Tero Jokela
**Studies on Multi-Device Usage Practices and Interaction Methods**

Tero Jokela

# Studies on Multi-Device Usage Practices and Interaction Methods

Thesis for the degree of Doctor of Science in Technology to be presented with due permission for public examination and criticism in Tietotalo Building, Auditorium TB109, at Tampere University of Technology, on the 30th of November 2018, at 12 noon.

Doctoral candidate:    Tero Jokela
                       User Experience Team
                       BHAG Realization Lab
                       Nokia Bell Labs
                       Finland

Supervisor:            Professor Kaisa Väänänen
                       Unit of Human-Centered Technology (IHTE)
                       Laboratory of Pervasive Computing
                       Tampere University of Technology
                       Finland

Pre-examiners:         Research Director Fabio Paternò
                       Human Interfaces in Information Systems Laboratory
                       Institute of Information Science and Technology (ISTI)
                       National Research Council (CNR)
                       Italy

                       Associate Professor Nicolai Marquardt
                       UCL Interaction Centre
                       Department of Computer Science
                       University College London
                       United Kingdom

Opponent:              Professor Enrico Rukzio
                       Human-Computer Interaction Group
                       Institute of Media Informatics
                       Ulm University
                       Germany

# Abstract

People today commonly have multiple information devices, including smartphones, tablets, computers, home media centers, and other devices. As people have many devices, situations and workflows where several devices are combined and used together to accomplish a task have become usual. Groups of co-located persons may also join their information devices together for collaborative activities and experiences. While these developments towards computing with multiple devices offer many opportunities, they also create a need for interfaces and applications that support using multiple devices together.

The overall goal of this doctoral thesis is to create new scientific knowledge to inform the design of future interfaces, applications, and technologies that better support multi-device use. The thesis belongs to the field of Human-Computer Interaction (HCI) research. It contains five empirical studies with a total of 110 participants. The study results have been reported in five original publications. The thesis generally follows the design science research methodology.

More specifically, this thesis addresses three research questions related to multi-device use. The first question investigates how people actually use multiple information devices together in their daily lives. The results provide a rich picture of everyday multi-device use, including the most common devices and their characteristic practices of use, a categorization of patterns of multi-device use, and an analysis of the process of determining which devices to use. The second question examines the factors that influence the user experience of multi-device interaction methods. The results suggest a set of experiential factors that should be considered when designing methods for multi-device interaction. The set of factors is based on comparative studies of alternative methods for two common tasks in multi-device interaction: device binding and cross-display object movement. The third question explores a more futuristic topic of multi-device interaction methods for wearable devices, focusing on the two most popular categories of wearable devices today: smartwatches and smartglasses. The results present a categorization of actions that people would naturally do to initiate interactions between their wearable devices based on elicitation studies with groups of participants.

The results of this thesis advance the scientific knowledge of multi-device use in the domain of human-computer interaction research. The results can be applied in the design of novel interfaces, applications, and technologies that involve the use of multiple information devices.

ii

# Preface

The work presented in this thesis was started at Nokia Research Center in 2012–14, was continued at Nokia Technologies in 2015–18, and was finally completed at Nokia Bell Labs in 2018. I would like to thank my laboratory directors Jyri Huopaniemi, Ville-Veikko Mattila, and Thierry Klein and my team leaders Vilja-Kaisa Aaltonen, Marja Salmimaa, and Arto Lehtiniemi for providing me the opportunity to work on this thesis.

I would like to express my sincere gratitude to my supervisor Professor Kaisa Väänänen for her guidance, encouragement, and patience during the course of this thesis work. I would like to thank my pre-examiners Research Director Fabio Paternò and Associate Professor Nicolai Marquardt for their careful reading of the manuscript and for their thoughtful comments. I am honored to have Professor Enrico Rukzio as my opponent in the public examination.

The publications presented in this thesis have been co-authored with Andrés Lucero, Jarno Ojala, Thomas Olsson, Guido Grassel, Petri Piippo, and Parisa Pour Rezaei. Many thanks are due to all of them. I would also like to thank the numerous other people who have helped me in the work presented in this thesis, including Arto Palin, Sampo Vesa, Johan Kildal, Akos Vetek, Jussi Holopainen, Severi Uusitalo, Jari Nikara, Peter Eskolin, Jyrki Kimmel, Toni Järvenpää, Jukka Saarinen, Hilkka Losoi, Juha Riippi, Iiro Vidberg, Arttu Pulli, Markus Rinne, Mikko Tolonen, Donald McMillan, Susan Fussell, Ming Ki Chong, and Hans Gellersen.

Finally, I would like to thank my parents for their continuous support and encouragement through the years. Most of all, I would like to thank my family, my dear wife Sanna and our sweet daughters Venla, Senni, and Iina, for their support, love, and understanding during my thesis work.

Tampere, October 17, 2018                    Tero Jokela

iv

# Contents

# List of Publications

This thesis consists of an introductory part and the following five original publications that are referred to in the text as P1–P5. The publications are reprinted with permission from the publisher.

**P1**   Tero Jokela, Jarno Ojala, and Thomas Olsson. 2015. A Diary Study on Combining Multiple Information Devices in Everyday Activities and Tasks. In *Proceedings of the 33rd SIGCHI Conference on Human Factors in Computing Systems (CHI '15)*, 3903–3912. DOI: https://doi.org/10.1145/2702123.2702211

**P2**   Tero Jokela and Andrés Lucero. 2013. A Comparative Evaluation of Touch-Based Methods to Bind Mobile Devices for Collaborative Interactions. In *Proceedings of the 31st SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*, 3355–3364. DOI: https://doi.org/10.1145/2470654.2466459

**P3**   Tero Jokela and Andrés Lucero. 2014. FlexiGroups: Binding Mobile Devices for Collaborative Interactions in Medium-Sized Groups with Device Touch. In *Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '14)*, 369–378. DOI: https://doi.org/10.1145/2628363.2628376

**P4**   Tero Jokela, Jarno Ojala, Guido Grassel, Petri Piippo, and Thomas Olsson. 2015. A Comparison of Methods to Move Visual Objects Between Personal Mobile Devices in Different Contexts of Use. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '15)*, 172–181. DOI: https://doi.org/10.1145/2785830.2785841

**P5**   Tero Jokela, Parisa Pour Rezaei, and Kaisa Väänänen. 2016. Natural Group Binding and Cross-Display Object Movement Methods for Wearable Devices. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '16)*, 206–216. DOI: https://doi.org/10.1145/2935334.2935346

# Author's Contribution to the Publications

This page summarizes the author's contribution to the publications of this thesis.

**P1**  The publication reports a user study on people's current practices of multi-device use. The author planned the study and collected and analyzed the data together with Jarno Ojala. Thomas Olsson also contributed to the data analysis. The author was the principal author of the publication.

**P2**  The publication presents a laboratory study that compares three different group binding methods. The author designed the study and the binding methods together with Andrés Lucero. The author conducted the experiments and collected the data. The author analyzed the data together with Andrés Lucero. The author was the principal author of the publication.

**P3**  The publication suggests a novel group binding method and evaluates it with a laboratory study. The author designed the binding method and the study together with Andrés Lucero. The author conducted the experiments and collected the data. The author analyzed the data together with Andrés Lucero. The author was the principal author of the publication.

