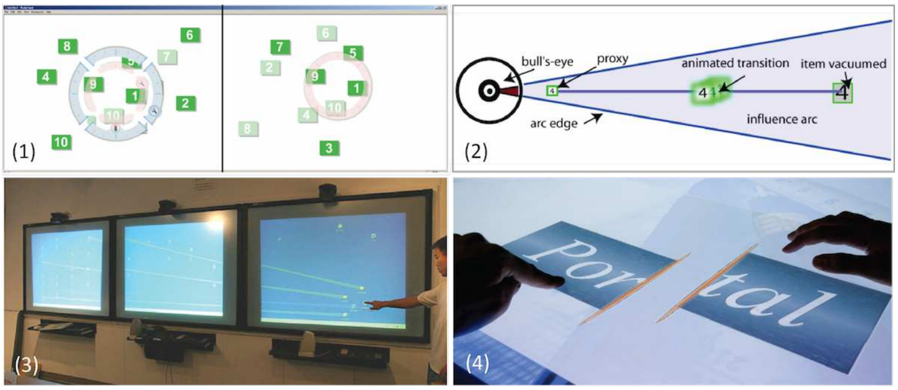


Figure 61: Examples of techniques for remote interaction on large displays. (1) Frisbee: the left widget is the "Telescope" and the right widget is the "target" [63]; (2) The Vaccum: the proxy of target 4 is generated close to the circular widget [11]; (3) Drag-and-Pop: proxies of remote targets are generated close to the moving object [25]; and (4) Dynamic Portals: portals are created by drawing lines on the display, and the content dragged into one portal comes out from the others [112].

pen. Similarly, in the context of wall displays, Nancel et al. [83] propose several dual-precision pointing techniques to reach small targets at a distance. Shadow Reaching [104] uses the shadow of the body of the user, as if it was located in front of a light source, to facilitate distant interaction.

Passing data among multiple users

There are also techniques helping multiple users to exchange data between each other. Dynamic Portals [112] allow users to create a teletransport "portal" just by drawing a line on the display (Figure 61-4). The transported content can be resized by changing the sizes of the portals and content can be transported to multiple portals. Pass-Them-Around [68] allows users to share photos across multiple mobile devices by "throwing" the photo towards another mobile phone's direction. Tilting the phone triggers different sharing interactions. Sparsh [75] presents the concept of using the human body as medium to pass data, by touching one device after another.

7.1.2 *Multiuser Cooperative Actions*

Multiple users are able to coordinate themselves to complete a task together. The scale ranges from carrying a large object to planning a complex task together. Due to the scope of the thesis, we focus on multi-user interaction in co-located environments, where users coordinate their actions to perform elemental tasks together.

Cooperative Gestures [77] explore gestures performed by multiple users cooperatively with a tabletop sketch application. As shown in Figure 62, two or more users interact simultaneously to complete a command. For example, a) is a gesture to establish partnership; b)

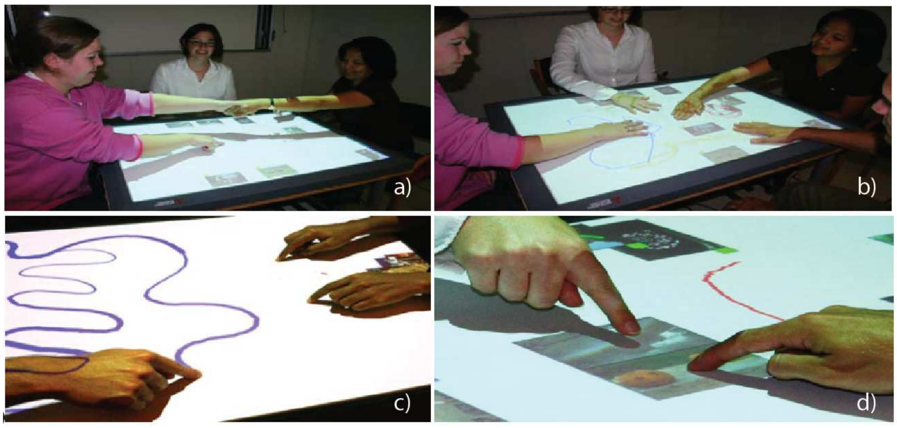


Figure 62: Cooperative Gestures. Image reproduced from [77]

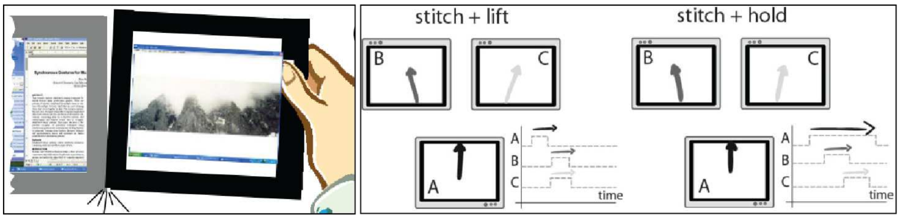


Figure 63: Cooperative stitching: left shows two tablets bump together to establish a connection, right shows Cooperative Stitching gesture. Image reproduced from [93]

shows that when all users are erasing strokes the display clears the entire board; c) shows one person drawing while another person controls the thickness of the stroke; d) shows automatic collage when two users push their pictures together.

Synchronous Gestures [93] provide a set of techniques requiring synchronous actions of multiple input resources (Figure 63). They can be used to establish connections between multiple surfaces and to transit information smoothly across them. Cooperative Stitching gestures allow one user to share content with multiple receivers by taking into account the timing overlap of the stitch actions.

HapticTurk [21] provides an interactive experience to a Virtual Reality game player by replacing motion platforms with multiple humans who manually carry the player together. They lift, tilt and push the player's limbs or torso, following the motion instructions displayed on a mobile device in front of each of them. HapticTurk explores the feasibility of providing interactive experiences through the guided cooperation of several "actuators", i.e., the carriers.

Using shared interaction techniques to support data exchange in different collaborative situations is an approach that has been rarely explored. This chapter describes the design and implementation of two sets of shared interaction techniques, focusing on close and loose collaboration respectively. An informal evaluation provides insights

Figure 64: Examples of cards used in the study. The cards are printed from the pictures of the Dixit game[1].

about the techniques and will help improving the design in future iterations.

7.2 OBSERVATIONAL STUDY

The collaboration experiment described in the previous chapter used the drop-for-partner technique to support shared interaction between pairs of users. This technique was tailored for pick-and-drop interaction in the experimental task. In order to gain inspiration for exploring the design space of shared interaction techniques, I conducted an informal study to observe people performing a task involving arranging paper cards.

7.2.1 *Design*

The task was inspired by a real task described by the sociologists I interviewed (Chapter 3). It was designed to operationalize a similar situation on interaction level, to the situation when the sociologists performed their real tasks.

Paper picture card arrangement task

Pairs of participants were asked to rearrange randomly placed paper picture cards[1] on a wall. Each paper card showed an artistic painting (Figure 64) which could be interpreted in several ways. For example, the picture in the middle of Figure 64 can mean knowledge, freedom, peace, cozy, magic and so on. Participants were asked to find similarities or connections between the pictures and to arrange them in any meaningful way they would agree on (e.g., by grouping and storytelling). They were encouraged to discuss in order to find out creative ideas and to collaborate in performing the task. The task was finished when they had arranged all pictures and agreed on the arrangement. At the end, they were asked to explain why they chose this layout to the moderator.

Two setups with different surface sizes were used with different groups in the study. In one setup, the cards were placed on a 3.3 × 1.6m wall with three white boards mounted side by side. Each card was at-

1 http://en.libellud.com/games/dixit

Figure 65: Layouts created by the four pairs of participants.

tached to the wall with a magnet. The card could be moved easily by pushing the card or the magnet. In the other setup, the cards were stuck to a 5.5 × 1.8m wall-sized display. Each card was attached to the display with patafix gluepads. In order to move a card, the participants had to remove it and then attach it at another position.

From small to large scale manipulation

Four pairs of participants were recruited for the study. Each pair was two colleagues who knew each other well. One pair performed the task with small size cards (9.5 × 7cm) on the smaller wall, two pairs used bigger cards (19 × 14 cm) on the smaller wall, and one pair used the bigger cards on the larger wall. The three situations correspond to small, medium and large scales of manipulation. I was interested to see what interaction problems might emerge when the manipulation scale increases, and how users would cope with them.

7.2.2 *Findings*

7.2.2.1 *Encode Meaning in Space*

The four pairs of participants reported on the layouts they created with the cards, as shown in Figure 65. Various meanings were encoded using the space through, e.g., relative positions.

Grouping

Group 1 chose to group the images using the keywords they came up with and agreed on. Group 2 had a similar grouping approach, except that they made a “bigger image” with the dominant colors of groups: blue pictures were on the top as they could be associated to the sky; green and brown colors went to the bottom as they were associated to the ground, etc.

Group 3 played a game of making a chain by linking the cards with common keywords. For example one participant would put a card and say a keyword, for instance “alone”, then both of them would be seeking for the next card related to either “alone” or a new keyword that could be interpreted from the previous card. Interestingly, they completed the task in a positive competitive approach, as every card became a quiz challenging the search of the next one.

Competitive cooperation

Group 4 created a story that branched from three parallel scenes. Each scene was illustrated with one line of cards from right to left. Two scenes described the two main character's lives in parallel and the third scene described the evolving background of the story. Eventually the three branches combined together when the two characters met in the story, which continued till the end at the left edge of the surface.

Story telling

7.2.2.2 *Combination of Resources*

Pairs were able to combine their knowledge and ideas in performing a creative task. By manipulating the positions of the cards, they encoded rich information in the space and used it to communicate ideas. I observed various ways of sharing knowledge and physical resources between partners.

DEICTIC ACTIONS All pairs frequently showed each other a card when it was relevant for the ongoing conversation or the ongoing action of the partner. A common reference appeared to be very important for communication and collaboration in this situation. Several types of deictic actions were observed:

1. Move the relevant card into the focus area of the partner (Figure 66-b1). This happened especially often in the smaller scale situation (smaller cards, smaller wall), as a participant could easily reach the partner's focus area.
2. Point to a card (Figure 66-a1, a2). This happened often when pairs were engaged in a conversation, perhaps because in these situations, they were sure that their movements would be noticed by their partner.
3. Take a card in hand and show it towards the partner (Figure 66-b2). This happened frequently in the large scale manipulation with a larger wall. It is difficult for participants to see a remote area on the wall. In this situation, the advantage of physical paper is well taken: a card can be detached from the wall and easily carried around. This is not as easy to achieve with digital items on a display.
4. Use deictic words. Pairs also talked about cards by naming a property or a keyword that would help to identify it. This usually happened for the cards that appeared in the conversation before, as partners would then be more likely to recall and identify them.

Figure 66: Observed behaviors from pairs of participants. a: deictic actions; b: showing cards to partner; c: passing cards; d: overview and close manipulation.

Deictic actions are powerful resources for embodied interaction. When one person points to a position or an object while talking to another person, the other person automatically knows where to look. When a target is involved in a conversation, people can quickly reach a common understanding about it. When a person shows another person an object from the other side of the room, the other person automatically starts to walk towards her, which commonly leads to both of them walking towards each other. People coordinate with each other subconsciously. Therefore, interaction techniques leveraging these resources could be beneficial, for instance a relatively low cognitive load.

PASSING OBJECTS The pairs passed objects to each other from time to time (Figure 66-c1, c2). Typically this happened when one participant heard that the partner was searching for one kind of cards and happened to have a matching card in hand. In close collaboration, the pairs were able to fluidly coordinate their actions to optimize the resources. For instance, when they were in a discussion and agreed to put a card at a given position, the person that was closer to the position would get the card and put it on the surface, if she was not engaged in another action. Here we see embodiment as well: when a person hands in an object to another, the other person automatically starts a gesture towards the object to accept it.

FAR + CLOSE COLLABORATION It also happened sometimes that one participant stepped back to have an overview of the arrangement and give suggestions to her partner, who was manipulating cards on the wall (Figure 66-d1, d2). When the far participant had a different idea, she would go back to the wall to move the cards while explaining her idea, instead of expressing it verbally from the far position.

MIXED COLLABORATION STYLES Consistent with findings in previous observational studies [57, 109], all groups performed the task with a mix of collaboration styles. They smoothly switched between loose and close collaboration during sessions. Although the extent of collaboration varied across groups, all of them appeared to be able to coordinate their actions to parallelize the work while initiating or reacting to requests or discussions.

7.2.2.3 *Interaction Challenges*

WALKING EFFORT Participants spent a lot of effort walking around while performing the task in the large scale manipulation task involving the wall surface. Typically, they would walk across the entire surface in order to find a suitable card to continue the story. Once a particpant found something interesting, he would bring the card to the working point of the story line and put it there. It also happened frequently that participants walked towards each other so that one could show a card to the other one. Physical movements and communication helped to compensate for the possible reduction of group awareness when the task was performed in a large space. Nevertheless, this caused physical fatigue over time.

TEMPORAL STORAGE NEEDED Participants sometimes held multiple cards in hand while walking or manipulating other cards (Figure 66-c2). This happened after they picked multiple cards that they anticipated to place soon. This helped reducing the need for walking back and forth, and also to not forget or lose a card. This suggests the need of a temporal storage or a queue that would be easily accessible when the users move around, or the ability to highlight objects to make it easy to find them later.

7.2.3 *Summary*

Experimental task vs. real user task

The task in the observational study was designed to operationalize the same kind of interactions as in the real task involving the sociologists (Section 3.2). The sociologists aimed to uncover hidden truths in a complex data set, while the goal of our participants was to be imaginative and creative. However, both tasks required users to look into details of each data item, to try finding relations between them and to use spatial arrangements to express these relations. Yet, discussions between users helped improve the quality of the results in both tasks. Both tasks involve similar data manipulation for assisting a high-level cognitive task that benefits from encoding meaning in space. Improving task quality is the motivation for users to collaborate in both tasks. Therefore, we expect that the interaction phenomena that emerged in the study reflect the real situations of the sociologists.

This observational study shows how users interact with each other and with physical artifacts for a task involving collaboration and data manipulation. Deictic actions leverage humans' common understanding about others' actions and facilitate their conversation by providing common references and understanding. Showing objects to each other enables a negotiation, while passing objects effectively delegates tasks. Holding multiple objects in hand creates an easy-to-access storage for a task queue. These insights gained from the study could help us design embodied interaction in an interactive environment with wall-sized displays.

In the following sections, I introduce the design, implementation and evaluation of a set of novel techniques that are designed based on the insights gained from this study.

7.3 COLLABORATIVE GESTURES

Collaborative Gestures are a set of techniques that facilitate data exchange between partners in small groups. Each collaborative gesture is composed of multiple individual gestures and triggers one command. For example, one user performs a throw gesture on a data item towards another user, another user can "catch" it by performing a catch gesture afterwards, the item would move to the position of the catch gesture. The following principles were applied in the design process:

1. Take advantage of co-located communication and coordination.

2. Support both synchronous and asynchronous actions for mixed collaboration styles and ease their transition.

3. Allow negotiation between partners for collaborative actions.

4. Support the forms of division of labor that facilitate the sharing of knowledge and physical resources.

7.3.1 *Design and Prototype*

The first prototype recognizes a set of basic hand gestures: *Finger Tap*, *Finger DoubleTap*, *Finger Drag*, *Finger Swipe*, *Finger Zigzag*, *Hand Tap*, *Hand DoubleTap*, *Hand Drag*, *Hand Swipe*, *Hand TouchHold*. Most of these are standard multi-touch gestures. *Finger* gestures are distinguished from *hand* gestures by the number of fingers touching the screen. Gestures performed with more than three fingers are recognized as *hand* gestures. *Finger Zigzag* is activated when a user draws a "zigzag" shape with one finger with at least three sharp corners (angle less than 1.2 in radian).

Three collaborative gestures: Preview, Throw and Catch, and shared grouping, are composed by each user performing above standard gestures. The user who starts performing the gesture is called the *action initiator*, and the user who performs the following gestures in order to complete the operation is called the *action follower*.

7.3.1.1 *Preview*

The *Preview* gesture allows one user to show an item to a partner by having both users touching the screen concurrently. The action initiator touches and holds one item on the screen with a *Hand TouchHold* gesture (dwelling for more than 400 ms). A *Preview* gesture gets recognized if an action follower performs another *Hand TouchHold* gesture anywhere else on the screen, before the action initiator releases her hand from the surface. The activation of this gesture creates a temporary copy of the item under the hand of the action follower. Figure 67 shows the state machine diagram for this copy. The copy appears under the hand of the action follower and can be dragged around like other items. The copy gradually vanishes in the absence of interaction, and fully disappears after 6 seconds. In case the action follower wants to keep the real item, a *Finger DoubleTap* on the copy "grabs" the real item over and deletes the copy.