**P4**  The publication reports a laboratory study that compares three different methods for moving virtual objects between mobile devices. The author designed the object movement methods together with Petri Piippo and Guido Grassel, and planned the study together with Jarno Ojala and Thomas Olsson. Jarno Ojala conducted the experiments and collected the data. The author analyzed the data together with Jarno Ojala. The author was the principal author of the publication.

**P5**  The publication presents an elicitation study that gathers interaction methods for common multi-device interaction tasks on wearable devices from groups of end users. The author planned the study together with Kaisa Väänänen. The author conducted the elicitation study sessions and collected the data. The author analyzed the data together with Parisa Pour Rezaei and Kaisa Väänänen. The author was the principal author of the publication.

# List of Acronyms

AR      Augmented Reality
BCI     Brain-Computer Interface
GUI     Graphical User Interface
HCI     Human-Computer Interaction
IoT     Internet of Things
IT      Information Technology
PAN     Personal Area Network
PC      Personal Computer
NFC     Near-Field Communication
RQ      Research Question
UCD     User-Centered Design
UI      User Interface
UX      User Experience
VNC     Virtual Network Computing
Wi-Fi   Wireless Fidelity

# Chapter 1

# Introduction

This chapter introduces the research presented in this thesis. Section 1.1 first explains the background and motivation for the research. Section 1.2 discusses the objectives and scope of this thesis, and Section 1.3 defines the key terms used. Section 1.4 then presents the research questions. Sections 1.5 and 1.6 summarize the research approach and the main contributions. Finally, Section 1.7 gives an outline of the structure of the thesis.

## 1.1   Background and Motivation

People today commonly have multiple information devices. In addition to personal computers, smartphones and tablets have established themselves as new general device categories that are widely used. Some conventional electronic devices, such as televisions, music players, and cameras, have become connected to the Internet and gained the ability to provide access to many of the same applications and services as computers. Emerging industry trends, including wearable devices, connected cars, and the Internet of Things, suggest continuing growth in the number and diversity of information devices in the future.

As people increasingly have multiple information devices, situations and workflows where several devices are combined and used together to accomplish a task have become usual. In some situations, a person might start a task with one device and then continue it with another device. For example, a person might find an interesting page when browsing the web with a smartphone on a bus and then continue reading the same page with a tablet at home; or a person might start writing an e-mail message on a smartphone but then switch to a laptop computer as the message grows longer and more complicated; or a person might take a photo with a cameraphone and then immediately attach it to a presentation that they are preparing on a laptop computer. In other situations, a person might use several devices at the same time. For example, while watching a movie on a television, a person might search for additional information related to the movie on a tablet; or a person might listen to music in the background with a smartphone

while writing a document on a desktop computer; or a person might share the smartphone's Internet connection with their other devices.

In addition to individual persons using multiple information devices, groups of co-located persons may also join their information devices together for multi-user multi-device applications. This enables a co-located group to engage in collaborative activities and experiences with their devices. For example, a group of friends in a café could view photographs and videos together with their smartphones; or the participants of a business meeting could use their laptops to share and collaboratively edit documents; or family members could play multi-player games with their smartphones and tablets in the living room.

While these developments towards computing with multiple devices offer many opportunities, they also bring about many challenges. This creates a need for novel interfaces and applications that better support multi-device use.

## 1.2  Objectives and Scope

This thesis belongs to the field of Human-Computer Interaction (HCI) research. Human-computer interaction is a multi-disciplinary field of research that studies the design, evaluation and implementation of interactive computing systems for human use and the major phenomena surrounding them (Hewett et al. 1992).

The overall goal of this thesis is to create new scientific knowledge to inform the design of future interfaces, applications, and technologies that better support using multiple information devices together. This new knowledge includes descriptive knowledge that reports and explains how and why people combine and use multiple information devices together in their everyday lives. It also includes prescriptive knowledge that aims at practical improvements in the user experience of multi-device use. In particular, this thesis investigates two common tasks in multi-device interaction in detail: device binding and cross-display object movement. Device binding refers to the problem of how persons create initial connections between their devices to enable multi-device interactions. Cross-display object movement refers to the problem of how to move virtual visual objects, such as content items or application sessions, between device displays in multi-device interaction.

This thesis approaches the problem of multi-device use primarily from the perspective of an end user who is using a multi-device system, rather than from the perspective of the technologies that are needed to enable multi-device computing. This thesis considers the use of multi-device systems that may consist of a wide variety of information devices that differ in terms of form factors, features, and capabilities. These devices include smartphones, computers, tablets, home media centers, smartwatches, and smartglasses. This thesis addresses both multi-device interaction by a single user and multi-device interaction by multiple users. However, the research presented in this thesis is limited to co-located interaction, that is, to interaction where all users and devices are at the same physical location in proximity of each other. Interaction between remote users and devices is beyond the scope of this thesis.

## 1.3  Terminology and Definitions

**Information device.**  Device that can be used to create or consume digital information. Examples of such devices include personal computers, smartphones, tablets, televisions, cameras, game devices, music players, navigation devices, and wearable devices.

**Wearable device.**  Information device that can be worn on the user's body. Examples of such devices include smartwatches and smartglasses.

**Human-computer interaction.**  Communication between humans and computers with the purpose of humans accomplishing some goals (Dix et al. 2004, p. 4).

**Interaction method.**  Way of using physical input and output devices to perform a generic task in human-computer interaction (Foley et al. 1990). Also known as interaction technique.

**Multi-device interaction.**  Interaction that happens with multiple information devices instead of a single device. Such interaction may involve starting a task with one device, switching the device in the middle of the task, and completing the task with another device, or it may involve interacting with multiple devices simultaneously.

**Multi-user interaction.**  Interaction that involves multiple users instead of a single user.

**Co-located interaction.**  Interaction where all users and devices are at the same physical location in proximity of each other.

**Experiential factor.**  Factor that influences the user experience of an interactive system. Such factors may be related, for example, to users, systems, tasks, physical environments, or social situations.

**Device binding.**  Creating initial connections between devices to enable multi-device interactions. Also known as device association, pairing, or coupling.

**Group binding.**  Binding devices of multiple users to enable multi-user multi-device interactions. Also known as group association.

**Cross-display object movement.**  Moving a virtual visual object, such as a content item or an application session, from a specific location on one device display to a specific location on another device display in multi-device interaction (Nacenta, Gutwin, et al. 2009).

## 1.4  Research Questions

According to User-Centered Design (UCD) approach, the design of interfaces, applications, and services should be based on understanding the users' behavior and needs. While a number of studies have addressed how people actually

combine and use multiple information devices together in real life, these studies are still relatively few given the broadness of the topic. Of these earlier studies, many have been made before the availability and widespread adoption of modern smartphones, tablets, and cloud services. Most of the earlier studies have also focused on technologically advanced users and use cases related to information work. Therefore, there is a need to provide an updated view into how people utilize and manage the new extended ecologies of devices and services, especially addressing diverse groups of users and exploring beyond work-related use. The first research question of this thesis focuses on this problem.

> RQ1  How do people combine multiple information devices in everyday activities and tasks?