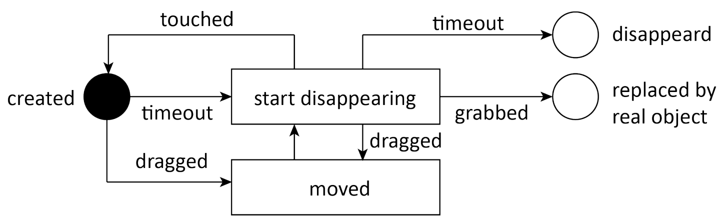


Figure 67: State machine of the "fake" object created for *Preview*.

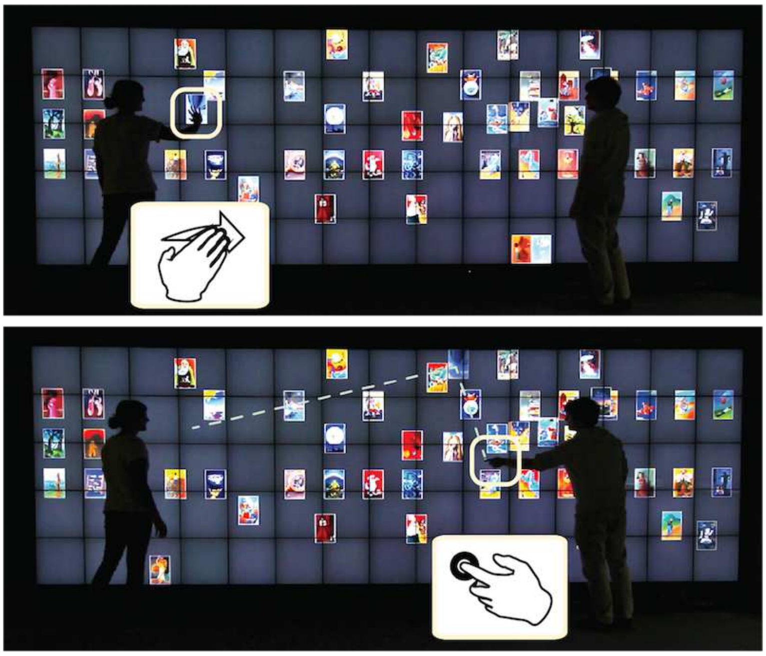


Figure 68: Throw and Catch gesture. On the top, the action initiator performs a *Hand Swipe* towards her partner. On the bottom, the action follower performs a *Finger DoubleTap* to move the image to his hand position. The dashed line shows the object trajectory in this process.

This enables one user to "show" an item to another user regardless of the distance between them. It is inspired by the observation that participants frequently showed a card to their partner.

7.3.1.2 *Throw And Catch*

As we can see in Figure 68, the action initiator throws an item towards the action follower with a *Hand Swipe* gesture. Figure 69 illustrates the state changes of the item. It "flies" in that direction with a friction that slows it down until it stops. The item blinks for 10 seconds, indicating that it is available to be caught. Before the timeout, the action follower can catch the item with a *Finger DoubleTap* gesture. To reduce false positives when more than two users work concurrently, a *Finger DoubleTap* is only recognized as a "catch" if it is performed at the direction where the item was thrown to, relative to the action initiator.

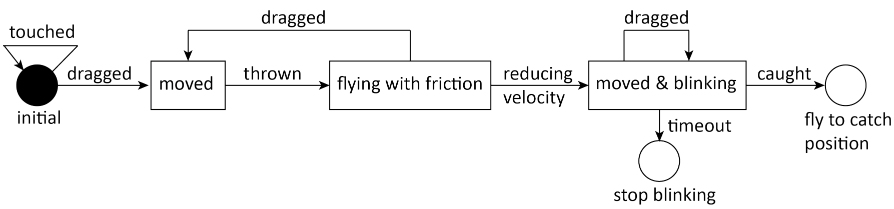


Figure 69: State machine of the object when performing *Throw and Catch*.

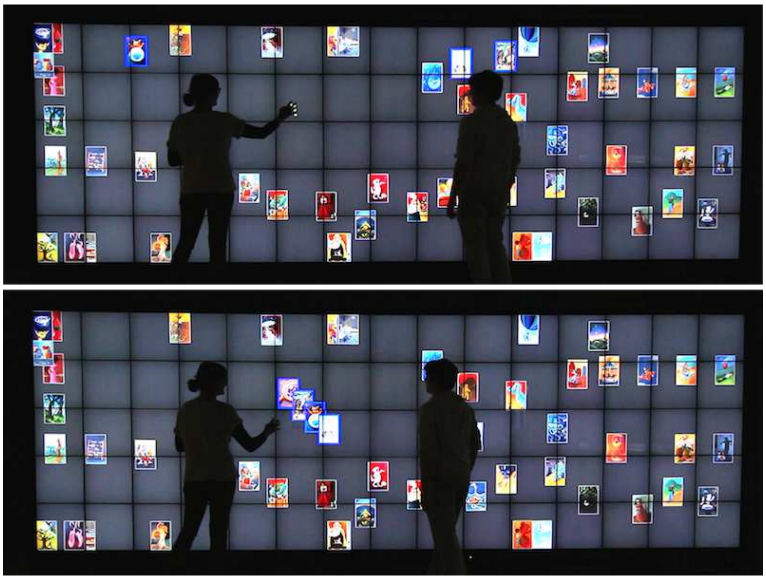


Figure 70: Shared Grouping gesture. Top: the images added to the shared group are highlighted with blue border. The girl is performing a collect action by a *Hand DoubleTap*. Bottom: after performing the collect action, all items are stacked around her hand.

In this prototype, only left and right directions are distinguished. The item then flies to the position where the *Finger DoubleTap* gesture was activated. After the timeout, the item stops blinking and the collaborative gesture is canceled.

7.3.1.3 *Shared Grouping*

By default, there is a "group" or "storage" for items to be shared by the pair of users. Before adding any item into it, it is empty and invisible. An item can be added by performing a *Hand Tap* on it by any user. The added items are highlighted with a thick blue border (Figure 70). Items can be also be removed from the group by tapping it with a *Hand Tap* again, which will remove the highlight effect. The items in the shared group stay where they are until they are *collected* by a *Hand DoubleTap* performed anywhere on the display. This will stack all the highlighted items at the position of the collect action. The *Hand DoubleTap* can be performed repeatedly on any position to re-stack and move all the highlighted items. Items can be added and removed before or after any collect action. A *Finger Zigzag* gesture performed anywhere ungroups all the items in the group, which will remove their highlight.

This gesture allows any user to select items for the shared group and stack them anywhere on the display. Dragging any item in the group with *Hand Drag* moves the entire group while preserving the relative positions between the items. The relative positions between the items within the group can be arranged by dragging single items with *Finger Drag*.

7.3.1.4 *Synchronization*

These techniques require different levels of synchronization between the collaborators' action. As illustrated in Figure 71, *Preview* requires the highest level of synchronization between the pair's actions because it is activated when both users touch the display concurrently. *Throw and Catch* requires less synchronization as there is a time tolerance of a few seconds between the throw and catch actions, which also makes it possible for a single person to perform it sequentially. *Shared Grouping* can be performed in a complete asynchronous way, without any time constraint. However as there is only one shared group at a time for a group of users, they need to finish with one group before starting another.

Synchronization tolerance aims to support both loose and close collaboration and to allow fluid transitions between them. However, as there are overlapping functionalities, users might need to coordinate about which gesture to use when both of them intend to start a gesture at the same time.

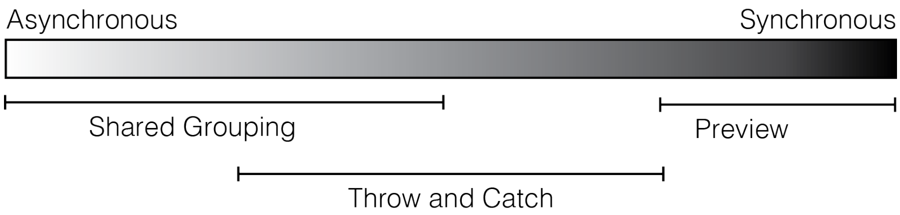


Figure 71: The range of synchronization between collaborators supported by each collaborative gesture.

7.3.1.5 *Implementation*

I implemented a prototype in JAVA based on the ZVTM engine[2], which supports graphical rendering on a cluster of computers. The prototype runs on a computer that communicates with the cluster controlling the wall display through a local network.

The application consists of two main components: graphical rendering and input handling. The InputHanler receives events from the gesture recognizer, which listens to TUIO[3] touch events from the the PQ Labs[4] driver through the network.

Divide Touches to TouchGroups

To enable the concurrent input of multiple users with both hands, each new touch is classified into a TouchGroup when a Touch_Down occurs. Each TouchGroup represents one hand. In this prototype, a

2 http://zvtm.sourceforge.net/
3 http://www.tuio.org/
4 http://www.multitouch.com/

simple distance threshold is used to determine whether a new touch belongs to an existing TouchGroup. Each TouchGroup runs its own recognizer to classify finger or hand gestures.

The Collaborative Gesture Recognizer listens to the finger and hand gesture events and activates collaborative gesture events. All gesture events are listened to by an InputHandler that processes these events and performs the corresponding operations, including calls to the rendering components to draw visual feedback. I use the Google Guava library[5] to facilitate event handling in the application. All gesture events are broadcast through the EventBus utility. Relevant components listen to them by subscribing for specific event types.

7.3.2 *User Study*

I conducted a user study to test the usability of the gestures and improve the design. The study uses the same open-ended task as in the observational study of arranging paper cards (Section 7.2), with slightly different instructions. The goal is to test whether the techniques can facilitate data manipulation under different collaboration strategies and to gain insights for design implications.

7.3.2.1 *Study Design*

A set of images from the Dixit game are placed on the wall display in random order. A few examples of the images are shown in Figure 64. As mentioned before, each image can have multiple interpretations. The task is to group them with creative or imaginary keywords by arranging their positions. Arranging the groups relative to each other is also encouraged. Participants are encouraged to discuss the images and the layouts to help generate more interesting insights.

The study compares two conditions: *Single-user Gestures* and *Collaborative Gestures*. Both conditions provide standard gestures performed by single users, including *Drag* with any number of fingers, *Finger Swipe* and *Hand Swipe*. In addition to these, the *Collaborative Gestures* condition provides all three shared interaction techniques described above: *Throw and Catch*, *Preview* and *Shared Grouping*.

In order to be fair between the two conditions, an alternative technique for grouping images with a *Two-hand Hold* gesture is provided in the *Single-user Gestures* condition. It mimics the real world action of pushing piles of paper on the table with two hands in order to move them. To perform the gesture, a user places their two hands close together on the screen and performs a *Hand TouchHold* with both hands on top of a pile of images. This creates a temporary group containing all the images under the touches of both hands as well as the images overlapping them. The user can then move all the images by drag-

5 https://code.google.com/p/guava-libraries/

Figure 72: Example of initial layouts to start a measured trial.

ging the hands. Considering the effort of dragging two hands, the prototype allows users to release some fingers and drag the whole group with the remaining fingers. But, once all the fingers that initiated the gesture are released from the screen, the temporal group is released. This allows a grouping quasimode which avoids performing explicit ungrouping after moving items, at the price of keeping fingers in contact with the screen to preserve the group.

The gesture is recognized in two situations. One case is when two TouchGroups, each of them with more than four touches, concurrently perform *Hand TouchHold*. The other case is when one TouchGroup occurs with more than 8 fingers performing *Hand TouchHold*. This is because the prototype just uses a distance threshold to distinguish between hands, so that touches of two closely placed hands may be recognized as a single TouchGroup.

PARTICIPANTS We recruited 6 pairs of participants aged between 23 to 38. One pair was a couple, three pairs were close colleagues and two pairs were acquaintances.

APPARATUS The tasks were performed with the 6×2 meter WILDER wall-sized display with multi-touch capability (Figure 72). Participants' positions in front of the display were tracked through a VICON motion-tracking system. The software builds on the prototype described in the previous section and runs on a front-end computer that also logs experimental data.

7.3.2.2 *Procedure*

Participants were instructed to perform the task collaboratively. They were allowed to use any strategy except to split the task and work completely separately. They were told to find creative or imaginary keywords to group them. Examples of both good and bad keywords were given by the instructor. For instance "peaceful" or "world inside another world" are more interesting keywords than "sky" or "red". They were also encouraged to make associations between groups.

They were encouraged to discuss with each other in order to find more interesting relationships between images.

The study consisted of two sessions, one for each condition, *Single-user Gestures* and *Collaborative Gestures*, and used two different sets of 54 Dixit pictures. Each session included a training trial with vacation pictures from the INRIA Holidays dataset[6]. The instructor introduced the techniques at the beginning of the training, and the participants learned and practiced them by performing a grouping task using the holidays pictures. The measured trial, which used the Dixit pictures[7], started after the users felt comfortable and sufficiently trained to use the given techniques. Figure 72 shows one example of the initial layout of the images. The presentation order of the 2 conditions and the 2 picture sets were counterbalanced across groups. Each measured trial was given 15 minutes.

After each measured trial, participants explained the layout and the keywords to the moderator. Completing both sessions took about 60 minutes. Afterwards, each participant filled a questionnaire with 7-points Likert scales for preferences and workload. Finally, there was a short interview with each group to discuss their experience and suggestions for improvement.

7.3.2.3 *Data Collection*

The experiment software logged kinematic data including participants' head movements, gesture activation and the movements of images on the display. Final image layouts created in the measured trials were captured and annotated with the keywords presented by the participants. The measured trials were video-taped and the interviews were audio-taped for future analysis.

7.3.3 *Results*

7.3.4 *Task Completion*

All groups finished the task successfully in about 15 - 20 minutes for each trial. Figure 73 shows some selected layouts created by participants. On the top, the layout created by Group 5 is a typical example of the layout results in this study.

For all the groups except Group 6, the created layouts do not show obvious differences in terms of the number of image group, nor, for a given group, obvious quality differences between the *Single-user Gestures* and *Collaborative Gestures* conditions. I will explain the behavior of Group 6 in later analysis.

6 http://lear.inrialpes.fr/ jegou/data.php
7 http://en.libellud.com/games/dixit

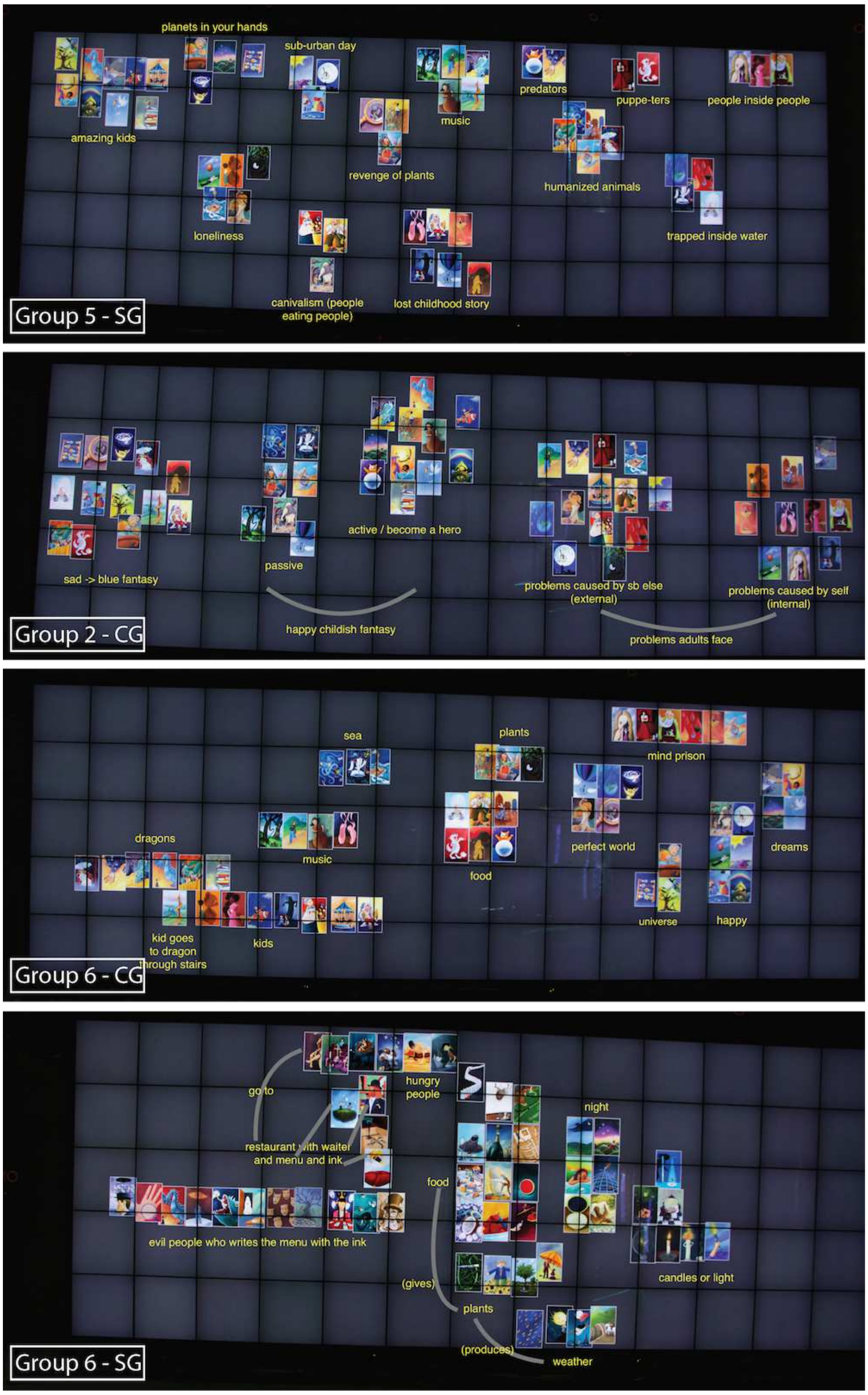


Figure 73: The selected final layouts created by participants. The yellow labels annotate the keywords spelled out by the participants. SG means *Single-user Gestures* and CG means *Collaborative Gestures*.