Compared to the design of traditional single-device user interfaces, the design of multi-device interfaces presents a difficult problem as there are multiple devices involved and the interaction is distributed across them. Information about the experiential factors that significantly influence the user experience of multi-device interaction in different situations of use would therefore be useful as these factors should receive particular attention in the design process. The second research question of this thesis examines these experiential factors. In particular, the question is addressed through investigating two common tasks in multi-device interaction: device binding and cross-display object movement.

> RQ2  Which experiential factors should be considered when designing methods for multi-device interaction?

Looking into the future, wearable devices are emerging as a new category of information devices that offer interesting opportunities to support collaborative activities and shared experiences between co-located persons through multi-user multi-device applications. To make such applications possible, the wearable devices need to support multi-device interactions. While there exists a wide variety of multi-device interaction methods that have been designed for conventional devices (such as smartphones, tablets, or computers), these methods may not be optimal for wearable devices as wearable devices have many differences compared to conventional devices. This creates a need for novel multi-device interaction methods for wearable devices. Ideally, these methods should be natural and intuitive, building on people's spontaneous perception of the characteristics and affordances of wearable devices. The third research question of this thesis explores this problem, focusing in particular on natural methods for initiating multi-device interactions between wearable devices. Regarding devices, the research in this thesis considers the two most common wearable device categories today: smartwatches and smartglasses.

> RQ3  How would a group of co-located persons naturally initiate multi-device interactions between their wearable devices?

| Research Question | P1 | P2 | P3 | P4 | P5 |
|---|---|---|---|---|---|
| RQ1 How do people combine multiple information devices in everyday activities and tasks? | • | | | | |
| RQ2 Which experiential factors should be considered when designing methods for multi-device interaction? | (•) | • | • | • | (•) |
| RQ3 How would a group of co-located persons naturally initiate multi-device interactions between their wearable devices? | | | | | • |

Table 1.1: Contributions of the publications to the research questions ("•" = primary contribution, "(•)" = supporting contribution).

## 1.5 Research Approach and Process

The research presented in this thesis was carried out within the general framework of design science research. According to the design science research approach, scientific research on information technology consists of two distinct types of research: behavioral science research and design science research. The first research question RQ1 of this thesis belongs to behavioral science research while the second research question RQ2 and the third research question RQ3 belong to design science research. Considering research methods, this thesis primarily uses qualitative research methods but the qualitative research methods are complemented with quantitative research methods, following a mixed methods approach.

This thesis consists of five empirical studies S1–S5 which are reported in five original publications P1–P5. Table 1.1 summarizes the contributions of each publication to the research questions. The first study S1 is a diary study that investigates how people combine multiple devices in their everyday lives and aims to provide a general understanding of the current practices and challenges in multi-device use. It primarily contributes to the first research question RQ1. The second study S2 is a laboratory evaluation that compares three different methods for group binding on smartphones. Based on the results of the study S2, study S3 suggests a novel more flexible method for group binding and evaluates it with a laboratory study. The fourth study S4 is a laboratory evaluation that compares three different methods for cross-display object movement on personal mobile devices. The studies S2, S3, and S4 contribute to the second research question RQ2. Finally, the fifth study S5 is an elicitation study that gathers interaction methods for common multi-device interaction tasks on wearable devices directly from end users. It primarily contributes to the third research question RQ3. Overall, a total of 110 persons participated in the studies presented in this thesis.

## 1.6    Contributions

Regarding the first research question RQ1, this thesis contributes descriptive knowledge that gives a rich picture of the current and emerging practices of multi-device use. It provides a three-level categorization of patterns of combining multiple information devices together. It also provides theoretical contributions that explain some of the observed practices, including how users decide which devices to use in a specific situation and how users decide to change the device they are currently using.

Regarding the second research question RQ2, this thesis contributes a set of experiential factors that should receive special consideration when designing methods for multi-device interaction. These factors are based on a series of studies evaluating interaction methods for device binding and cross-display object movement tasks. These studies also provide a number of secondary contributions, including novel methods for multi-device interaction and comparisons of different multi-device interaction methods against each other.

Regarding the third research question RQ3, this thesis contributes a categorization of user-generated methods for initiating multi-device interactions between wearable devices. It also provides initial evaluations of the strengths and weaknesses of the different methods. The categorization gives an overview of the design opportunities for solving the target identification problem in multi-device interaction methods for wearable devices.

## 1.7    Outline of the Thesis

This thesis consists of seven chapters that are structured as follows.

The first chapter, Introduction, explains the background and motivation for the research, defines the scope and the research questions, and summarizes the research approach and the contributions. The second chapter, Related Work, presents a summary of prior research on topics related to this thesis, including technologies, systems, and interaction solutions that have been developed to enable and support multi-device use as well as studies that have investigated how people actually combine and use multiple devices together in real-life situations. The third chapter, Research Approach and Methods, describes the general research approach and the detailed research methods that were used in the research presented in this thesis, and discusses the related ethical considerations. The fourth chapter, Research Process, first gives an overview of the research process and then describes in more detail the five empirical studies that form the core of this thesis. The fifth chapter, Results, summarizes the results of the research presented in this thesis. The sixth chapter, Discussion, discusses the contributions of this thesis, their reliability, validity, and generalizability, as well as potential directions for

future research. Finally, the seventh chapter, Conclusion, concludes this thesis by summarizing its key contributions.

The five original publications P1–P5 are included as an appendix at the end of this thesis.

# Chapter 2

# Related Work

This chapter presents a summary of prior research on topics related to this thesis. Section 2.1 first discusses the early visions that have had major influence on the development of information technology towards computing with multiple devices. Section 2.2 then provides an overview and gives examples of different kinds of multi-device computing systems and applications presented in the prior literature. Section 2.3 surveys interaction methods and technologies that have been developed to enable and support multi-device use. Sections 2.4 and 2.5 cover in more detail device binding and cross-display object movement, two interaction tasks that are in the focus of this thesis. Beyond multi-device technologies, systems, and interaction solutions developed in research laboratories, Section 2.6 reviews studies that have investigated how people actually combine and use multiple devices together in real-life situations. Finally, Section 2.7 identifies and discusses some gaps in prior research that this thesis aims to fulfill.

## 2.1   Early Visions of Multi-Device Computing

Over the past seventy years, computing has evolved from mainframe computing through personal computing towards computing with multiple devices (Weiser and Brown 1997). The early days of computing can be called the **mainframe computing** era. Mainframe computers were scarce and expensive devices that were operated by experts. Each computer was shared between many people. The second era of computing can be called the **personal computing** era. In that era, it became possible for an ordinary person to have their own personal computer that contained their data and applications. There was one computer per person. In the third era of computing, there are large numbers of computers everywhere around us. These computers have many different forms and sizes and they are all connected through a network. Each person is shared between multiple computers.