7.3.4.1 *Gesture Acceptance and Usability*

All of the participants preferred to perform the task with *Collaborative Gestures* rather than with *Single-user Gestures*. Several reasons were mentioned by the participants in the questionnaire and interviews. Firstly, *Collaborative Gestures* helped them reach remote areas. Secondly, they improved the positioning accuracy as the partner could "catch" a thrown object at a precise position. As one participant described, "*Collaborative gestures make it possible to have more accurate positioning of the pictures the first time we decided to move them. It's easier to have an image precisely reaching the other side of the screen.*" Thirdly, they reduced the physical effort of walking around, thus supporting activities such as negotiation, especially when the partner was far away. As one participant said, "*Even if they demand more coordination, the collaborative gestures allow us to be more efficient and more organized. Especially the preview mode which allows one user to gather one category, so that the other can quickly ask his opinion about another category at the other side of the screen.*" In the end, all participants mentioned they were playful and fun to use.

Benefits of collaborative gestures

This was confirmed by the ratings in the questionnaire (Figure 74): the ratings of enjoyment was significantly higher for *Collaborative Gestures* than for *Single-user Gestures*. We also observe that *Single-user Gestures* was rated more tiring than *Collaborative Gestures* (discussed in detail later), which might be a major reason for the preference for *Single-user Gestures*. No difference in mental load was observed.

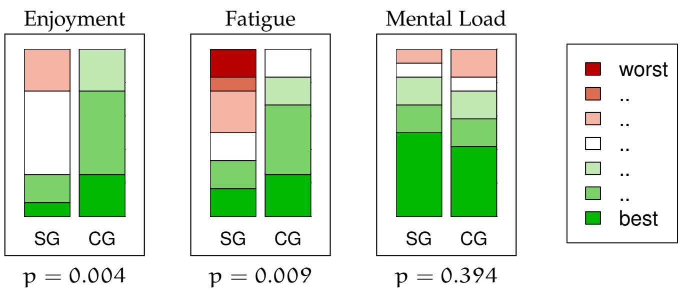


Figure 74: Comparison of the *Single-user Gestures* (SG) and *Collaborative Gestures* (CG) conditions (twelve participants) regarding enjoyment, fatigue and mental load. The p's give the result of the comparison of SG and CS using a paired Wilcoxon rank sum test.

Participants found that *Collaborative Gestures* helped with the existing collaborative operations. One participant who performed *Single-user Gestures* prior to *Collaborative Gestures* said, "*I felt that some collaborative gestures supported the kind of collaboration I had already held with my partner during the single-user gesture session. For example, instead of asking my partner to come to see a picture, I could "mirror" it for him with 5 fingers holding it. And it was easier to create groups together by selecting pictures with a hand tap over the whole space without walking.*"

Supporting existing collaborative operations

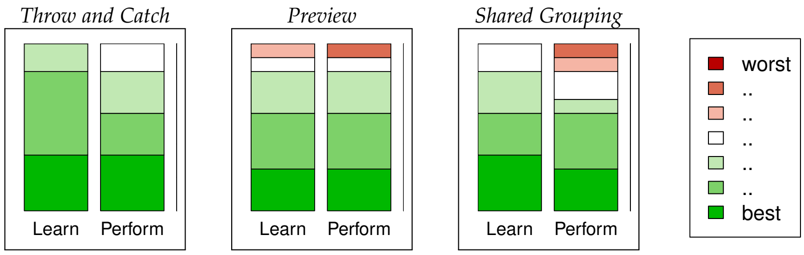


Figure 75: Results of the questionnaire for the twelve participants regarding easiness of learning and easiness to perform the gestures, for the three collaborative gestures.

Overall, the three collaborative gestures were accepted by the participants. Figure 75 shows the ratings of "Easy to learn" and "Easy to perform" for each collaborative gesture. *Throw and Catch* was rated the most easy to learn (average 6.2 out of 7) and easy to perform (average 5.6). Similar ratings are given for *Shared Grouping* (average 5.7 for easy to learn and 5.25 for easy to perform) and *Preview* (average 5.6 for easy to learn and 5.4 for easy to perform). The participants who found *Preview* not easy to perform attributed it to the overhead of coordination, while the participants who found *Shared Grouping* not easy to perform complained about the error recovery cost when they forgot to ungroup an existing group before starting a new group.

Participants found *Throw and Catch* useful when they "have a clear criteria of grouping category and just want to send a picture to the other side." The *Shared Grouping* technique allowed selecting multiple items across the display and stacking them together. Participants found it convenient to "assemble images when a theme was obvious". However, it happened from time to time that participants forgot to ungroup selected images after using the *Shared Grouping* technique. Then they selected images to create a new group, but in fact added them to the existing shared group.

Preview was more suitable when they needed to discuss an item at a distance. One participant particularly liked the *Preview* technique because it was "more comfortable to negotiate". However it was designed for the situation where a user wants to show an item to her partner thus initiates the gesture. The sequence of actions cannot be reversed in the current design. This is to distinguish which image to send over when both users have their hand on top of a data item. This introduces more coordination effort when one user requires the other to show an image, as they have to take care of who placed the hand first. Future designs should take this limitation into consideration.

Regarding gesture usage, Table 5 provides the number of times each collaborative gesture was activated in each condition per group. On average, they were used about 30 times per trial, which corresponds to about two collaborative gestures per minute. On average, the most used collaborative gesture is *Throw and Catch*, while two

Groups	*Preview*(grab)	*Throw and Catch*	*Shared Grouping*	Total
1	5(2)	5	9	21
2 (*)	4(2)	23	0	29
3	10(2)	5	9	26
4 (*)	5(4)	17	10	36
5	7(0)	11	18	36
6 (*)	4(3)	26	2	35
Mean±SD	5.8±2.3 (2.2±1.3)	12.8±7.2	8.2±6.2	30.5±6.2

Table 5: Number of time each collaborative gestures was used by the groups. Groups with a (*) started with *Collaborative Gestures*.

	Single-user Gestures	*Collaborative Gestures*	
Groups	Swipe & Drag	Swipe & Drag	Collab. Gestures
1	303	194	73 (27%)
2 (*)	339	410	53 (11%)
3	215	212	81 (28%)
4 (*)	239	127	111 (47%)
5	233	174	137 (44%)
6 (*)	352	168	74 (26%)
Mean±SD	280±59	214±100	88±30 (30%±13)

Table 6: Comparing the numbers of elemental operations: Drag-and-Drop (including two hands stack drag-and-drop) vs. each component of the collaborative gestures (e.g., for *Preview* we count two gestures and for *Shared Grouping* we count the select and unselect gestures). Groups with a (*) started with *Collaborative Gestures*.

groups used *Shared Grouping* the most and one group used *Preview* the most. Table 6 compares the number of elemental actions performed for collaborative gestures (all three of them) and for single-user gestures (hand or finger gestures for swipe and drag including *Two-hand Hold* for stacking images) in each condition. Although there is a large variability between groups, on average, 30% of the gesture actions were performed to activate collaborative gestures. Therefore we can see that *Collaborative Gestures* were used frequently during the study.

As a side note, no participant complained about false positives caused by unintentional gestures. The moderator observed three accidental activation of unwanted gestures through the entire study. As this study focused on pairs of users interacting together, false positive were seldom encountered. However, this may not be the case with a larger number of users, a problem that will need to be handled in future prototypes.

7.3.4.2 *Effort for Navigation and Drag*

As mentioned before, the *Single-user Gestures* condition was significantly more tiring than the *Collaborative Gestures* condition (Figure 74). Some participants mentioned in the interviews that *Collaborative Gestures* saved much of the effort of walking around. As one participant

Navigation effort

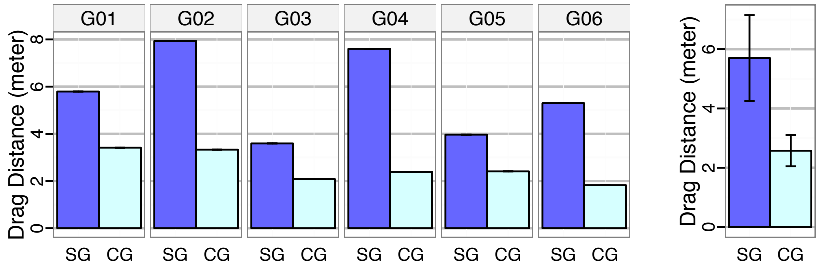


Figure 76: Drag distance in meter for *Single-user Gestures* (SG) and *Collaborative Gestures* (CG) for each group (left) and on average (right).

mentioned, "*It's easier to see preview of the pictures, I do not have to walk around.*" However, the average travel distance per participant was 183 meters with *Single-user Gestures* and 161 meters with *Collaborative Gestures*. For a duration of 15 minutes, this is a small difference.

It is known that dragging on a surface can cause fatigue [50]. Figure 76 shows that participants drag on the surface more under the *Single-user Gestures* condition than under the *Collaborative Gestures* condition. While there is a lot of variance between groups, on average, the drag distance was reduced more than half with *Collaborative Gestures*. Therefore, the subjective reduction of fatigue with *Collaborative Gestures* is likely to be attributed to reduced drag effort. As a side remark, the differences in drag distance are larger for the groups that started with *Collaborative Gestures* and these groups also dragged more in *Single-user Gestures* than the groups that started with *Single-user Gestures*. There is possibly an order effect here. The way in which the groups performed the task with *Collaborative Gestures* might have influenced it with *Single-user Gestures*, possibly leading to more engaged manipulation.

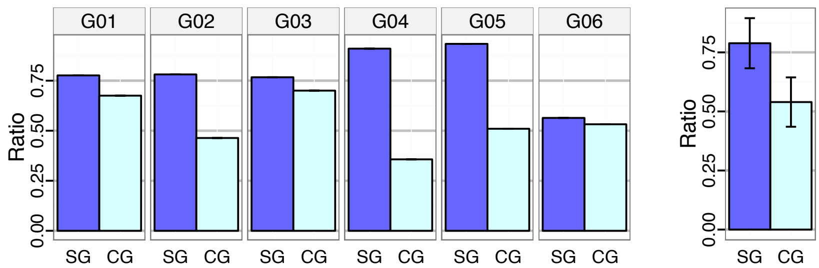


Figure 77: Ratio of drag distance divided by images movement for *Single-user Gestures* (SG) and *Collaborative Gestures* (CG) for each group (left) and in average (right).

As expected, collaborative gestures reduce the need of dragging images with direct touch. However, some groups performed the task differently between the two gestures condition (for instance, they would move less images). In Figure 77, we plot the average ratio of the drag distance divided by the object movement. This indicates better the drag effort spent for reaching the same distance of object movement.

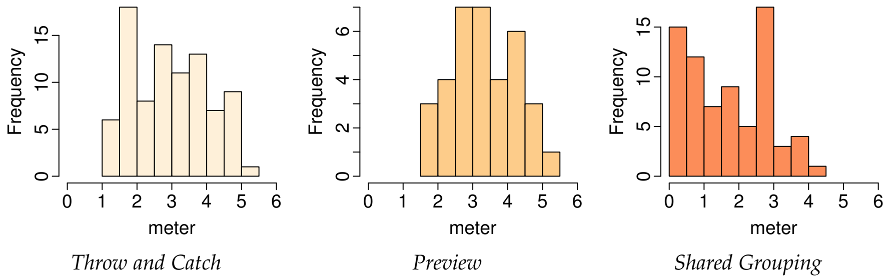


Figure 78: Distribution of the distance in meter between the two touch gestures (e.g., throw and catch) needed to perform the *Throw and Catch* and the *Preview* collaborative gestures. For *Shared Grouping* we consider the average distance between the selected objects and the gesture that group the selected images.

Note that we do not count the drag on the temporal objects created for *Preview*. A small ratio indicates that a small fraction of object movements are performed by direct drag. A ratio of 1 means all object movements are performed by dragging. In the *Single-user Gestures* condition, the ratio is less than 1, because with swipe gestures, objects continue to move after fingers leaves the screen. Dragging image stacks is also a source for this.

For Group 2, 4 and 5, the ratio is much smaller for *Collaborative Gestures* than for *Single-user Gestures*, showing the operational power of *Collaborative Gestures*. However, again, the variance across groups is large. For Group 6 the ratio is almost the same for both conditions. In fact the object movement distance is much shorter for *Collaborative Gestures* than for *Single-user Gestures* (3.5m vs. 9.4m) for this group. Indeed, Group 6 massively used hand swipes (30 times) under the *Single-user Gestures* condition to move objects and pass them between partners. The ratios for Group 3 are also close for both conditions. They used *Throw and Catch* moderately and mainly used *Preview* and *Shared Grouping*.

7.3.4.3 *Pair distances and Territoriality*

As shown in Figure 78, all pairs mainly used *Throw and Catch* and *Preview* when they are relatively far away from each other. The effect is obvious for *Preview*, which is performed with a minimum distance of 1.63 meter and an average distance of 3.2 meter. *Throw and Catch* is performed with an average distance of 2.92 meter. According to the interviews, this technique was not only used to send objects to a remote area, but also used to move objects to a precise position by "catching" them. This measure for *Shared Grouping* has a different distribution. This is due to the fact that participants often selected objects from several areas and then collected them together in one position.

The average distance between a pair of participants is about 40cm larger with *Collaborative Gestures* than with *Single-user Gestures*. Indeed, *Collaborative Gestures* let pairs work closely even when they are far away from each other. This is also the case for the two groups that collaborated very closely for both sessions. The participants from both groups noticed that they were further away from each other in the session with *Collaborative Gestures*.

This is consistent with the findings in the experiment of the previous chapter about the fact that shared interaction techniques enable close collaboration even when the users are far away. However in this study, since different collaboration styles are mixed together in each session, it is difficult to see a clear distinction between navigation patterns and to understand them. Figure 80, Figure 81 and Figure 82 (next page) show the traces of drag gestures on the wall and the physical navigation over time for three groups. There is no obvious pattern for several groups, for instance Group 4. However, these graphs visualize certain aspects of users' behaviors and provide some insights that will be useful for explaining collaborative behaviors later in this chapter.

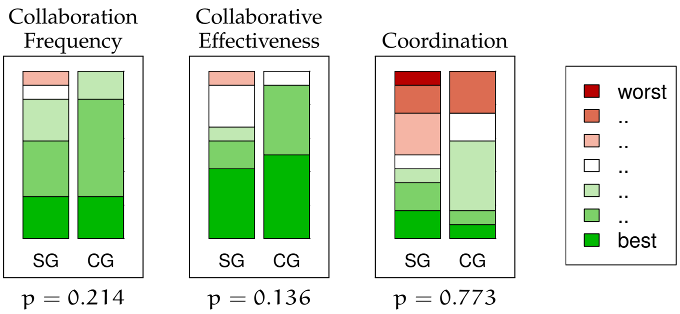


Figure 79: Results of the questionnaire for the twelve participants for the ratings regarding collaboration and coordination, comparing the *Single-user Gestures* (SG) and *Collaborative Gestures* (CG) conditions. The p-values give the result of the comparison of SG and CS using a paired Wilcoxon rank sum test.

7.3.4.4 *Collaboration Strategy*

Participants were asked about their collaboration strategies and had mixed answers. Two groups (Group 1 and 4) mentioned that they collaborated more with the *Collaborative Gestures* session than in *Single-user Gestures*. Three groups (Group 2, 3 and 5) used the same strategy and had a similar frequency of collaboration for both conditions. One group (Group 6) collaborated more in the *Single-user Gestures* condition. Overall, the participants' ratings for collaboration frequency and effectiveness show no significant differences between the conditions. However, *Collaborative Gestures* received slightly better scores (Figure 79).

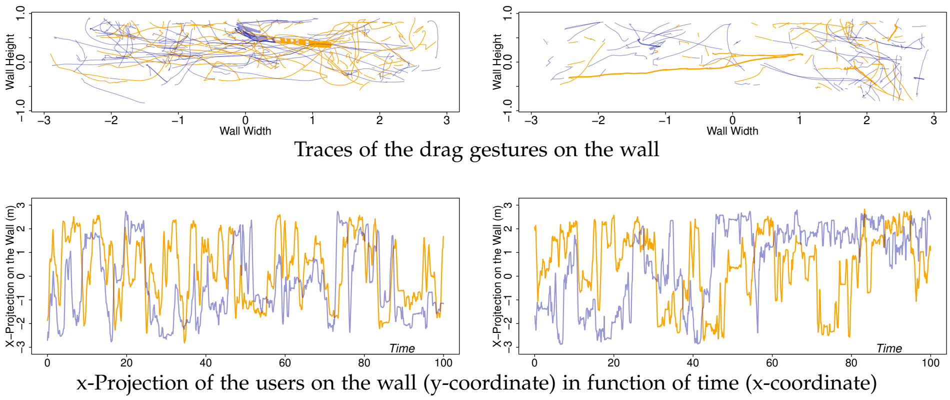


Figure 80: Gestures on the wall (top) and physical movement of the participants in front of the wall (bottom) for group 4. The left column is for *Single-user Gestures*, the right column is for *Collaborative Gestures*. Different participants are distinguished with color. The ownership of gesture traces is calculated by simply choosing the closest head position.

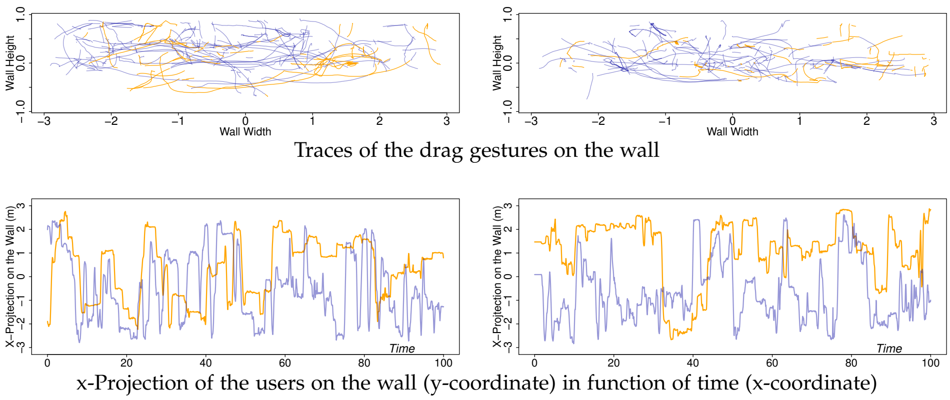


Figure 81: As above, gesture traces and physical movement of Group 5.

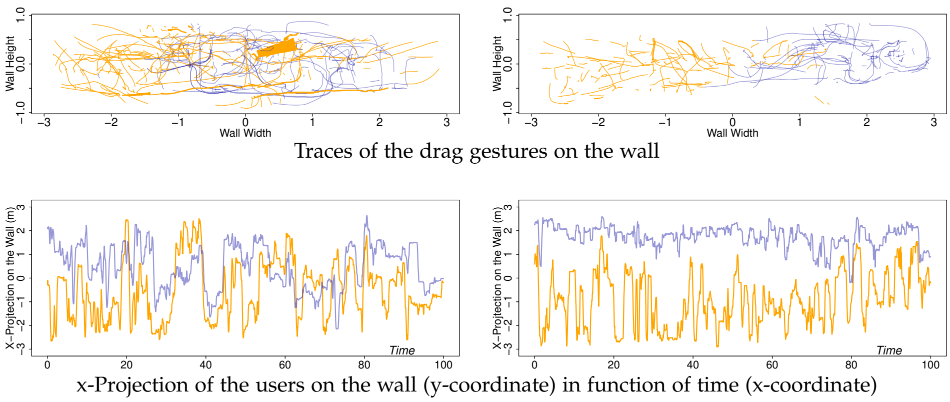


Figure 82: As above, gesture traces and physical movement of group 6.

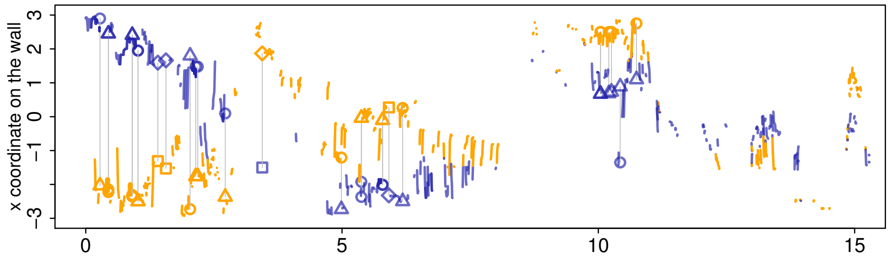


Figure 83: This is the gesture trace and object movement trance of Group 2 stretched over a time line on X axis. It takes a bird's eye view from the top of the wall display, thus the traces are all straight lines. Thicker lines are the trace of drag, distinguishing the user by color. Object movements are the gray lines between the gestures. *Throw and Catch*(○ → △) and *Preview* (□ → ◇).

For groups that collaborated more in *Collaborative Gestures*, one participant said, "*I did something without discussion in Single-user Gestures but with Collaborative Gestures we were working together on the same thing.*" His partner also mentioned, "*in Single-user Gestures it was more difficult to collaborate, because it is more difficult to put images to another side or to share with another person. We collaborated less efficiently.*" Another participant also indicated that "*(with Collaborative Gestures) presenting the idea about a group to my collaborator was more interesting, because he could send me pictures that he thought were appropriate*".

As groups that used the same strategy (group 2, 3 and 5) in both conditions, they were collaborating closely all the time, leaving little area to improve in terms of frequency of collaboration. Although using the same strategy, the participants mentioned differences in terms of effort between the conditions. One participant said, "*With Single-user Gestures I felt more like an individual, I grouped the pictures before giving them. With Collaborative Gestures I swiped them over one by one, ... and I knew people will be able to receive it and place it at the right place.*" "*With Single-user Gestures, the decisions (about moving pictures) were YES or NO. With Collaborative Gestures, we had more complex reflection.*" However, another participant felt differently: "*(With Collaborative Gestures) I felt I was more alone in my task because I could get preview from the other. It allowed me to focus more on my theme. I got less disturbed.*" This may depend on the personality of users and the subtle social relations between partners.

Group 2 decided on a strategy through discussion at the beginning of the first measured session and used it throughout the whole study. This pair decided to categorize pictures into two large sets first, for example "happy" and "sad". They used *Throw and Catch* extensively in the beginning with one person on the left and the other on the right side of the wall. Afterwards, they worked shoulder-by-shoulder for each subset and created smaller subsets within them. We can see the trace of this behavior in Figure 83 and the end result of this group in Figure 73.

Group 5 is another special case since the pair is a couple. The way they performed the task seemed to be influenced by their life habits as a couple. This group extensively used the *Shared Grouping* technique, which was frequently performed by the husband alone. As we can see in Figure 81, the husband (in blue) performed most of the interaction. The moderator observed a frequent situation, in which the husband was proposing ideas of grouping the images and performing the manipulation while the wife was standing back and giving opinions. Interestingly, with *Collaborative Gestures*, the wife started to perform tasks on the other side of the wall and controlled fewer of the husband's moves. Furthermore, it was observed that the husband frequently walked over to his wife when she intended to show him an image through *Preview* technique. The husband explained that he was used to get closer physically because it felt more like they were solving the problem together and he was more comfortable with the physical proximity.

In contrast, Group 6 collaborated more in the *Single-user Gestures* session. They explained that because they expected it to be uneasy to move items around, they often stepped back to discuss and plan how to layout objects. The fact that they had to walk around "forced" them to view the same overview and increased collaboration. As we can see from the final layout they created (Figure 73), they created many between-group associations, probably because of their planning activity while looking at the overview together. Whereas with *Collaborative Gestures*, they would rather stay in their own spot, so that they eventually talked less, as one of them said, "because we didn't have to walk more towards each other to communicate". This difference in behavior indeed largely affected their results. Figure 82 shows their trajectories over time. We can see a clear difference between the two conditions. With *Collaborative Gestures*, one participant barely moved for most of the session. The participants also explained that they learned a lot about performing the task in the first session (*Collaborative Gestures*) so they tried to use a better strategy in the next session. The learning effect for this group is large, including learning strategies to perform the task and getting familiar with each other (they were strangers before the study).

7.3.4.5 *Coordination and Communication*

Collaborative Gestures certainly require coordination between partners. Two types of coordination were observed during the study. One was coordination for *performing* the gestures. For example, after one person threw an image, another should catch it before the event timeout. The *Preview* gesture requires both collaborators to touch and hold the screen concurrently, thus needing a higher level of coordination. Participants needed to learn them and train to perform them.

Coordination for performing a collaborative gesture

Coordination for which gesture to use

The other type of coordination is about which gesture to use for one operation. Since all three collaborative gestures provide a function to move an image to another position, a participant "*had to be clear on which gesture the other had started to know how to finish it*". For example, one incident was recorded as follows: Participant A asked: "do you want to give me a preview"? At this time his partner had already selected a few pictures with the *Shared Grouping* technique, so A gave up with Preview and grabbed the group with a double tap instead.

Learning to use *Collaborative Gestures* involves not only performing the gestures, but also establishing a communication protocol between partners. As one participant said, "in the beginning, coordination is costly, but after learning it, it is not that much because you establish a word protocol like 'look' or 'catch' to give another the cue."

In fact, coordination issues also exist in *Single-user Gestures* sessions. A participant described an example: He had piled some images together somewhere on the display and planned to deal with them later. They disappeared after he came back, because his partner took them for something else. In fact surprisingly, Figure 79 shows that *Collaborative Gestures* even received a slightly better rating than *Single-user Gestures*. Some participants found coordination of *Collaborative Gestures* costly, while some others found it encouraged discussion.

7.3.4.6 *User Adaptation*

Preview and Shared Grouping gestures used by single users alone

Given the semi-synchronous nature of *Throw and Catch* technique, participants sometimes used it as a single-user gesture. They threw a picture and caught it by themselves, just to avoid dragging the image for a long path. It was used like a pick-and-drop technique [97]. Similarly, *Shared Grouping* can be performed in an asynchronous way. Participants sometimes used it to group multiple pictures for themselves. Interestingly, conflicts due to both participants using the same *Collaborative Gestures* alone were rarely observed. This is probably due to group awareness in a co-located environment.

Prepare for actions

Participants also adapted their actions to ease coordination. For example, one participant put a hand on an image whenever she started to discuss it with her partner. This way her partner could get a preview by placing his hand whenever he wanted to look at the image closely, without having to explicitly ask her to start a *Preview* gesture. If a discussion is not needed for this image, she could immediately throw it.

7.3.4.7 *Areas for Improvement*

Standard multi-touch gestures have been widely used for horizontal surfaces. In this study, some ergonomic problems appeared because gestures were performed on a vertical surface. I noticed the following problems in terms of performing multi-touch gestures on a wall dis-

Difficulties of performing multi-touch gestures on a wall display

play. First, they seemed to be relatively difficult to perform on some parts of the screen. For example, a five-finger gesture seemed to be performed well with the fingers pointing upwards when the hand was placed above the participants' shoulder. The lower the hand was placed, the more the wrist had to bend. The participant tilted the direction of the hand to compensate with the angle of the wrist. Finally when the hand was placed on the lower area of the display, participants performed the gestures with fingers pointing towards the ground, or with the body lowered. Second, performing a *Hand Swipe* gesture towards the fingers' pointing direction was difficult. The relatively large friction between the fingers and the touch surface, which is made of glass, might be part of the reason in this case.

Feedback about partners' action

Three participants suggested to provide more visual feedback or feedforward to improve coordination. For instance, when the action initiator starts the first part of a collaborative gesture, there could be visual feedback in the action follower's peripheral view and feedforward to help him complete the gesture.

One participant suggested that it would be useful to send content to partners without needing immediate confirmation or acceptance. Hence, asynchronous data exchange seems to be needed in certain cases.

7.3.5 *Discussion*

7.3.5.1 *Reflection on Methodology*

The controlled experiments in the previous chapter evaluated the effect of a shared interaction techniques with an abstract classification task (Section 6.4.1), while a semi-structured observational study in this chapter evaluated such techniques with an image arrangement task. Some findings are common to these two studies: such techniques reduce physical effort, allow close collaboration when users are far away and provide a fun experience. Similar patterns of territoriality were also found for some groups.

However, there are also differences in the findings. In the controlled experiment, a shared interaction technique, drop-for-partner, significantly increased the amount of collaboration. This effect is not obvious in this study. Based on our observations, the result of this study is largely influenced by the personalities of the participants and the social relations between pairs, which leads to different problem-solving styles. Learning effect seems to be another issue, as two groups mentioned that they would have been more efficient using *Collaborative Gestures* after longer training. Indeed, in contrast to the controlled experiment which was relying on an abstract task and instructions to enforce strategies, participants of this study were free to choose their strategy. Participants may have been able to quickly adapt their behavior to optimize interaction efficiency in the controlled experiment because of the simplicity of the task and of the given instructions. This was not necessarily the case in this study because the task was more complex and participants experimented various strategies with different goals. For instance, in the beginning a participant might suggest a strategy that might or might not be optimal but accepted by the partner for social reasons and used for the whole session. Moreover, in a task involving intellectual discussions, people may or may not optimize their behavior for interaction efficiency. More considerations are involved in the choice of strategies, such as the comfortableness in conversation, etc.

This example shows that observational studies capture different aspects of a task than controlled experiments. Controlled experiments help understand causal effects while observational study give insights about real-world situations. In our cases, we can see that the choice of a collaboration strategy is a collective decision influenced by social factors in the observational study, while interaction efficiency was optimized in the controlled experiment. Interaction designers should provide support for different social situations.

7.3.5.2 *Enable Collaboration at Various Distances to Data*

This first prototype is a stepping stone to explore the usability of shared interaction techniques. Each technique covers a different range of synchronization level between collaborators (Figure 71), but they all rely on direct touch interaction on the wall itself. To explore the design space further, the next prototype will enable such collaboration when users are at different distances from the wall display.

The observational study with the paper cards showed various physical arrangements of the users. They were sometimes both close or far away from the wall, or one was close and the other was far. The next prototype will enable users to collaborate with shared interaction techniques when they are at different distances from the wall. Variations of the collaborative gestures will be performed on the wall display and on a tablet simultaneously.

For example, the *Preview* gesture can be performed when one user touches the wall and another user touches a tablet. The data item touched by the person on the wall gets shown on the tablet, so that the tablet holder can see the overview of the display and the detail of a data item without walking back and forth.

7.3.5.3 *Beyond Pairs*

The current prototype focuses on two users collaborating with *Collaborative Gestures* at a time. While it supports more than two users working in the same workspace, it does not support multiple pairs performing the same collaborative gestures at the same time.

Identifying users is often important to support collaborative interaction. It has been a challenging problem on multi-touch interactive surfaces. For example, DiamondTouch [28] detects which user the touch belongs to and draws a shadow towards the user's position. iDWidget [100] uses the concept that a widget is aware of the user who interacts with it and responds accordingly.

Technically, in this prototype, the user of a touch event could be identified by tracking the hand position with optical markers on the wrist and mapping it with touch positions. As previously observed in the conference scheduling task (Section 3.4), people frequently paired up for discussion or close collaboration. Understanding the collaboration between pairs is essential before going into more complex situations with more users.

With increased number of users, one technical issue is the potentially higher false positive rate of gesture recognition. In terms of design, how to enable dynamic grouping between users? Groups could be predefined but might be cumbersome to change. Temporal partnership could be established via particular gestures, such as hand shaking between users. While humans can manage partnership dynamically with subtle social cues such as eye contact or approaching each other, how to indicate this to computer systems is a challenge. I will discuss this in the next chapter.