In the seminal paper "The Computer for the 21st Century", Mark Weiser (1991) presents his vision of the era of computing with multiple devices. He calls it the era of **ubiquitous computing**. In this era, computing technologies become so

ubiquitous and integrate so seamlessly into the fabric of everyday life that they disappear—people use them unconsciously to accomplish everyday tasks. Weiser envisions computing devices in three different scales: tabs, pads, and boards. Each scale of devices is suited to a particular task. Tabs are inch-scale devices approximately the size of a sticky note: they can be used for simple applications such as a pocket calculator, be attached to persons or objects to track them, or provide physical representations of shrinked application windows or files in a similar way as icons on a traditional computer desktop. Pads are foot-scale devices comparable to a sheet of paper or a book or a magazine: a person may have many pads, each representing a different task like a pile of papers on a real desk or a window on a traditional computer desktop. Boards are yard-scale devices equivalent to a whiteboard or a bulletin board: they support collaboration and information sharing between multiple people. All devices are interconnected in a ubiquitous network, and their real power comes from interactions between them rather than from any single device alone. Devices are also aware of their location and surroundings, and can adapt their behavior accordingly. Devices may have no individual identity or importance: they are everywhere and can be just grabbed and used like scrap paper.

Weiser's paper has been highly influential in guiding information technology research and industry over the last two decades. His vision of computing with devices of different scales has been realized in the form of smartphones, tablets, and a variety of wall-sized and tabletop computers. Practically all devices have become connected by ubiquitous cellular and wireless networks and supported by cloud infrastructure. However, there are also important differences between Weiser's vision and the reality of computing today. For example, today's practices of tablet use (Müller, Gove, and Webb 2012) resemble more the conventional practices of personal computer use than the ways of using pads described by Weiser. Most devices are also highly valued personal items that people are not willing to share (Hang et al. 2012) rather than scrap devices envisioned by Weiser.

In his book "The Invisible Computer", Donald Norman (1999) criticizes the complexity of general-purpose personal computers and suggests families of **information appliances** as a solution. (The term information appliance was invented by Jef Raskin already in 1978 (Norman 1999, p. 51).) Information appliances are computing devices that have been designed to perform a specific function or activity. Examples of information appliances include digital musical instruments, cameras, navigation devices, calculators, and health monitoring devices. In the appliance model of computing, each task has its own device that has been particularly designed for that task, and learning to use the device equals to learning the task. In order to change tasks, the user changes devices, and to resume a task, the user moves back to the appropriate device. While individual information appliances can be used on their own, they are really intended to be used as families of related devices that have been designed to work together smoothly and effortlessly. For example, a family of photography appliances could consist of different kinds of cameras, viewing appliances, storage appliances, and printing appliances. All devices are connected and supported by universal and invisible infrastructure to allow seamless information exchange between them. New applications, arts, and industries can be created by combining and sharing information between devices in novel and innovative ways.

Since the publication of Norman's book, we have seen the diversity of form factors in computing devices to remarkably increase and many new device categories have emerged alongside the personal computers. But controversially, the most popular new device categories, namely smartphones and tablets, have not been specialized devices for a particular purpose but general-purpose devices that integrate a wide range of hardware capabilities and can support numerous tasks and activities through different software applications in a way similar to personal computers. Especially smartphones have to a large degree replaced and taken over the roles of a myriad number of other more specialized digital and non-digital devices, including telephones, address books, calendars, watches, alarm clocks, calculators, music players, portable game consoles, cameras, camcorders, maps, navigation devices, flashlights, and wallets. Norman (1999, pp. 60–62) also discusses the tradeoffs between specialized and general-purpose devices. While specialized devices are tailored to the task and provide ease of use and simplicity, general-purpose devices can offer cost advantages, can allow easy information transfer and unexpected synergies between tasks, and can be extended to support new purposes. Especially in mobile situations, general-purpose devices provide great convenience as a single small and lightweight device may offer the functions of many specialized devices. This illustrates that both specialized and general-purpose devices have their benefits and drawbacks. Rather than being mutually exclusive alternatives, they can complement each other in different situations of use.

## 2.2 Multi-Device Systems and Applications

Inspired by visions of computing with multiple devices, a large amount of research and development work has been invested into building such computing systems in practice. In this section, we aim to provide an overview and give examples of different kinds of multi-device and multi-display computing systems and applications that have been presented in prior literature.

**Multi-display systems** have long existed in particularly demanding operating environments, such as in control rooms overseeing air traffic or factory production. With the decreasing cost of display hardware, multi-monitor workstations have become relatively inexpensive and are a common sight in today's office environments (Grudin 2001). Other kinds of multi-display systems may consist of a single device with multiple independent displays that the user can configure in different ways. For example, Codex (Hinckley, Dixon, et al. 2009) is a dual-display book-like device that has a hinged binding with embedded sensors. It demonstrates a wide variety of rich interactions between two displays. Surface-Constellations (Marquardt, Brudy, et al. 2018) is a modular hardware platform with a comprehensive library of link modules to easily connect mobile devices to create different kinds of multi-display environments.

A large part of the research on multi-device systems has focused on development of interactive **smart spaces** to support collaborative work. Colab (Stefik et al. 1986) was an early experimental meeting room environment that connected personal workstations of two to six persons with a large shared touch display. Later projects, such as the i-Land (Streitz et al. 1999) and the Stanford iRoom (Johan-

son, Fox, and Winograd 2002), have built entire physical spaces that combine a wide variety of displays and devices of different types and sizes, including personal mobile devices, interactive tables, and large wall-sized displays. ReticularSpaces (Bardram et al. 2012) is a multi-display smart space based on the principles of activity-based computing.

Another research direction that has received a lot of attention is multi-device systems that combine **handheld devices**, in particular personal mobile devices, with other devices. Pebbles (Myers, Stiel, and Gargiulo 1998) demonstrates ways how handheld computers can serve as useful adjuncts to personal computers. Pass-Them-Around (Lucero, Holopainen, and Jokela 2011) allows a group of co-located persons to share photographs between their smartphones using various interaction techniques that have been inspired by practices of sharing paper photos. Mobicomics (Lucero, Holopainen, and Jokela 2012) is a system for collaboratively creating and editing photo comics on smartphones and for sharing them through public displays. United Slates (N. Chen, Guimbretiere, and Sellen 2012) is a multi-tablet reading environment that supports active reading. PaperWindows (Holman et al. 2005) explores interactions between multiple flexible paper-like devices. Siftables (Merrill, Kalanithi, and Maes 2007) are tiny computers with graphical displays that can sense and communicate with each other. They can be physically manipulated as a group to interact with digital information.

In **second screen** systems, the experience of watching a television is enhanced with the use of other devices, typically personal mobile devices. Robertson et al. (1996) report an early experiment of combining a television with a handheld device for an interactive real estate information service. Cruickshank et al. (2007) present a system that allows the user to view the electronic program guide and to control the content that is shown on the television using a handheld device. Mate, Chandra, and Curcio (2006) analyze scenarios of transferring content between the television and a handheld device, for example, continuing to watch a television show on a personal mobile device when leaving the place where the television is located. Nandakumar and Murray (2014) describe a companion application that provides time-synchronized information to support the viewing of a complex long arc television series. FanFeeds (Basapur et al. 2012) allows the viewer to augment television shows with time-synchronized comments and links to external content and to share the augmentations within their social circle.

Some researchers have also explored multi-device systems that involve **wearable devices**, either in combination with conventional devices or in combination with other wearable devices. Duet (X. A. Chen et al. 2014) demonstrates joint interactions between a smartwatch and a smartphone. MultiFi (Grubert et al. 2015) is an interactive system that combines smartwatches with smartglasses. Budhiraja, Lee, and Billinghurst (2013) suggest using a handheld device with a head-mounted display to control augmented reality applications. Benko, Ishak, and Feiner (2005) describe a system that integrates head-mounted displays with a large tabletop display.