7.4 SUMMARY

This chapter began with a survey of related work on interaction techniques for assisting large scale interaction on large displays as well as existing techniques that implement a similar concept to *shared interaction techniques* defined in this dissertation. It is followed by the description of an observational study of pairs arranging paper pictures on vertical surfaces. This study gained insights of users' needs while collaborating on a large wall surface for this task and inspired the design of a set of shared interaction techniques - called *Collaborative Gestures*.

From an embodiment perspective, Collaborative Gestures leverage the social resources of co-located collaboration while enlarging users' interaction capabilities. As we saw in the first observational study with paper cards, various deictic actions facilitate discussion with common references and enable smooth transitions between their own tasks and shared tasks. Users shared the physical cards and kept them in hand to manage a task queue. They were able to coordinate their actions and avoid conflicts. This is a form of embodiment in a collaborative environment that involves interaction with physical objects.

Collaborative Gestures build on these embodied direct interaction between users in a co-located space and attempt to mimic users' operations with physical objects while augmenting their effects to empower users. The goal is to blend this augmentation with the collaborative practices and to use minimal technical support to leverage users' skills learned from interacting with the real world.

The three techniques - *Throw and Catch*, *Preview* and *Shared Grouping*, performed in different synchronization levels of actions, support both close and loose collaboration and enable a smooth transition between them. Informal evaluation of the first prototype showed that the techniques assist collaboration by reducing the physical effort of collaborative data manipulation on a wall-sized display, particularly by reducing the need for direct dragging on the surface. Participants were able to learn and perform the collaborative gestures, which supported different collaborative strategies and allowed smooth transition between them. After a short training session, participants were able to use them while focusing on the high-level cognitive task. The interaction with artifacts could blend in the interaction between users, as most participants did not feel more coordination effort with these techniques.

Along the line of the study of embodied interaction in group work in this chapter, next chapter introduces another set of techniques that augment direct gestural interaction between people to facilitate asynchronous data exchange and task delegation.

8 LEVERAGING DIRECT HUMAN-TO-HUMAN INTERACTION

This chapter introduces a novel concept - PoPle, which enables the exchange of digital information between people by augmenting direct gestural interaction between users. A set of proof-of-concept techniques are implemented as prototypes. I introduce the concept and explain the design choices based on embodiment and collaboration.

In a co-located collaborative situation, when one user needs to deliver information or delegate a task to another user, she needs to notify this user through verbal communication or by moving the data items close to this person, hence disturbing this user in performing their current task. If the user is not ready to perform the requested task right away, he will have to remember where the data item was placed and go back to it at a later time. This is especially problematic when users work on separate tasks in loose collaboration.

As mentioned before, some participants from the previous study missed a way of giving data to their partner without having to disturb their current task. It also happened that one participant could not find some prearranged images after coming back from another task, because his partner was not aware of it and used the images for something else. Given the frequent shifts in territory when users work on a wall display for such tasks, users need easy ways to handle content ownership and access.

This chapter introduces a set of techniques to enable asynchronous data exchange between collaborators while considering the content placement problem. It leverages direct human-to-human interaction in a co-located space for digital data exchange.

Tracking technology lets us create a large interactive space in front of a wall display. Interaction is not restricted to be with the screen, but also in the entire space. Proxemic interaction, for example, explores the use of relative physical position and orientation between users and devices as input parameters for sharing information or switching interaction modes [42, 70]. Eyeblog [27] augments the communication between people by recording video whenever it detects eye contact between the user and another person. SixthSense [74] mentions scenarios in which the name of a person is projected on his body when the user is close to him. However, existing techniques only take advantage of limited resources, such as position and orientation. The resources in co-located environments are not fully explored for interaction design.

In human-to-human interaction, there are many ways to directly interact with another person, such as looking at people, pointing to objects or other people, approaching while speaking, or giving people physical objects. As mentioned in the previous chapter, interaction between people is highly embodied and often done subconsciously. Here I explore an approach that capitalizes on the human skill of interacting with others by augmenting direct human-to-human gestures to transfer digital data. The prototypes are built in the context of collaborative data manipulation.

8.1 POPLE CONCEPT

I introduce a set of novel techniques called PoPle – Pointing to People – that enable one user to pick up data items with a pointer and send them to another person by directly pointing to the person with the same pointer.

The concept is that users are physical indexes of associated digital information such as data or tasks. The goal is to support the division of labor between collaborators and to facilitate asynchronous data exchange. From a human point of view, a user can transfer data to another user by performing a socially understandable "give" gesture. From a system point of view, users can be seen as "widgets with social properties", that are part of the whole system. The relationships between data and users are managed by the system and can be retrieved by physical interaction between users.

The PoPle conceptual model (Figure 84) supports interaction between users as follows: an *action user* selects a *target user* and issues a command to transfer data to the target user or ask the target user to perform an action. Feed-forward and feedback help the action user

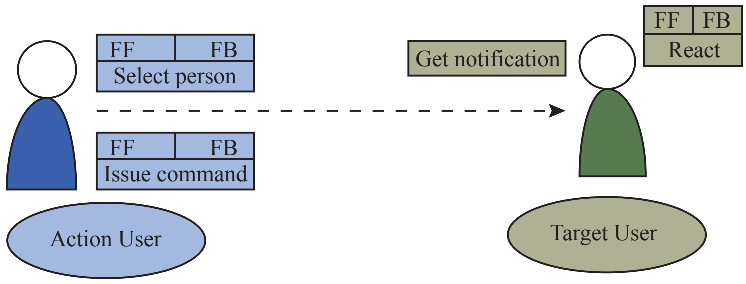


Figure 84: PoPle conceptual model: An *action user* selects a *target user* and issues a command. The target user is notified and can react to it. (FF means feed-forward and FB feedback).

select the target user and issue the command. The target user receives a notification as well as feed-forward and feedback about the data being transfered or the action to be performed, if any. The target user may act immediately or at a later time. The system may also automatically buffer the request if it determines that the user is busy.

8.2 PROTOTYPE

To explore the PoPle concept we designed a prototype, which is a bug-report classification application running on a wall-sized display. The categories are displayed in the top row of the wall display, the bug reports in the remaining space. Several users can interact simultaneously by using a mobile device. Each item is a bug report that must be classified into a category. Users can send a report to another person who has more knowledge about the topic. The goal is to emulate a collaborative task performed by people with different expertises.

Users interact with items on the wall through a motion-tracked pointer device that was used in the controlled experiment (described in Section 6.3.2). A user picks an item by clicking a Pick/Drop button. Then instead of dropping it on the display, the user can send it to another person by pointing to the user in the room with the same pointer and clicking the same Pick/Drop button. With the bug report classification application, users pick a bug report and either directly classify it, hand it off to another expert whom they think can handle it, or drop it back on the wall.

This prototype was implemented on the WILD display, which is a 5.5m × 1.8m tiled display with 140-million pixels. The position and orientation of each user's head and mobile device are tracked by a VICON motion-capture system. The pointer consists of a mobile phone with a motion-tracked stick attached to it (Figure 85). The software is written in Java, runs on a front-end computer driving the display cluster and communicates with the Android mobile devices through OSC messages.

8.2.1 *Interaction Techniques*

In order to explore the design choices and interaction scenarios, I designed and implemented several example techniques that demonstrate and instantiate the concept. In the following, I describe the various interactions I have designed and implemented.

8.2.1.1 *Sending Data To Another User*

In order to send data to another user (Figure 86a), the action user (left, in blue) picks an item on the wall using her mobile device, then points to the target user (in green) and sends him the selected item. In order to assist the selection of the target user, the mobile device vibrates when a user is in the line of sight of the device. The action user can then tap a button to transfer the item. The action is confirmed by a "drop" sound on the sender's device. The target user is notified of the transfer by a different sound on her device.

8.2.1.2 *User as Temporal Data Storage*

The previous chapter mentioned users' needs of having a easy-to-access storage for themselves to carry temporal data while walking around and manipulating data on a wall-sized display (Section 7.2.2.3). During a loosely-coupled collaborative task, a user's status switches between idle and occupied. In order to enable asynchronous data exchange and minimize interruptions, I designed a "buffering" approach to manage items received from other users. When the receiver is not available to deal with other tasks, the object is buffered in his "virtual space" and added to his task queue, which is displayed on the screen of his pointer. In Figure 85, we can see that the owner of the pointer had 5 data items buffered and three of them came from the "green" user (green item icons) and two came from the "blue" user (blue item icons). The user can pick an item on his cursor at any time by tapping it.

Users can choose among three levels of availability: *Immediate*, *Queue* and *Busy*. *Immediate* mode is for users working in tight collaboration: The transfered item appears directly on the wall display and is attached to the target user's cursor so that he does not have to pick it up. If the target user already has an item attached to the cursor, the received item is added to the queue. Buffered items are automatically retrieved and attached to the cursor (one by one, in receiving order) as soon as the user releases the item attached to his cursor.

Queue mode is intended for loose collaboration: Items are always buffered and must be manually retrieved by clicking on the desired item in the queue. This makes it possible to process items in a different order than the receiving order.

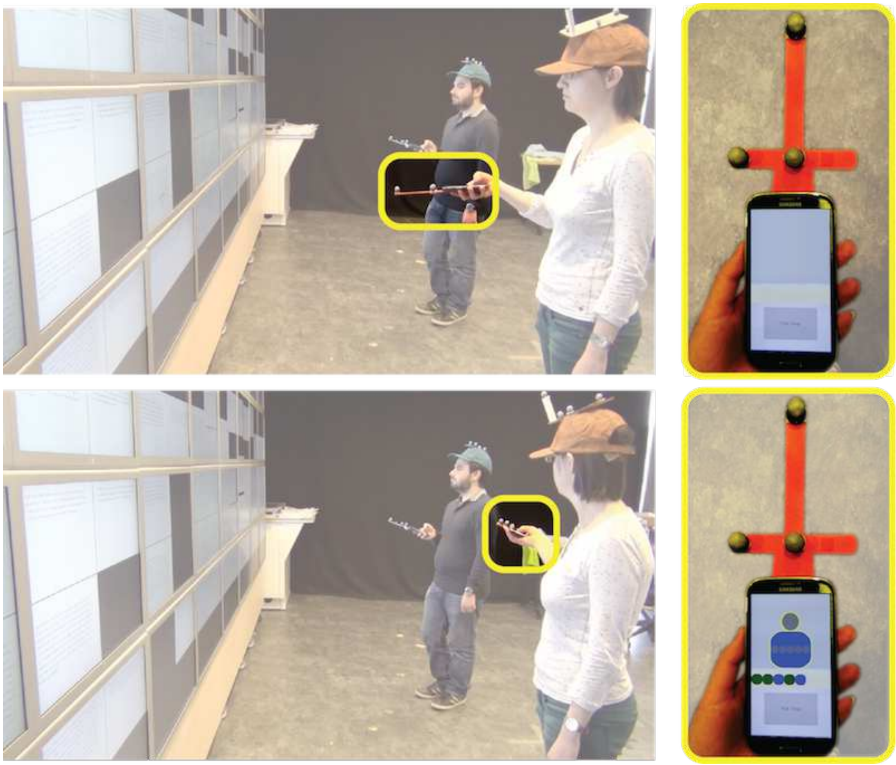


Figure 85: Top: a user (she) picks an item by clicking the Pick/Drop button on the bottom of the pointer. Bottom: afterwards she sends it to her co-worker by pointing to his direction and clicking the Pick/Drop button while pointing. Bottom right: while pointing on a user, the top of the mobile interface shows an icon of the pointed person with a task queue inside (gray dots), the middle shows a bar with the holder's own task queue, in which the items are colored corresponding to their sender.

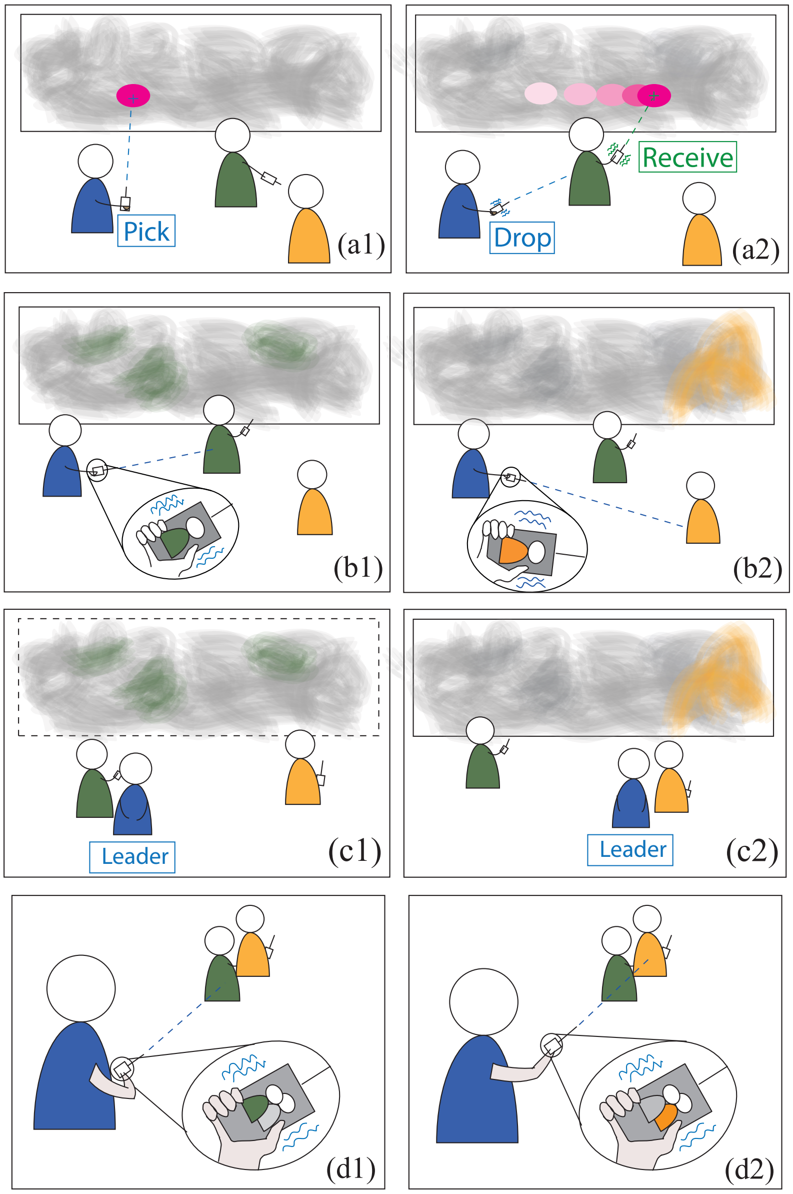


Figure 86: PoPle Techniques: Sending information (a), getting user information by aiming (b) or by proximity (c), disambiguating users (d).

Finally, *Busy* mode means that the target user does not accept items. Users who attempt to transfer items get a different vibration pattern as well as visual feedback indicating that this user does not accept requests.

8.2.1.3 *Showing Information About a User*

The prototype also lets users get real-time information about other users' tasks in order to provide awareness of the progress of each user. This provides a chance for users to acquire background information about the task progress of others without interrupting them. Such information is important to help users monitor and equalize their workload.

As for sending data, getting information about a user is achieved by pointing to that user (Figure 86b). Rather than using different buttons on the mobile device, I use the vertical direction of the aim to distinguish between these two functions: pointing to the upper part of the body sends data while pointing to the lower part retrieves user information. Technically it is easy to distinguish the vertical direction of the aim, so other functions could be provided in future designs, for instance to indicate priority (with high priority when pointing at the head and low priority at the feet). Similarly, pointing at the front vs. the back of a user could also lead to different functions.

User information can be displayed in a variety of ways. In this prototype, I use the wall display to visualize the overall activity of the target user and the mobile device to display the tasks queued on the target user's buffer. For example in Figure 85, the "blue" user (the user with blue hat) is the target user, thus the pointer's screen shows the visual feedback of a person in blue with 5 items in his queue (five gray dots inside the blue square).

An alternative way to get information about a user is to use proximity (Figure 86c): When a user (here the task leader, in orange) approaches another user, the wall display overlays activity data related to this user. This technique can also serve to highlight the data related to several users when they are close to each other.

These techniques provide a chance for users to acquire background information about the task progress of others without interrupting them. This may provide more participation equity and lure other users to help the one with the higher work load.

8.2.1.4 *Prioritizing With Face-to-face Communication*

Gestures in human-to-human communication often have subtle social meanings and implications for coordination. For example, when a user wants to delegate an urgent task to another user, she will probably talk to the user and ask to prioritize it. As an example, this prototype automatically prioritizes items when users are face-to-face when transmitting data. By turning his body towards the action user, the target user implicitly allows the system to insert the delivered item at the top of his queue. This interaction typically happens naturally after an oral agreement. It is implemented by tracking the orientations of users to detect the face-to-face configuration.