Terrenghi, Quigley, and Dix (2009) present a taxonomy of multi-person-display ecosystems. Their taxonomy categorizes multi-person-display ecosystems according to two attributes: the scale of the ecosystem and the nature of the social interaction. The scale of the ecosystem refers to the physical dimensions of the

entire ecosystem, including the displays, the people, and the space in which they are placed. There are five possible scales: inch scale, foot scale, yard scale, perch scale, and chain scale. The first three scales of Terrenghi et al. map to the three scales of ubiquitous computing devices by Weiser (1991) but note that Weiser's scales are scales of devices while the scales of Terrenghi et al. are scales of ecosystems. The scale of the whole ecosystem tends to be one scale larger than the scale of the largest device in the ecosystem. The nature of the social interaction refers to number of persons who are actively controlling content on the displays, the number of persons who are primarily receiving content from the displays, and their relationship. There are five possible categories of social interaction: one-to-one, one-to-few, few-to-few, one/few-to-many, and many-to-many.

## 2.3   Interfaces to Support Multi-Device Use

In this section, we survey interaction technologies and methods that have been developed to support multi-device use. Device binding and cross-display object movement, two interaction tasks that are in the focus of this thesis, are covered separately in more detail in Sections 2.4 and 2.5.

In multi-device systems, the user interface is distributed across several devices. Such user interfaces are commonly referred to as **distributed user interfaces**. The distribution of the user interface may happen on several levels (Elmqvist 2011). The input and the output of the system may be split between many devices. The devices may operate on different computing platforms that vary in terms of hardware architectures, operating systems, and networking technologies. The interaction may be synchronous and happen simultaneously on all devices, or it may be asynchronous and be distributed over time. The system may be used by a single user or by multiple users. Some definitions of distributed user interfaces also include user interfaces for interaction between geographically distributed spaces, for example, interfaces that have been developed to support access to remote devices or remote collaboration. While similar user interface and technology solutions could also be used to support remote interaction, in this thesis we focus on co-located interaction, that is, to interaction where all users and devices are located within the same physical space. Interaction between remote users and devices is beyond the scope of this thesis.

**Migratory applications** (Bharat and Cardelli 1995) are applications that are not tied to one computer but can freely roam over the network from one computer to another, taking their user interface and application contexts with them. However, they require that all computers operate on the same computing platform. In **migratable user interfaces** (Grolaux, Van Roy, and Vanderdonckt 2004), the distribution is implemented on the user interface component level instead of entire applications. Many commercial technologies, such as the X Window System, Windows Remote Desktop Services, and Virtual Network Computing (VNC) also support transferring application windows or entire desktops between devices.

One of the fundamental challenges in multi-device interaction is how to adapt the user interface to different device configurations. Individual devices may differ in many dimensions, including the display size, available input devices, processing

power, physical size, and form factor. In a multi-device configuration, there may be numerous ways to distribute the functions of the user interface across different devices. The user interface may also be adapted to the more general environment and context of use. **Plasticity** (Thevenin and Coutaz 1999) refers to the capacity of a user interface to withstand variations of both the system physical characteristics and the environment while preserving usability. **Responsive design** (Marcotte 2011) is a pragmatic design approach that aims to produce web pages that render well on a variety of devices with different display sizes, including computers and handheld mobile devices. Nebeling and Dey (2016; 2017) suggest systems that allow ordinary users to adapt existing single-device web interfaces for multi-device use manually and semi-automatically.

To address the challenge of user interface adaptation, many researchers have explored **model-based approaches** to user interface specification (Paternò 2000). In model-based design, the user interface ideally needs to be specified only once and can then be automatically generated numerous times in different device configurations, always in an optimal way. Examples of model-based frameworks, languages, and tools for multi-device user interface design include TERESA (Mori, Paternò, and Santoro 2004), MARIA (Paternò, Santoro, and Spano 2009), Panelrama (Yang and Wigdor 2014), and AdaM (Park et al. 2018).

Denis and Karsenty (2004) introduce the concept of **inter-usability** to refer to the ease with which users can reuse their knowledge and skills as they change from one device to another when using of a multi-device service. Denis and Karsenty identify two dimensions that influence inter-usability: knowledge continuity and task continuity. Knowledge continuity relates to the knowledge constructed from the use of the devices. In order to maintain knowledge continuity, ideally all devices should present the service in the same way and allow access to the same data and functions. Task continuity, on the other hand, relates to the memory of the last operations performed with the service independently of the device that was used. To ensure task continuity, it should be possible to recover the state and context of the service when changing the device. Wäljas et al. (2010) present another conceptual framework of cross-platform service user experience. They identify three main elements of cross-platform user experience: fit for cross-contextual activities, flow of interactions and content, and perceived service coherence.

Spatial relationships play an important part in everyday interactions between humans. Similar concepts can also be exploited in multi-device user interfaces. In **proxemic interaction** (Ballendat, Marquardt, and Greenberg 2010), the interaction methods build on detailed information about the position, identity, movement, and orientation of the nearby devices and users. This information can be used, for example, to trigger interactions, to mediate interactions between different users, and to detect user's attention to other users or objects. F-formations refer to the physical arrangements that groups of people adopt when they engage in focused conversational encounters with each other (Marquardt, Hinckley, and Greenberg 2012). **3D tracking technologies** (Rolland, Davis, and Baillot 2001) that make it possible to follow the position and orientation of devices and users in physical space are a key enabler for proxemic interaction.

A common technique in multi-user interfaces in **input and output redirection**. In input redirection, the input events from one device are sent to another device

| Dimension | Description |
|---|---|
| UI Distribution | How the user interface elements can be distributed across multiple devices at a given time (for example, dynamic distribution that can be changed at runtime or static distribution that cannot be changed)? |
| UI Migration | What is the level of continuity when user changes the device? |
| UI Granularity | At which granularity the user interface can be distributed or migrated between devices (for example, entire user interface, groups of user interface elements, single user interface elements, or components of user interface elements)? |
| Triggers | How changes in the configuration of a multi-device user interface are triggered (for example, manually by user push or pull or automatically by the system)? |
| Device Sharing | How devices can be shared between multiple users? |
| Timing | When should changes in the configuration of a multi-device user interface occur (for example, immediately or later at a user-specified time)? |
| Interaction Modalities | What modalities does the multi-device user interface support (for example, all devices support the same modality, different devices support different modalities, or the devices support multiple modalities simultaneously)? |
| UI Generation | When the user interface is generated for a specific device configuration (for example, at design time or at runtime)? |
| UI Adaptation | How the user interface is adapted to a specific device configuration (for example, by scaling, by transducing, or by transforming)? |
| Architecture | What is the underlying platform architecture (for example, client-server or peer-to-peer)? |

Table 2.1: Dimensions of multi-device user interfaces. Based on Paternò and Santoro (2012).

for processing, and in output redirection, the output primitives from one device are sent to another device for rendering. iStuff (Ballagas et al. 2003) is a toolkit that allows dynamic mapping of input and output devices to different applications. ARIS (Biehl and Bailey 2004) is a space window manager that allows the user to control applications across different devices in an interactive space. Interaction methods for moving user interface elements across different displays are covered in more detail in Section 2.5. **Display tiling** refers to combining the displays of multiple devices into a single large display surface.