8.2.1.5 *Disambiguating Between Pointed People*

A consequence of using raycasting is that several people may be in or near the line of sight of the action user's device. When this is the case, the pointer screen displays the icons of the pointed users in a queue with increasing distance, with the selected one highlighted (Figure 86d). The user can change the selection simply by moving the device closer or further away from him, then taps the screen to send the item.

8.2.2 *Embodiment*

This set of techniques feature embodied interaction in several ways. First, a user's presence is the physical index to access all the associated digital data, including tasks and meta information. To exchange data with another user, instead of searching for a representation of this person, i.e., an id or icon, one can simply point to the user and click a button. Second, as exchanged data can be added to a queue managed by the receiver's pointer interface, it is as if the user contains a data storage that moves with him and can be easily accessed while walking around. Third, the technique takes advantage of the fact that people communicate face-to-face when a matter is important or special, highlighting the exchanged data in this situation. Finally, the techniques use the relative positions of users to distinguish between them when there is an ambiguity.

The goal is to allow users to exchange digital information in a more *direct* fashion. The data exchange can happen as a side effect of the users' interaction in the real world.

8.3 DESIGN CHOICES

When considering users themselves as interactive components in an environment, there are many possible ways for them to interact with each other. While in the presented proof-of-concept prototype I focused on a pointing gesture, there are many other alternatives. For instance one can interact with proxies of other users such as icons on their input devices, or with their digital shadows on the wall or on the ground, etc. Designers need to make choices depending on the tasks and user experience they want to support. This section takes an overview of various design choices I came across during the design process, and discuss their possible impact on collaboration.

Beyond the prototype, our concept considers a new way of collaborating in front of wall-size displays, and as such it can affect different aspects of collaboration. The design choices affect the awareness of user action and its mutuality, whether users are aware of others' awareness, as well as the availability of background information. These are important factors in collaboration according to the literature [117, 44]. In the PoPle prototype, the directness of actions between users, the system feedback of actions and notifications of receiving actions are the three important aspects that impact these factors. In the following, I discuss the design choices in each of these aspects.

8.3.1 *Directness of Action*

In distributed situations with standard GUI interface, users assign documents to proxies of remote users represented by icons or names [42]. In co-located environments such as tabletop interfaces, users normally share documents by putting them somewhere in the shared workspace [103]. These operations can go unnoticed or ignored and later forgotten by the target user when working with a wall-sized display, where people move around frequently, especially if the user is involved in another task.

I define the *Directness* of interaction between co-located users as one design dimension (Figure 87). Note that this is not about how direct the user interacts with the data, but how directly it is with relation to the target user. Some interactions involve the target user more directly, due to the natural movement and proximity of the action user.

Choices of directness

When both users operate on mobile devices, the interaction can be completely private. If operations ought to be anonymous (e.g. voting) or confidential (e.g. sensitive information exchange between the pair), this is a good choice. Interacting with controllers on the wall to initiate actions towards other people is more public than when using a mobile device and provides a higher degree of awareness, even though there is no guarantee that other users would indeed see them. Proximity is an even more direct interaction, but it is rather implicit, so other users may not notice it, unless it is followed by, e.g., a touch action towards the target user. On the other hand, if the user is ap-

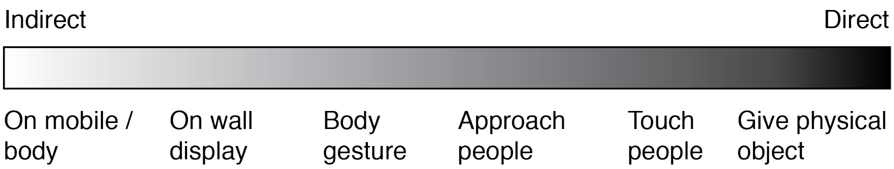


Figure 87: Spectrum of directness of the actions initiated by the action user.

proached, touched, or is given a physical document by hand, she is more likely to pay attention and remember the exchange.

The more direct the interaction, the more likely the target user is to get interrupted, but also the more aware she is of actions related to her. In this prototype I provide a “give” gesture performed by an action user to “select” a target user. This is more direct than using a proxy, and less direct than touching or giving physical objects. This choice favors embodiment compared to current GUIs, thus avoids the cognitive steps of finding a user through representations. In addition, it leverages a real-world practice - giving something to someone.

8.3.2 *System Feedback*

System feedback is crucial to notify users about the recognition of their commands to the system. As PoPle augments interactions between users to command to system, it is important for the action user to be notified about the system status.

Table 7 summarizes the general options of system feedback considered in the design of PoPle. When the action user initiates an action towards a target user, such as pointing, she needs to receive feedback that the action is recognized by the system.

Choices of feedback modality and location affect privacy and public awareness

Different modalities can be used for the feedback depending on how much awareness is needed for the action. Vibration feedback and sound feedback with an earphone is private, only perceived by the action user. Visual feedback provides more information, but requires attention. The notability then depends on where the action user’s current focus is. While visual feedback on mobile device is private, it would be more public if shown on the wall display as well as around the target user. The feedback location and privacy also affect the awareness of the action. For example a document that follows a user on the wall provides higher awareness information as to the ownership of this document, compared to a task icon displayed on the user’s mobile device or a document placed at a fixed position on the wall. Similarly a public sound ensures everyone hears it, but can be disruptive. Designers need to choose an appropriate feedback

Feedback / Awareness	Vibration on body/device	Visual on device	Visual on wall	Visual on/around target user	Sound
Action user	high	low	medium	medium	high
Target user	none	none	low	low	high

Table 7: General options for providing system feedback to the action user when or after she performs an action towards a target user. They likely provides different awareness to the action user and the target user.

location and modality depending on where they expect the user's attention to be focused.

For the current prototype I chose a combination of visual, vibration and audio feedback. When an action user points toward a target user, vibration feedback on the mobile device is provided to indicate that the pointing action is recognized by the system. This is to avoid distracting other users, at the expense of reducing awareness. A sound feedback is played on the action user's mobile device when data is sent. Informal tests showed that action users were able to send an item to another user without looking at their input device. In addition, the users could retrieve information about others' work load by pointing at their upper body. The workload information is displayed on the action user's mobile device (Figure 86A), while some overall task history related to the target user is shown on the wall. While we use explicit actions to get information about user workload, we considered displaying this information on the floor next to them but we lacked the technology in current prototype.

8.3.3 *Notification to Target User*

Not only does an action user need feedback from the system about the success of their action, a target user also needs notification about the actions happening to them, e.g., data transmission. The design of such notification likely influences the mutual awareness between users, which can have subtle affects on users' behavior depending on the task and situation.

The mutuality of awareness depends on the privacy choices of input and output, as well as anonymity choices. In this prototype, I use a sound feedback on the target users' mobile device to notify the receipt of data. It is a different sound from the feedback when sending data as an action user.

Anonymity choices determine how clear it is for the target user to know who initiated the action. The level of anonymity can be controlled by the design of the notification they receive. For example a different color or an audio message can represent each action user, if the action should not be anonymous.

In such collocated environments, some feedback can be perceived by both action and target users. For example a projection at the feet of a target user can be noticed and meaningful for both the action and target user. Note that anonymity can be affected by how private or public the feedback is for the action user. If the action user has a feedback sound, then the target user is likely to relate to a newly arriving task even if it was sent anonymously. Similarly, this depends on how direct the action user's interaction is. In addition, the location where a sound comes from is important for identifying the user of the action. Plain sound from public speakers would make it hard to identify the source.

In this prototype, different sound feedback is provided for sending and receiving data, on the pointer devices. Our informal tests show that this effectively provides mutual awareness. Current implementation does not provide explicit notification for being pointed. As show-info displays the target user's task history on the wall, he may be aware of this action by noticing changes on the wall.

8.4 BEYOND WALL DISPLAYS

While the prototype technique was designed for interactive rooms with large wall-size displays where groups of users collaborate and exchange data, the PoPle concept of augmenting human-to-human communication can be generalized to other co-located multi-user environments.

For example in an office environment, digital content could be shared between co-workers by augmenting inter-user interactions, for instance by pointing or swiping to the target user's direction. This could also be useful in classrooms for exchanging digital content between classmates or with the teacher.

Today people are connected through multiple networks and devices. When we need to exchange data with another user we need to specify a name or an email address, select an icon or an avatar, etc. We do not only use these channels when we are kilometers away from each other, but also when we are close by in the same room. For instance in an office, colleagues often talk about something interesting and then say:"I will send this to you via email". Why not take a more direct approach when we are co-located, instead of using remote communication channels?

Specifying which user to interact with the system, is a "people selection" action from the system's point of view. Let us first think about how people "select" another person to interact in the real world. When people start talking with each other, they use eye contact to establish a communication channel. Talking while looking at another person shows whom this conversation is targeting. If the target person is focused on something else, you may need to call her name to get her attention. Similarly, proximity or physical interaction can also imply the target of an action or conversation.

Continuous conversation with the "selected" people implicitly retain them as the target person. Once the conversation and eye contact stop for a while, the communication channel implicitly disappears and a target person of the next conversation needs to be "reselected", unless there is only one person around. Furthermore, people can also be "unselected" in other ways. For example, change of language in a conversation "unselects" the people who do not understand the language. Putting earphone “unselects” oneself from a conversation. The deictic words in the conversation can “select” and “unselect” people too.

The reason why I look at the physical ways of establishing communication channels is to be able to augment such embodiment resources with technology. The advance of tracking and sensing technology enables computerization of various human activities. Users' motion, ranging from body scale to room scale and then to transportation, can all be detected precisely. Eye contact can be detected through eye-tracking technology. We need to find ways to augment such activities to overcome physical constrains without disrupting their existing activities, while keeping users in control of the technology power.

8.5 SUMMARY

This chapter introduces a novel interaction technique - PoPle - to support asynchronous data exchange in small-group collaboration with data manipulation. PoPle allows users to transfer information and delegate tasks by simply pointing to each other with the same pointer device they use to interact with data on the wall display. A queuing mechanism is introduced to assist asynchronous task execution between users and reduce disruption. Users' territories are shifted or extended from the shared surface to the space on and around their bodies. Therefore the idea of what is a territory differs from existing work on tabletop [103], as well as on ambient wall displays [4]. We expect PoPle to facilitate a fluid division of labor in collaborative tasks with data exchange. In situations where content is difficult to process (e.g. long textual reports) and if there are users with different expertise, collaborators can save time by giving content to the right person to process it. These techniques may potentially enhance and smooth the collaboration between users, especially for time critical tasks with complex data sets.

The PoPle concept leverages direct human-to-human gestural interaction to carry out an action with virtual content, thus blending physical and virtual interaction. PoPle is a form of embodied interaction, but the embodiment is different from that of Collaborative Gestures. While Collaborative Gestures attempt to mimic user actions performed with physical objects, PoPle directly augments users' social actions. Collaborative Gestures face the challenge of striking a balance between maintaining the metaphor and adding magic. PoPle requires a careful design of the system feedback to notify the users about the system's behavior. It also requires to give users full control of the augmentation and to avoid interference with the users' social interactions.

"We know more than we can tell."

Michael Polanyi – The Tacit Dimension (1967) [91].

9 CONCLUSION AND PERSPECTIVES

Large wall-sized displays can display a large amount of information over space, so that users can navigate the data by moving their body and benefit from smooth transitions for viewing different parts of it. In contrast to a desktop environment where users are sitting in front of a window using a mouse and a keyboard to explore data, large wall-sized displays allow users to walk closer or further away to see different parts of the data at different levels of detail.

Interaction is not limited to the display surface in such environments. Location-tracking technologies such as VICON motion systems or Kinect depth cameras make it possible to capture various body- or room-scale actions occurring in front of the display. These technologies enable the augmentation of users' actions both for interacting with the display and with each other.

High-resolution wall displays promote *physical navigation*, which essentially allows users to pan and zoom virtual content by performing a familiar physical activity: walking around to explore a space. This situation can be seen as a case of *embodied interaction*. In this thesis, I describe "embodied interaction" as interaction that engages users' body for sensing and acting in an environment, while taking advantage of their skills learned from interacting with the real world. In the context of large data manipulation, I focus on embodiment in performing actions for interaction with virtual content, by either single-users alone or co-located multiple users working in collaboration. Therefore I look at embodied interaction from small to large scale, and from individual to social level.

9.1 CONTRIBUTION

This dissertation explores the benefits of embodied interaction in large interactive environments such as a room with wall-size displays. It increases our understanding of interaction phenomena and introduces new interaction techniques for supporting collaboration in such environments. It contributes to the HCI literature in several ways.

Firstly, this work contributes to fundamental research by exploring the benefits of interacting with large wall-sized displays. I conducted a series of controlled experiments to evaluate and better understand these benefits for data manipulation tasks. In the single-user case, they seem to be attributed to the embodied interaction that lets users navigate the data with their head and body movements and manipulate items with their hands at the same time.

In collaborative situations, I show that collaboration between users working in parallel can incur costs on interaction efficiency because of multi-tasking and disruption. Nevertheless, I showed the benefits of providing an interaction technique that allows each user to collaboratively perform single actions. I call such techniques *Shared Interaction Techniques*. A shared interaction can be highly embodied, when it takes advantage of the co-located resources, such as communication with direct deictic actions, awareness and coordination between users, etc. This approach had not been explored in research on wall-sized displays.

Secondly, I have explored different shared interaction techniques on a touch sensitive wall display. For this purpose, I designed, implemented and evaluated *Collaborative Gestures* to support data manipulation and exchange between users in close and loose collaboration. Moreover, I have explored the concept of leveraging and augmenting co-located communication between users with the PoPle technique, which allows users to exchange data asynchronously by pointing to another user's position with the same pointer as for interacting with the wall display. The embodiment of this technique is that it integrates users' physical presence as part of the interface to the system, thus enabling data exchanges with direct interaction between users.

Thirdly, I designed an abstract classification task that operationalizes various factors that affect interaction with a wall-sized display. It was designed by removing the task-dependent cognitive part of a real task while preserving interaction with the data. This task relies on labels for classifying data items and on font sizes to control information density. Thus, time performance can be measured and compared across conditions. This task provides a testbed that can be easily replicated and modified for future experiments. Indeed, Jakobsen and Hornbæk [58] used it to evaluate the effects of locomotion on interaction efficiency on wall-sized displays. Moreover, collaboration can be evaluated in a novel way using this task because different collaboration styles are intentionally separated. This offers a way to study collaboration in a more controlled manner.

Lastly, my observations and interviews of real users performing real tasks with high-resolution wall displays provide insights about the use of such devices.Encoding meanings in space helps users to associate data with positions, and provides common references and even vocabularies for discussion in collaborative cases. Arranging data items is not only a mechanical action, but also a thinking process as well as a discussion and negotiation process.

9.2 PERSPECTIVES

9.2.1 *From Real World to Abstraction*

My work in this dissertation began with the observation of a conference scheduling task. The insights I gained inspired the first research question about comparing a wall display and a desktop computer for manipulating a large amount of data. In order to be able to compare time performance across conditions, I constructed an abstract classification task based on the real task.

The choice of real or abstract tasks for experiments involves trade-offs between external and internal validity. The more abstract it is, the more likely the data is less affected by irrelevant noise, increasing internal validity. External validity may decrease with the level of abstraction, as such tasks may not reflect rich real-world situations. On the other hand they may be generalizable to more tasks.

I used a semi-structured observation study to test Collaborative Gestures (Section 7.3.2). The purpose was to provide an informal evaluation and to get insights for future design iterations. This study preserves the complexity of the real world, including influences caused by the relationships between the pairs of participants, content-based biases, etc. It has a high external validity at the cost of a low internal validity, as, for instance, repeating the experiment may generate different results, due to, e.g., different characteristics of the pairs.

For the controlled experiments, the abstract task allowed us to find a middle point on the spectrum and to reach a balance between both aspects. The large effects in the measured results of our controlled experiments attest their internal validity. External validity comes from two aspects. On the one hand, classification is one of the basic tasks that are involved in many real tasks with a large amount of data, as shown in Chapter 3. On the other hand, this task preserves some real world complexity, e.g., searching, memorizing and moving items for single-users, and co-located communication and coordination in the case of multiple users. The value of such controlled methods is to attribute causes to a phenomenon. Such findings can be complementary to the insights gained from observational studies. I believe both observational studies and controlled experiments are needed to provide a comprehensive understanding of phenomena. Both types of understanding may inform better interaction design.

Since one experiment always necessarily chooses instances of some situations, a more complete understanding requires future experiments that replicate the studies with different factors. For instance, data sets with geographical shape might require different sorts of interaction. Such use cases need further investigation in future work. Also, the single-user experiment can be replicated by comparing the wall display with a larger desktop screen. With the highest information density in Experiment 1, the reach range on the wall display was 6 times as large as the desktop. Would a power-desktop with the size of 6 desktop screens reach the performance of a wall display?

For the multi-user experiment, other tasks involving intellectual communication and negotiation, or generative tasks such as brainstorming could be studied with similar approaches. In these cases, the measure could be the task quality instead of the task completion time. Moreover, the drop-for-partner interaction is a simple example of a shared interaction technique tailored to the experimental task and pick-and-drop interaction between pairs. Other techniques need to be designed for different tasks and collaboration scenarios, e.g., more users with different roles or levels of expertise. Multi-selection, grouping or editing could be tested, as they may lead to different trade-offs and collaboration styles. Also, remote collaboration could evaluated with a similar approach.

Another future direction of this work is to build models for such a classification task in the single-user and collaborative cases. Building a mathematical model reflects a deep understanding of a phenomenon as it explains the relations between all its parameters. For instance, the performance in time of a single-user classification task is a function of several parameters: pick-and-drop, search and (spatial) memory. Their relations are not simply additive or subtractive because they happen at the same time. Two users performing this task might take half of the time of one user. Increasing the number of users would introduce a cost, due to coordination, disruption and other reasons, while introducing shared interaction techniques would introduce a benefit. Further investigations could lead to predictive models reflecting the understanding of these phenomena and their design implications.

9.2.2 *From Understanding to Innovation*

The single-user experiments increased our understanding of the benefits of wall displays for manipulating large data sets. The first experiment showed an increasing performance gap between the wall display and the desktop for higher information density and more difficult tasks. The analysis attributed the major reason to the fact that users navigate data with their body movements and that walking increases their interaction area with data. We also found that this effect could not be achieved by using state-of-the-art navigation techniques on desktop computers for tasks involving manipulation. Our understanding of this effect may inspire and inform novel interaction techniques for the desktop, which might simulate a large display effect.

In collaborative situations, another controlled experiment showed the benefits of providing a shared interaction technique, including encouraging collaboration, reducing fatigue and improving interaction efficiency. Such techniques have not been explored in wall-display environments. These findings led to the design of shared interaction techniques, which support and blend in with users' collaborative behavior, for arranging data items on a multi-touch wall-sized display. I conducted an observational study and uncovered some frequent collaborative operations, including showing data to each other, passing data and queuing tasks by holding data in hands, etc. I designed and implemented Collaborative Gestures - a set of novel interaction techniques to assist collaborative operations while reducing the interaction effort for large-scale manipulation. I also introduced PoPle - an interaction technique that allows users to exchange data asynchronously using a "queue" mechanism which enables easy access for users to their data while moving around.

Future work should improve the design of Collaborative Gestures and PoPle and extend the design space. It includes overcoming the design and technical challenges to support more users and multi-party collaboration.

Embodiment in social context shall be explored further. This part of the work in this dissertation is only a start point. Inspirations can be taken from here and applied to other co-located environments than with a wall display, as mentioned in Section 8.4. This line will be continued in my future research.

[1] Johnny Accot and Shumin Zhai. Performance evaluation of input devices in trajectory-based tasks: An application of the steering law. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '99, pages 466–472. ACM, 1999. http://doi.acm.org/10.1145/302979.303133.

[2] Paul André, Robert E. Kraut, and Aniket Kittur. Effects of simultaneous and sequential work structures on distributed collaborative interdependent tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pages 139–148. ACM, 2014. http://doi.acm.org/10.1145/2556288.2557158.

[3] Christopher Andrews, Alex Endert, and Chris North. Space to think: large high-resolution displays for sensemaking. In *Proceedings of the 28th international conference on Human factors in computing systems*, CHI '10, pages 55–64. ACM, 2010. http://doi.acm.org/10.1145/1753326.1753336.

[4] Alec Azad, Jaime Ruiz, Daniel Vogel, Mark Hancock, and Edward Lank. Territoriality and behaviour on and around large vertical publicly-shared displays. In *Proceedings of the Designing Interactive Systems Conference*, DIS '12, pages 468–477. ACM, 2012. http://doi.acm.org/10.1145/2317956.2318025.

[5] R. Bakeman. Recommended effect size statistics for repeated measures designs. *Behav. Res. Meth.*, 37(3):379–384, 2005.

[6] Robert Ball and Chris North. The effects of peripheral vision and physical navigation on large scale visualization. In *Proceedings of graphics interface 2008*, GI '08, pages 9–16. CIPS, 2008. http://dl.acm.org/citation.cfm?id=1375714.1375717.

[7] Robert Ball, Chris North, and Doug A. Bowman. Move to improve: promoting physical navigation to increase user performance with large displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '07, pages 191–200. ACM, 2007. http://doi.acm.org/10.1145/1240624.1240656.

[8] Olivier Bau, Ivan Poupyrev, Ali Israr, and Chris Harrison. Teslatouch: Electrovibration for touch surfaces. In *Proceedings of the 23Nd Annual ACM Symposium on User Interface Software and Technology*, UIST '10, pages 283–292. ACM, 2010. http://doi.acm.org/10.1145/1866029.1866074.

[9] Patrick Baudisch and Ruth Rosenholtz. Halo: A technique for visualizing off-screen objects. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '03, pages 481–488. ACM, 2003. http://doi.acm.org/10.1145/642611.642695.

[10] M. Beaudouin-Lafon, O. Chapuis, J. Eagan, T. Gjerlufsen, S. Huot, C. Klokmose, W. Mackay, M. Nancel, E. Pietriga, C. Pillias, R. Primet, and J. Wagner. Multi-surface interaction in the WILD room. *IEEE Computer*, 45(4):48–56, 2012. http://dx.doi.org/10.1109/MC.2012.110.

[11] Anastasia Bezerianos and Ravin Balakrishnan. The vacuum: facilitating the manipulation of distant objects. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '05, pages 361–370. ACM, 2005. http://doi.acm.org/10.1145/1054972.1055023.

[12] Anastasia Bezerianos and Petra Isenberg. Perception of visual variables on tiled wall-sized displays for information visualization applications. *IEEE TVCG*, 18(12):2516–2525, 2012.

[13] Xiaojun Bi and Ravin Balakrishnan. Comparing usage of a large high-resolution display to single or dual desktop displays for daily work. In *Proceedings of the 27th international conference on Human factors in computing systems*, CHI '09, pages 1005–1014. ACM, 2009. http://doi.acm.org/10.1145/1518701.1518855.

[14] Norman Biggs. *Algebraic graph theory*. Cambridge university press, 1993.

[15] Renaud Blanch and Eric Lecolinet. Browsing zoomable treemaps: Structure-aware multi-scale navigation techniques. *IEEE Transactions on Visualization and Computer Graphics*, 13(6): 1248–1253, November 2007. http://dx.doi.org/10.1109/TVCG.2007.70540.

[16] J. Borchers, M. Ringel, J. Tyler, and A. Fox. Stanford interactive workspaces: A framework for physical and graphical user interface prototyping. *Wireless Commun.*, 9(6):64–69, December 2002. http://dx.doi.org/10.1109/MWC.2002.1160083.

[17] Lauren Bradel, Alex Endert, Kristen Koch, Christopher Andrews, and Chris North. Large high resolution displays for co-located collaborative sensemaking: Display usage and territoriality. *International Journal of Human-Computer Studies*, 71(11): 1078–1088, 2013.

[18] British Standards Institution. Test charts for clinical determination of distance visual acuity, 2003. London.

[19] Stuart K. Card, Thomas P. Moran, and Allen Newell. The keystroke-level model for user performance time with interactive systems. *Commun. ACM*, 23(7):396–410, July 1980. http://doi.acm.org/10.1145/358886.358895.

[20] Olivier Chapuis, Renaud Blanch, and Michel Beaudouin-Lafon. Fitts' Law in the Wild: A Field Study of Aimed Movements. Technical Report 1480, LRI, Univ Paris-Sud, 11 pages, December 2007. https://hal.archives-ouvertes.fr/hal-00612026.

[21] Lung-Pan Cheng, Patrick Lühne, Pedro Lopes, Christoph Sterz, and Patrick Baudisch. Haptic turk: A motion platform based on people. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pages 3463–3472. ACM, 2014. http://doi.acm.org/10.1145/2556288.2557101.

[22] Andy Cockburn, Amy Karlson, and Benjamin B. Bederson. A review of overview+detail, zooming, and focus+context interfaces. *ACM Comput. Surv.*, 41(1):2:1–2:31, January 2009. http://doi.acm.org/10.1145/1456650.1456652.

[23] A. Cohe, B. Liutkus, G. Bailly, J. Eagan, and E. Lecolinet. Schemelens: A content-aware vector-based fisheye technique for navigating large systems diagrams. *Visualization and Computer Graphics, IEEE Transactions on*, PP(99):1–1, 2015.

[24] J. Cohen. Eta-squared and partial eta-squared in fixed factor anova designs. *Educational and Psychological Measurement*, 33(1): 107–112, 1973.

[25] Maxime Collomb, Mountaz Hascoët, Patrick Baudisch, and Brian Lee. Improving drag-and-drop on wall-size displays. In *Proceedings of Graphics Interface 2005*, GI '05, pages 25–32. CHCCS, 2005. http://dl.acm.org/citation.cfm?id=1089508.1089514.

[26] Mary Czerwinski, Greg Smith, Tim Regan, Brian Meyers, George Robertson, and Gary Starkweather. Toward characterizing the productivity benefits of very large displays. In *Proc. Interact*, INTERACT '03, pages 9–16. IOS, 2003.

[27] Connor Dickie, Roel Vertegaal, Jeffrey S. Shell, Changuk Sohn, Daniel Cheng, and Omar Aoudeh. Eye contact sensing glasses for attention-sensitive wearable video blogging. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '04, pages 769–770. ACM, 2004. http://doi.acm.org/10.1145/985921.985927.

[28] Paul Dietz and Darren Leigh. Diamondtouch: A multi-user touch technology. In *Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology*, UIST '01, pages

219–226. ACM, 2001. http://doi.acm.org/10.1145/502348.502389.

[29] Paul Dourish. *Where the Action is: The Foundations of Embodied Interaction*. MIT Press, Cambridge, MA, USA, 2001.

[30] Paul Dourish. Seeking a foundation for context-aware computing. *Hum.-Comput. Interact.*, 16(2):229–241, December 2001. http://dx.doi.org/10.1207/S15327051HCI16234_07.

[31] Paul Dourish and Victoria Bellotti. Awareness and coordination in shared workspaces. In *Proceedings of the 1992 ACM Conference on Computer-supported Cooperative Work*, CSCW '92, pages 107–114. ACM, 1992. http://doi.acm.org/10.1145/143457.143468.

[32] Olive Jean Dunn. Estimation of the medians for dependent variables. *Ann. Math. Statist.*, 30(1):192–197, 03 1959. http://dx.doi.org/10.1214/aoms/1177706374.

[33] T Edward. Hall, the hidden dimension, 1966.

[34] Niklas Elmqvist and Jean-Daniel Fekete. Hierarchical aggregation for information visualization: Overview, techniques, and design guidelines. *IEEE Transactions on Visualization and Computer Graphics*, 16(3):439–454, May 2010. http://dx.doi.org/10.1109/TVCG.2009.84.

[35] Alex Endert, Christopher Andrews, Yueh Hua Lee, and Chris North. Visual encodings that support physical navigation on large displays. In *Proceedings of Graphics Interface 2011*, GI '11, pages 103–110. CHCCS, 2011. http://dl.acm.org/citation.cfm?id=1992917.1992935.

[36] Jean-Daniel Fekete and Catherine Plaisant. Interactive information visualization of a million items. In *Proceedings of the IEEE Symposium on Information Visualization (InfoVis'02)*, INFOVIS '02, pages 117–. IEEE Computer Society, 2002. http://dl.acm.org/citation.cfm?id=857191.857732.

[37] Paul M Fitts. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology: General*, 121(3):262, 1992.

[38] Michael Fitze. Discourse and participation in esl face-to-face and written electronic conferences. *Language Learning & Technology*, 10(1):67–86, 2006.

[39] George W. Fitzmaurice, Hiroshi Ishii, and William A. S. Buxton. Bricks: Laying the foundations for graspable user interfaces. In

Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '95, pages 442–449. ACM Press/Addison-Wesley Publishing Co., 1995. http://dx.doi.org/10.1145/223904.223964.

[40] John C Flanagan. The critical incident technique. *Psychological bulletin*, 51(4):327, 1954.

[41] David Frohlich and Rachel Murphy. The memory box. *Personal Ubiquitous Comput.*, 4(4):238–240, January 2000. http://dx.doi.org/10.1007/PL00000011.

[42] Saul Greenberg, Nicolai Marquardt, Till Ballendat, Rob Diaz-Marino, and Miaosen Wang. Proxemic interactions: The new ubicomp? *interactions*, 18(1):42–50, January 2011. http://doi.acm.org/10.1145/1897239.1897250.

[43] Yves Guiard, Michel Beaudouin-Lafon, Julien Bastin, Dennis Pasveer, and Shumin Zhai. View size and pointing difficulty in multi-scale navigation. In *Proceedings of the working conference on Advanced visual interfaces*, AVI '04, pages 117–124. ACM, 2004. http://doi.acm.org/10.1145/989863.989881.

[44] Carl Gutwin and Saul Greenberg. The effects of workspace awareness support on the usability of real-time distributed groupware. *ACM Transactions on Computer-Human Interaction*, 6(3):243–281, September 1999. http://doi.acm.org/10.1145/329693.329696.

[45] Carl Gutwin and Saul Greenberg. A descriptive framework of workspace awareness for real-time groupware. *Computer Supported Cooperative Worr*, 11(3-4):411–446, 2002.

[46] Chris Harrison and Anind K. Dey. Lean and zoom: Proximity-aware user interface and content magnification. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 507–510. ACM, 2008. http://doi.acm.org/10.1145/1357054.1357135.

[47] Mountaz Hascoët. Throwing models for large displays. In *Proceedings of the 11th International Conference on Human Computer Interaction*, HCI'03, pages 73–77. British HCI Group, April 2008. http://hal-lirmm.ccsd.cnrs.fr/lirmm-00269751.

[48] Kirstie Hawkey, Melanie Kellar, Derek Reilly, Tara Whalen, and Kori M. Inkpen. The proximity factor: Impact of distance on co-located collaboration. In *Proceedings of the 2005 International ACM SIGGROUP Conference on Supporting Group Work*, GROUP '05, pages 31–40. ACM, 2005. http://doi.acm.org/10.1145/1099203.1099209.

[49] Steven J. Henderson and Steven K. Feiner. Augmented reality in the psychomotor phase of a procedural task. In *Proceedings of the 2011 10th IEEE International Symposium on Mixed and Augmented Reality*, ISMAR '11, pages 191–200. IEEE Computer Society, 2011. http://dx.doi.org/10.1109/ISMAR.2011.6092386.

[50] Juan David Hincapié-Ramos, Xiang Guo, Paymahn Moghadasian, and Pourang Irani. Consumed endurance: A metric to quantify arm fatigue of mid-air interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pages 1063–1072. ACM, 2014. http://doi.acm.org/10.1145/2556288.2557130.

[51] Kasper Hornbæk, Benjamin B. Bederson, and Catherine Plaisant. Navigation patterns and usability of zoomable user interfaces with and without an overview. *ACM ToCHI*, 9(4):362–389, December 2002. http://doi.acm.org/10.1145/586081.586086.

[52] Eva Hornecker, Paul Marshall, Nick Sheep Dalton, and Yvonne Rogers. Collaboration and interference: Awareness with mice or touch input. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work*, CSCW '08, pages 167–176. ACM, 2008. http://doi.acm.org/10.1145/1460563.1460589.

[53] P. Isenberg, D. Fisher, S.A Paul, M.R. Morris, K. Inkpen, and Mary Czerwinski. Co-located collaborative visual analytics around a tabletop display. *IEEE Transactions on Visualization and Computer Graphics*, 18(5):689–702, May 2012.

[54] Hiroshi Ishii and Brygg Ullmer. Tangible bits: Towards seamless interfaces between people, bits and atoms. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, CHI '97, pages 234–241. ACM, 1997. http://doi.acm.org/10.1145/258549.258715.

[55] Hiroo Iwata, Hiroaki Yano, Fumitaka Nakaizumi, and Ryo Kawamura. Project feelex: Adding haptic surface to graphics. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '01, pages 469–476. ACM, 2001. http://doi.acm.org/10.1145/383259.383314.

[56] Robert J.K. Jacob, Audrey Girouard, Leanne M. Hirshfield, Michael S. Horn, Orit Shaer, Erin Treacy Solovey, and Jamie Zigelbaum. Reality-based interaction: A framework for post-wimp interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 201–210. ACM, 2008. http://doi.acm.org/10.1145/1357054.1357089.