Development of multi-device applications and user interfaces where input and output are distributed across several devices presents a major challenge also from **software development** perspective. Many software platforms and development tools have been proposed to support programmers in the development of such applications. Examples of software platforms for multi-device applications include iROS (Johanson, Fox, and Winograd 2002), Gaia (Román et al. 2002), and Impromptu (Biehl, Baker, et al. 2008). WatchConnect (Houben and Marquardt 2015) and Weave (Chi and Li 2015) are more recent software toolkits intended for the development of multi-device applications that particularly target wearable devices. XDStudio (Nebeling, Mintsi, et al. 2014) is a visual user interface builder designed to support interactive development of multi-device interfaces. Chi, Li, and Hartmann (2016) suggest an automatically-generated multi-device storyboard visualization to illustrate the execution of multi-device applications.

Finally, Paternò and Santoro (2012) present a framework for analysis, design, and comparison of multi-device user interfaces. Their framework identifies ten key design dimensions of multi-device interfaces. The dimensions are summarized in Table 2.1.

## 2.4 Device Binding

Before a set of devices can be used for multi-device interactions, they must first be joined together into a multi-device ecosystem. This procedure of connecting devices for multi-device interactions is generally called **device binding** or ecosystem binding, but it is also known as device association, pairing, or coupling. If the devices to be connected are owned by several different persons, the procedure may be called **group binding** or association. Device binding is a complex technical procedure that involves multiple steps. The required system and application software must be initiated on all devices. The devices need to discover the other devices in proximity and the devices that are intended to join in the multi-device ecosystem need to be identified. A channel for communication between the devices taking part in the ecosystem then has to be set up to enable exchange of data and coordination of user interactions. Short-range radio technologies, such as Wi-Fi or Bluetooth, are typically used for communication between devices. As the goal is to make possible spontaneous interactions between devices, the users should be able to bind devices that have no earlier knowledge of each other in a quick and easy manner. If the procedure of device binding becomes too difficult or cumbersome, the users may lose their interest in multi-device interactions entirely. The binding methods should also provide sufficient indications and security, so that the users can be sure that the right devices are safely connected.

From interaction design perspective, the problem of device binding involves two important sub-problems: device identification and device authentication. **Device identification** refers to designating the devices that are to be bound with each other. There may be a large number of devices available in the proximity and the users need to indicate by some means which of these devices should be connected into a multi-device ecosystem. **Device authentication** refers to making sure that the users are really connecting the devices they intend to. Because wireless communications is invisible, the users cannot see the wireless links between the devices and therefore cannot be sure which devices are actually being connected. This opens a possibility for so called Man-in-the-Middle attacks: a malicious third party may secretly relay, monitor, and possibly modify communication between the two parties who assume that they are directly communicating with each other. To prevent such attacks, a wide variety of methods that authenticate the wireless connection over an auxiliary communication channel has been proposed. An **auxiliary communication channel**, also known as an out-of-band channel, is a secure, typically low bandwidth communication channel that the human users can easily perceive and manage.

The problem of device binding in ubiquitous computing environments without a trusted third party was first identified by Frank Stajano (2000). To solve the problem, he suggested a device binding method called the Resurrecting Duckling (Stajano and R. Anderson 2002): when a device is initiated, it associates to the first device it connects to, in a similar way as a duckling assumes that the first moving object it sees is its mother. Stajano's method was based on using electrical contact, for example, physical cables, as an auxiliary communication channel to securely connect the devices, which makes the method impractical in many real-life applications that use wireless communication.

The most common device binding method today is the **Scan & Select** method. It is widely used in different variations in applications that use Bluetooth or Wi-Fi technologies for communication. In the Scan & Select method, the device first scans its environment for other available devices. The list of the devices that were found is then presented to the user and the user can select the other devices they want to bind with. For authentication, short strings (also known as PIN codes) are typically used. The user is expected to copy or compare these strings between devices to authenticate the connection. In the user interface, the authentication strings can be represented in different forms, for example, as numbers, words, graphical images, or audio signals.

The problem of device binding has been extensively studied in the fields of ubiquitous computing, human-computer interaction (Terrenghi, Quigley, and Dix 2009; Chong, Mayrhofer, and Gellersen 2014), and security research (Suomalainen, Valkonen, and Asokan 2009), and a wide variety of alternative methods for binding devices has been proposed. These methods vary in terms of device hardware requirements, amount of user involvement, and level of provided security. Some of the proposed methods encompass only device identification or authentication, while others combine both identification and authentication into a single user action. We can identify several different categories of proposed device binding methods: methods based on synchronous user actions, methods based on spatial alignment, and methods based on the use of auxiliary devices. We will next discuss and provide examples of methods in each of these categories.

The first category of methods proposed for device binding is based on **synchronous user actions**. In SyncTap (Rekimoto 2004), the user can bind two devices by pressing buttons on both devices at the same time. By comparing the timing information of the button presses, the devices can correctly identify and authenticate each other. The user may also bind two devices by touching them simultaneously. A Personal Area Network (PAN) (Zimmerman 1995) can then be used to create an auxiliary communication channel between the devices through the user's body. Alternatively, it is possible for two users to connect wearable devices on their bodies by touching each other, for example, by shaking hands. In Smart-Its Friends (Holmquist et al. 2001), the user binds several devices by holding them together and waving or shaking them. This imposes a similar movement pattern on all devices that are held together, and this pattern is likely to be different from the movement pattern of any other device at the same time, at least in the proximity. The movement pattern can be measured with a motion sensor and compared to identify and authenticate the devices. Hinckley (2003) suggests a similar method where devices are connected by bumping them together. It is possible to detect bump events with motion sensors as simultaneous and roughly equal but opposite patterns of forces on each device. In Stitching, (Hinckley, Ramos, et al. 2004), the user may connect two devices equipped with touch displays with a continuous touch gesture spanning both displays. The stroke gesture begins from the screen of the first device, crosses over the bezels, and ends on the display of the second device. The gesture can be detected by comparing the timestamped touch events observed on both displays.

The second category of methods proposed for device binding is based on **spatial alignment** of the devices. Mayrhofer and Welch (2007) suggest a method where the user binds devices by pointing one device at another. An auxiliary communication channel for authenticating the connection is established between the devices with visible laser light, with one of the devices incorporating a laser transmitter and the other device incorporating a laser receiver. In Proximal Interactions (Rekimoto et al. 2003), the user can bind two devices by using one device to touch another device. A near-field communication channel implemented, for example, with RFID technology, is used as an auxiliary communication channel. Security is based on the limited transmission range of near-field communication. Rukzio et al. (2006) compare touching and pointing with Scan & Select for device identification. Their results indicate that people prefer touching for devices that are near, pointing for devices that are further away, and scanning only if touching and pointing are not possible, for example, if the device is far away and cannot be seen. Kray et al. (2008) propose a system where the devices are bound when the user places them in proximity of each other. Each device has a number of spatial proximity regions around it. The regions are associated with different multi-device interactions, and putting another device into one of these regions connects the devices and triggers the particular interaction between them. The devices are tracked with external cameras and the proximity regions are visualized with a projector system.