[57] Mikkel R. Jakobsen and Kasper Hornbaek. Up close and personal: Collaborative work on a high-resolution multitouch wall display. *ACM Trans. Comput.-Hum. Interact.*, 21(2):11:1–11:34, February 2014. http://doi.acm.org/10.1145/2576099.

[58] Mikkel R. Jakobsen and Kasper Hornbæk. Is moving improving?: Some effects of locomotion in wall-display interaction. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 4169–4178. ACM, 2015. http://doi.acm.org/10.1145/2702123.2702312.

[59] Rønne Mikkel Jakobsen and Kasper Hornbæk. Sizing up visualizations: effects of display size in focus+context, overview+detail, and zooming interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 1451–1460. ACM, 2011. http://doi.acm.org/10.1145/1978942.1979156.

[60] Y. Jansen and P. Dragicevic. An interaction model for visualizations beyond the desktop. *Visualization and Computer Graphics, IEEE Transactions on*, 19(12):2396–2405, Dec 2013.

[61] Sergi Jordà, Günter Geiger, Marcos Alonso, and Martin Kaltenbrunner. The reactable: Exploring the synergy between live music performance and tabletop tangible interfaces. In *Proceedings of the 1st International Conference on Tangible and Embedded Interaction*, TEI '07, pages 139–146. ACM, 2007. http://doi.acm.org/10.1145/1226969.1226998.

[62] Jeevan J. Kalanithi and V. Michael Bove, Jr. Connectibles: Tangible social networks. In *Proceedings of the 2Nd International Conference on Tangible and Embedded Interaction*, TEI '08, pages 199–206. ACM, 2008. http://doi.acm.org/10.1145/1347390.1347434.

[63] Azam Khan, George Fitzmaurice, Don Almeida, Nicolas Burtnyk, and Gordon Kurtenbach. A remote control interface for large displays. In *Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology*, UIST '04, pages 127–136. ACM, 2004. http://doi.acm.org/10.1145/1029632.1029655.

[64] Juho Kim, Haoqi Zhang, Paul André, Lydia B. Chilton, Wendy Mackay, Michel Beaudouin-Lafon, Robert C. Miller, and Steven P. Dow. Cobi: A community-informed conference scheduling tool. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology*, UIST '13, pages 173–182. ACM, 2013. http://doi.acm.org/10.1145/2501988.2502034.

[65] Scott R. Klemmer, Björn Hartmann, and Leila Takayama. How bodies matter: Five themes for interaction design. In *Proceedings of the 6th Conference on Designing Interactive Systems*, DIS '06, pages 140–149. ACM, 2006. http://doi.acm.org/10.1145/1142405.1142429.

[66] Can Liu, Stephane Huot, Jonathan Diehl, Wendy Mackay, and Michel Beaudouin-Lafon. Evaluating the benefits of real-time feedback in mobile augmented reality with hand-held devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 2973–2976. ACM, 2012. http://doi.acm.org/10.1145/2207676.2208706.

[67] Can Liu, Olivier Chapuis, Michel Beaudouin-Lafon, Eric Lecolinet, and Wendy Mackay. Effects of display size and navigation type on a classification task. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pages 4147–4156. ACM, 2014. http://doi.acm.org/10.1145/2556288.2557020.

[68] Andrés Lucero, Jussi Holopainen, and Tero Jokela. Pass-them-around: Collaborative use of mobile phones for photo sharing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 1787–1796. ACM, 2011. http://doi.acm.org/10.1145/1978942.1979201.

[69] I. Scott MacKenzie, Tatu Kauppinen, and Miika Silfverberg. Accuracy measures for evaluating computer pointing devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '01, pages 9–16. ACM, 2001. http://doi.acm.org/10.1145/365024.365028.

[70] Nicolai Marquardt, Ken Hinckley, and Saul Greenberg. Cross-device interaction via micro-mobility and f-formations. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology*, UIST '12, pages 13–22. ACM, 2012. http://doi.acm.org/10.1145/2380116.2380121.

[71] Leslie Marsh. Michael Wheeler: Reconstructing the cognitive world: The next step. *Phenomenology and the Cognitive Sciences*, 7(1):147–149, 2008.

[72] Paul Marshall, Alissa Nicole Antle, Elise van den Hoven, and Yvonne Rogers. Introduction to the special issue on the theory and practice of embodied interaction in HCI and interaction design. *ACM Trans. Comput.-Hum. Interact.*, 20(1):1, 2013. http://doi.acm.org/10.1145/2442106.2442107.

[73] P. Milgram and F. Kishino. A taxonomy of mixed reality visual displays. *IEICE Trans. Information Systems*, E77-D(12):1321–1329, dec 1994.

[74] Pranav Mistry and Pattie Maes. Sixthsense: A wearable gestural interface. In *ACM SIGGRAPH ASIA 2009 Sketches*, SIGGRAPH ASIA '09, pages 11:1–11:1. ACM, 2009. http://doi.acm.org/10.1145/1667146.1667160.

[75] Pranav Mistry, Suranga Nanayakkara, and Pattie Maes. Sparsh: Passing data using the body as a medium. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work*, CSCW '11, pages 689–692. ACM, 2011. http://doi.acm.org/10.1145/1958824.1958946.

[76] Meredith Ringel Morris, Anthony Cassanego, Andreas Paepcke, Terry Winograd, Anne Marie Piper, and Anqi Huang. Mediating group dynamics through tabletop interface design. *IEEE Computer Graphics and Applications*, 26(5):65–73, September 2006. http://ilpubs.stanford.edu:8090/787/.

[77] Meredith Ringel Morris, Anqi Huang, Andreas Paepcke, and Terry Winograd. Cooperative gestures: Multi-user gestural interactions for co-located groupware. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '06, pages 1201–1210. ACM, 2006. http://doi.acm.org/10.1145/1124772.1124952.

[78] Tomer Moscovich, Fanny Chevalier, Nathalie Henry, Emmanuel Pietriga, and Jean-Daniel Fekete. Topology-aware navigation in large networks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 2319–2328. ACM, 2009. http://doi.acm.org/10.1145/1518701.1519056.

[79] Jörg Müller, Florian Alt, Daniel Michelis, and Albrecht Schmidt. Requirements and design space for interactive public displays. In *Proceedings of the International Conference on Multimedia*, MM '10, pages 1285–1294. ACM, 2010. http://doi.acm.org.gate6.inist.fr/10.1145/1873951.1874203.

[80] Miguel A. Nacenta, David Pinelle, Dane Stuckel, and Carl Gutwin. The effects of interaction technique on coordination in tabletop groupware. In *Proceedings of Graphics Interface 2007*, GI '07, pages 191–198. ACM, 2007. http://doi.acm.org/10.1145/1268517.1268550.

[81] Mathieu Nancel, Julie Wagner, Emmanuel Pietriga, Olivier Chapuis, and Wendy Mackay. Mid-air pan-and-zoom on wall-sized

displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 177–186. ACM, 2011. http://doi.acm.org/10.1145/1978942.1978969.

[82] Mathieu Nancel, Olivier Chapuis, Emmanuel Pietriga, Xing-Dong Yang, Pourang Irani, and Michel Beaudouin-Lafon. High-precision pointing on large wall displays using small handheld devices. In *Proceedings of the 31th international conference on Human factors in computing systems*, CHI '13, pages 831–840. ACM, 2013. http://doi.acm.org/10.1145/2470654.2470773.

[83] Mathieu Nancel, Emmanuel Pietriga, Olivier Chapuis, and Michel Beaudouin-Lafon. Mid-air pointing on ultra-walls. *ACM Trans. Comput.-Hum. Interact.*, 22(5):21:1–21:62, August 2015. http://doi.acm.org/10.1145/2766448.

[84] Dmitry Nekrasovski, Adam Bodnar, Joanna McGrenere, François Guimbretière, and Tamara Munzner. An evaluation of pan & zoom and rubber sheet navigation with and without an overview. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '06, pages 11–20. ACM, 2006. http://doi.acm.org/10.1145/1124772.1124775.

[85] Tao Ni, Doug Bowman, and Jian Chen. Increased display size and resolution improve task performance in information-rich virtual environments. In *Proceedings of Graphics Interface 2006*, GI '06, pages 139–146. CIPS, 2006. http://dl.acm.org/citation.cfm?id=1143079.1143102.

[86] J. Karen Parker, Regan L. Mandryk, and Kori M. Inkpen. Tractorbeam: Seamless integration of local and remote pointing for tabletop displays. In *Proceedings of Graphics Interface 2005*, GI '05, pages 33–40. Canadian Human-Computer Communications Society, 2005. http://dl.acm.org/citation.cfm?id=1089508.1089515.

[87] Emmanuel Pietriga, Caroline Appert, and Michel Beaudouin-Lafon. Pointing and beyond: an operationalization and preliminary evaluation of multi-scale searching. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '07, pages 1215–1224. ACM, 2007. http://doi.acm.org/10.1145/1240624.1240808.

[88] Emmanuel Pietriga, Stéphane Huot, Mathieu Nancel, and Romain Primet. Rapid development of user interfaces on cluster-driven wall displays with jbricks. In *Proceedings of the 3rd ACM SIGCHI symposium on Engineering interactive computing systems*, EICS '11, pages 185–190. ACM, 2011. http://doi.acm.org/10.1145/1996461.1996518.

[89] Cyprien Pindat, Emmanuel Pietriga, Olivier Chapuis, and Claude Puech. Jellylens: Content-aware adaptive lenses. In *Proceedings of the 25th annual ACM symposium on User interface software and technology*, UIST '12, pages 261–270. ACM, October 2012. http://doi.acm.org/10.1145/2380116.2380150.

[90] David Pinelle, Carl Gutwin, and Saul Greenberg. Task analysis for groupware usability evaluation: Modeling shared-workspace tasks with the mechanics of collaboration. *ACM Transactions on Computer-Human Interaction*, 10(4):281–311, 2003.

[91] Michael Polanyi. *The Tacit Dimension*. Doubleday, New York, 1967.

[92] Gonzalo Ramos, Kenneth Hinckley, Andy Wilson, and Raman Sarin. Synchronous gestures in multi-display environments. *Human–Computer Interaction*, 24(1-2):117–169, 2009. http://www.tandfonline.com/doi/abs/10.1080/07370020902739288.

[93] Gonzalo Ramos, Kenneth Hinckley, Andy Wilson, and Raman Sarin. Synchronous gestures in multi-display environments. *Human–Computer Interaction*, 24(1-2):117–169, 2009.

[94] Hanae Rateau, Laurent Grisoni, and Bruno De Araujo. Mimetic interaction spaces: Controlling distant displays in pervasive environments. In *Proceedings of the 19th International Conference on Intelligent User Interfaces*, IUI '14, pages 89–94. ACM, 2014. http://doi.acm.org.gate6.inist.fr/10.1145/2557500.2557545.

[95] Khairi Reda, Andrew E. Johnson, Michael E. Papka, and Jason Leigh. Effects of display size and resolution on user behavior and insight acquisition in visual exploration. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 2759–2768. ACM, 2015. http://doi.acm.org/10.1145/2702123.2702406.

[96] 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, pages 31–39. ACM, 1997. http://doi.acm.org/10.1145/263407.263505.

[97] 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, pages 31–39. ACM, 1997. http://doi.acm.org/10.1145/263407.263505.

[98] Chris Rooney, Alex Endert, Jean-Daniel Fekete, Kasper Hornbæk, and Chris North. Powerwall: int. workshop on interactive,

ultra-high-resolution displays. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '13, pages 3227–3230, 2013. http://doi.acm.org/10.1145/2468356.2479653.

[99] Daniela K. Rosner and Kimiko Ryokai. Spyn: Augmenting the creative and communicative potential of craft. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 2407–2416. ACM, 2010. http://doi.acm.org/10.1145/1753326.1753691.

[100] Kathy Ryall, Alan Esenther, Clifton Forlines, Chia Shen, Sam Shipman, Meredith Ringel Morris, Katherine Everitt, and Frederic D. Vernier. Identity-differentiating widgets for multiuser interactive surfaces. *IEEE Comput. Graph. Appl.*, 26(5):56–64, September 2006. http://dx.doi.org/10.1109/MCG.2006.108.

[101] Kimiko Ryokai, Stefan Marti, and Hiroshi Ishii. I/o brush: Drawing with everyday objects as ink. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '04, pages 303–310. ACM, 2004. http://doi.acm.org/10.1145/985692.985731.

[102] Manojit Sarkar and Marc H. Brown. Graphical fisheye views of graphs. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '92, pages 83–91. ACM, 1992. http://doi.acm.org/10.1145/142750.142763.

[103] Stacey D. Scott, M. Sheelagh T. Carpendale, and Kori M. Inkpen. Territoriality in collaborative tabletop workspaces. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*, CSCW '04, pages 294–303. ACM, 2004. http://doi.acm.org/10.1145/1031607.1031655.

[104] Garth Shoemaker, Anthony Tang, and Kellogg S. Booth. Shadow reaching: A new perspective on interaction for large displays. In *Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology*, UIST '07, pages 53–56. ACM, 2007. http://doi.acm.org/10.1145/1294211.1294221.

[105] Randall B. Smith. Experiences with the alternate reality kit: An example of the tension between literalism and magic. *SIGCHI Bull.*, 18(4):61–67, May 1986. http://doi.acm.org/10.1145/1165387.30861.

[106] Daniel Spelmezan, Mareike Jacobs, Anke Hilgers, and Jan Borchers. Tactile motion instructions for physical activities. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 2243–2252. ACM, 2009. http://doi.acm.org/10.1145/1518701.1519044.

[107] Desney S. Tan, Darren Gergle, Peter Scupelli, and Randy Pausch. Physically large displays improve performance on spatial tasks. *ACM ToCHI*, 13(1):71–99, March 2006. http://doi.acm.org/10.1145/1143518.1143521.

[108] Desney S Tan, Darren Gergle, Regan Mandryk, Kori Inkpen, Melanie Kellar, Kirstie Hawkey, and Mary Czerwinski. Using job-shop scheduling tasks for evaluating collocated collaboration. *Personal and Ubiquitous Computing*, 12(3):255–267, 2008. http://research.microsoft.com/apps/pubs/default.aspx?id=101716.

[109] Anthony Tang, Melanie Tory, Barry Po, Petra Neumann, and Sheelagh Carpendale. Collaborative coupling over tabletop displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '06, pages 1181–1190. ACM, 2006. http://doi.acm.org/10.1145/1124772.1124950.

[110] Theophanis Tsandilas, Catherine Letondal, and Wendy E. Mackay. Musink: Composing music through augmented drawing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 819–828. ACM, 2009. http://doi.acm.org/10.1145/1518701.1518827.

[111] Edward Tse, Jonathan Histon, Stacey D. Scott, and Saul Greenberg. Avoiding interference: How people use spatial separation and partitioning in SDG workspaces. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*, CSCW '04, pages 252–261. ACM, 2004. http://doi.acm.org/10.1145/1031607.1031647.

[112] Simon Voelker, Malte Weiss, Chat Wacharamanotham, and Jan Borchers. Dynamic portals: A lightweight metaphor for fast object transfer on interactive surfaces. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, ITS '11, pages 158–161. ACM, 2011. http://doi.acm.org/10.1145/2076354.2076384.

[113] Mark Weiser. The computer for the 21st century. *SIGMOBILE Mob. Comput. Commun. Rev.*, 3(3):3–11, July 1999. http://doi.acm.org/10.1145/329124.329126.

[114] Malte Weiss, Julie Wagner, Yvonne Jansen, Roger Jennings, Ramsin Khoshabeh, James D. Hollan, and Jan Borchers. Slap widgets: Bridging the gap between virtual and physical controls on tabletops. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 481–490. ACM, 2009. http://doi.acm.org/10.1145/1518701.1518779.

[115] Beth Yost, Yonca Haciahmetoglu, and Chris North. Beyond visual acuity: the perceptual scalability of information visualizations for large displays. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '07, pages 101–110. ACM, 2007. http://doi.acm.org/10.1145/1240624.1240639.

[116] Nicola Yuill and Yvonne Rogers. Mechanisms for collaboration: A design and evaluation framework for multi-user interfaces. *ACM Transactions on Computer-Human Interaction*, 19(1):1:1–1:25, May 2012. http://doi.acm.org/10.1145/2147783.2147784.

[117] Nicola Yuill and Yvonne Rogers. Mechanisms for collaboration: A design and evaluation framework for multi-user interfaces. *ACM Trans. Comput.-Hum. Interact.*, 19(1):1:1–1:25, May 2012. http://doi.acm.org/10.1145/2147783.2147784.

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede.

January 5, 2016