The third category of methods proposed for device binding is based on using various **auxiliary devices** to bind devices. In tranSticks (Ayatsuka and Rekimoto 2005), the user may bind two devices by placing one of a pair of physical tokens with the same identifier into each device. The method exploits the familiar metaphor of a physical cable, with the tokens representing the endpoints of a

"virtual cable" that connects the devices. In Seeing-is-Believing (McCune, Perrig, and Reiter 2005), the user takes a photo of the devices they want to bind with their cameraphone. The cameraphone scans barcodes attached to the devices and uses that information to securely connect the devices. The camera image forms a human-perceivable visual auxiliary channel for authenticating the connection.

Chong, Mayrhofer, and Gellersen (2014) suggest an alternative classification of device binding methods. In their classification, the device binding methods are divided into four categories: guidance, input, enrollment, and matching. The guidance category encompasses techniques where users act in the real world to bind devices together. The methods are based on concepts that are external to the involved devices, such as contact, alignment, and proximity. The methods in the input category focus user actions on the involved devices. They build on conventional user interface concepts such as triggering commands, entering data, and direct manipulation. In the enrollment category, the user attaches an identity to a device and then presents this identity to other devices to bind them with it. Finally, the matching category describes methods where users compare output of the involved devices to confirm or reject a connection.

The development of new device binding methods has been to a large extent driven by technology, focusing on aspects such as security, efficiency, and novel sensing techniques. Chong and Gellersen (2011; 2013) present two studies with a more user-centered approach: using the **elicitation study methodology**, they asked users to suggest methods that the users would spontaneously and intuitively use to connect a wide variety of different device combinations. The devices were represented by low-fidelity acrylic prototypes. Instead of a single prevalent method, several common categories of methods were found, including Scan & Select, proximity, button event, device touch, and gesture. Physical device attributes, such as form factor, mobility, and flexibility, as well as prior knowledge and experiences of technology were found to influence users' expectations of how device binding should be done.

Device binding methods are not just pragmatic ways to connect devices—their user experience is also affected by many **non-pragmatic factors**. In realistic situations, the users do not always use the easiest or the fastest method available, nor the one they think is technically the best. Many factors influence the users' choices of binding methods, including sensitivity of the data, time constraints, and social factors such as the place, the other people present, and the social conventions considered appropriate (Ion et al. 2010; Rashid and Quigley 2009). Users have been found to be willing to accept security risks in order to comply with social norms (Ion et al. 2010).

Most research on device binding methods has focused on scenarios of an individual user binding two devices with each other, for example, a user binding a headset with a smartphone or a user binding a laptop computer with a wireless access point. More complex scenarios involving **multiple users and devices** have been studied significantly less. Device binding methods that have been designed for individual users may not necessarily be applicable to scenarios involving multiple users as multi-user scenarios are in many respects different from single-user scenarios. In multi-user scenarios, communication between the users provides an additional source for potential errors (Kainda, Flechais, and Roscoe 2010). On

the other hand, the users are typically willing to help each other and make decisions by mutual agreement, which reduces the amount of errors (Kainda, Flechais, and Roscoe 2010). Users may not be willing to hand in their personal devices to other people who they do not know very well, not even for short moments, which can make methods that involve physical exchange of devices unacceptable (Uzun, Saxena, and Kumar 2011; Hang et al. 2012). Group binding is not necessarily just a sequence of binding operations between pairs of devices. It can also be considered as a single procedure where all the devices are bound together at the same time (Terrenghi, Quigley, and Dix 2009; Chong, Mayrhofer, and Gellersen 2014).

Chong and Gellersen (2012) present a framework that summarizes and categorizes the factors that influence the usability of spontaneous device binding. They identify three main categories of factors: technology, user interaction, and application context. The technology category includes technical factors related to the involved devices and the used communication channel. The user interaction category covers factors that influence the user's actions. Finally, the factors in the application context category address the conditions under which the device binding task takes place.

## 2.5 Cross-Display Object Movement

In situations where several devices are used together, there is frequently a need to move virtual visual objects from the display of one device to the display of another device. This common procedure in multi-device use is generally known as **cross-display object movement**. Depending on the application and the user interface, there can be many different types of objects that the users should be able to move between devices. These objects may include content items, for example, text documents, photographs, audio or video clips, or items of application-specific content types. The objects to move may also include application sessions or user interfaces, for example, entire application windows or sub-parts of the user interface.

As moving objects between devices is a very common problem, a large number of different solutions has been developed for it. **Conventional solutions** that are widely commercially available include connecting the devices directly for data transfer using physical cables or wireless short-range radio technologies, such as Bluetooth or Wi-Fi. Objects can also be moved between devices by copying them to a portable storage medium, such as a memory card, a USB memory stick, or an external hard drive, that is then removed and attached to another device. Another common practice to move objects between devices is to send them as e-mails or other kinds of messages to oneself. Further, it is possible to move objects by saving them in a network folder on a file server and then opening them with another device. More recently, similar cloud-based storage services have become popular. In addition to general-purpose cloud storage services, many cloud-based content-centric applications support transferring content of specific types, for example, photographs, music, or web pages, between devices. Beyond moving objects manually, there is also a wide variety of solutions that automatically synchronize data between devices.

In the field of human-computer interaction research, the general problem of cross-display object movement has been studied in the context of multi-monitor computer setups, large composite displays, collaborative interactive spaces, tabletop displays, and ad-hoc ecosystems of mobile devices (Nacenta, Gutwin, et al. 2009). A wide variety of different methods for moving visual objects between displays has been proposed in the prior literature. Perhaps the most widely known cross-display object movement method suggested in the human-computer interaction literature is the **Pick-and-Drop** method by Rekimoto (1997). In Pick-and-Drop, the user can pick up an object by touching it with a digital pen and then drop it by repeating the touch action on another display. Pick-and-Drop has been implemented in many different variations, including variations that work with finger touch instead of a digital pen. In Touch & Interact (Hardy and Rukzio 2008), the user can move objects between their phone and another display by touching the display with the corner of their phone.

Other methods suggested in the human-computer interaction literature include Conduit (N. Chen, Guimbretiere, and Sellen 2012) that uses the metaphor of transferring the object through the user's body. The user first designates the target display by touching it with the non-dominant hand and then selects the object to be moved with the dominant hand on the source display. Objects can also be moved from one display to another by making a throw touch screen gesture towards the target (Geißler 1998). Many variations of the basic throw gesture, for example, Drag-and-Throw and Push-and-Throw (Hascoët 2003), have been suggested, aiming to provide better control over and to improve the accuracy of the throw. Further, objects can be manipulated across displays by using pointing gestures, for example, with laser pointers (Olsen and T. Nielsen 2001).

Drag-and-Pop (Baudisch et al. 2003) extends traditional drag-and-drop to work across multiple displays by bringing the possible targets from other displays near the object when the user starts to drag it. This allows the user to complete the movement action on the same display. In ConnecTables (Tandler et al. 2001), the displays of two devices positioned side-by-side form a single continuous display area and objects can be directly dragged from one display to another. Objects can also be transferred between device displays by stacking the devices on top of each other (Girouard, Tarun, and Vertegaal 2012). In Pocket Transfers (Mäkelä et al. 2018), the user can move objects from large displays in the environment to their personal mobile device while keeping the device in a pocket or in a bag. Several different interaction modalities are supported, including touch, mid-air gestures, gaze, and their multi-modal combinations.

In ARIS (Biehl and Bailey 2004), the user can manage objects across screens on a radar-like map view showing all displays in proximity. Synchronized clipboards (Miller and Myers 1999) are shared between multiple devices, so that an object copied to the clipboard on one device can be pasted on another device. In Conductor (Hamilton and Wigdor 2014), the user may broadcast objects to all other devices as cues which can then be acted upon or ignored. In MediaBlocks (Ullmer, Ishii, and Glas 1998), objects can be bound to physical tokens that can be moved between displays.

Nacenta, Gutwin, et al. (2009) present a cognitive model of cross-display object movement. The model identifies four different processes in the cross-display ob-

ject movement interaction: first, the user transforms the task requirements or environmental constraints into an intention to move an object to a certain target display; second, the user formulates an action plan for moving the object to the target; third, the user executes the movement; and fourth, the user monitors and adjusts the movement action through a feedback loop. These four processes are not necessarily strictly sequential but they may partly happen in parallel.

Based on this model, Nacenta, Gutwin, et al. (2009) propose a taxonomy that classifies cross-display object movement methods on three conceptual levels: referential domain level, display configuration level, and control paradigm level. The referential domain level relates to how the user and the system reference to different displays. The references can be either spatial or non-spatial (for example, textual names, colors, shapes, or hierarchies). The display configuration level relates to the mapping between the physical arrangement of the displays and the input model of the interaction method. There are three possible mappings: planar mapping where the objects move across displays as if they were all arranged on a two-dimensional plane; perspective mapping which depends on the point of view of the user; and literal mapping that relies on the physical context. Finally, the control paradigm level relates to how the actual movement takes place. There are three different control possibilities: closed loop that allows the user to adjust the execution of the movement action before it is completed; open loop that does not; and intermittent that varies between closed loop and open loop depending on the phase of the interaction.

Despite the large variety of proposed methods for cross-display object movement, only a few studies have systematically compared multiple different methods against each other. These studies have mostly focused on moving objects between handheld devices and large tabletop or wall displays. Nacenta, Aliakseyeu, et al. (2005) report an early study that evaluates the performance of six different methods for moving objects with a stylus from a tablet computer to a tabletop display at different distances within and beyond hand's reach. Substantial differences between the methods were identified. Radar View and Pick-and-Drop were found to be the fastest methods and were also preferred by the participants.

Bachl et al. (2011) present a study where pairs of users solved collaborative tasks which involved moving objects from a shared tabletop display to personal tablet computers. Three different one-way object movement methods were included in the study: pressing a button attached to the object, dragging the object on top of an icon representing the target device, and using the tablet as a lens to capture content beneath it on the tabletop display. The participants considered the slowest button method to be the easiest to use, while the fastest lens method was considered too complicated.

Scott, Besacier, and McClelland (2014) report another study where groups of users played a game that involved moving virtual cards between tablet computers and a large tabletop display. Three different cross-display object movement methods were included in the study: two variants of the Pick-and-Drop method and a method called Bridges where objects were moved through a virtual portal area shared between devices. All methods were found to effectively support the movement tasks. The participants' preferences were equally divided between the methods, with each method providing unique advantages that suited different

participants' playing styles.

Paay et al. (2017) present a study where individual users moved different shapes in both directions between handheld devices and large displays. The study compared four different cross-display object movement methods that combine touch and mid-air gestures: pinching, swiping, swinging, and flicking. Swiping and swinging were found to be the most successful methods.

## 2.6   Studies of Real-Life Multi-Device Use

In addition to the extensive effort to develop technologies, systems, and interaction solutions to enable and support multi-device use, a number of researchers have also investigated how people actually combine and use multiple devices together in real life situations.

The early studies on multi-device use by Oulasvirta and Sumari (2007) and Dearman and Pierce (2008) identified a multiplicity of **reasons and benefits** that motivate people to have and use multiple devices instead of a single one. First of all, different devices have different characteristics and capabilities. This makes it beneficial to have multiple devices as different devices are suitable to different tasks, physical environments, and social situations. For example, when traveling, a person might choose to carry a lightweight but less powerful device with them, while at home, they might select a heavier but more powerful device. The setup overhead, that is, the effort needed to prepare the device for use, relative to the length of the task is also an important consideration when people decide which device to use: for short tasks, a person might pick a less powerful device with low setup overhead, while for longer tasks, a more powerful device might be preferred despite a high setup overhead. Having several devices may also help to organize and separate tasks: for example, a person might have one device for work-related use and another device for home use. In some situations, there may be no single device that could provide all the functions and data needed to complete the task, and this may force a person to use multiple devices. Some people decide to have extra devices as fallback devices, for example, as data or battery backups. People may also have personal preferences and habits that influence their decisions to use different devices for different tasks: for example, when a person purchases a new device, they might still keep and use the old device for certain tasks.

**Practices and workflows** in using multiple devices have been found to vary between different individuals and professional groups (Dearman and Pierce 2008; Karlson et al. 2010; Santosa and Wigdor 2013). People tend to divide tasks between devices, assigning each device a specific role within the workflow (Grudin 2001; Dearman and Pierce 2008; Santosa and Wigdor 2013). In addition to serial patterns of multi-device use (Karlson et al. 2010), also parallel patterns have recently become more common (Santosa and Wigdor 2013; Google 2012; Müller, Gove, and Webb 2012).

Accessing and **managing content** across different devices has been found to be one of the key challenges and sources of frustration in multi-device use (Oulasvirta and Sumari 2007; Dearman and Pierce 2008; Müller, Gove, and Webb 2012).

People who actively use several devices need to constantly make decisions related to content management, such as where to store content and how to transfer and synchronize it between devices. Despite a large variety of commercially available solutions (see Section 2.5), people still commonly encounter problems in accessing content between devices (Santosa and Wigdor 2013). While various cloud-based storage and synchronization solutions have recently become popular, they have many limitations regarding reliability, privacy, capacity, and connectivity, and therefore they supply only partial solutions to the problem (Santosa and Wigdor 2013). People have been found to assemble their own personal patchworks of solutions to manage content between their devices by combining multiple different tools and approaches, often in creative ways (Dearman and Pierce 2008). In addition to content, other kinds of information such as interaction histories should also be synchronized between devices (Oulasvirta and Sumari 2007; Dearman and Pierce 2008; Kane et al. 2009).

In mobile context, it may require significant physical effort and planning to manage the configurations of devices that a person carries with them and uses in different situations. Oulasvirta and Sumari (2007) found that rather than constantly making conscious decisions about which devices to carry and use, people develop habits and preparatory strategies to balance the risk of not having the right data or functionality available and the effort required to manage the devices. Oulasvirta and Sumari observed three different strategies: the conservative strategy of always carrying the same configuration of devices; the planful opportunism strategy of making just-in-case preparations for potential situations; and in some cases, the strategy of making careful advance planning and preparations. A **mobile kit** refers to a reasonably stable set of personal information devices that a person keeps together and carries with them while traveling (Mainwaring, K. Anderson, and Chang 2005). Jung et al. (2008) suggest examining the set of digital devices that a person uses as an ecology of interactive artifacts, in order to understand how people experience and strategize the use of interactive artifacts and the development of their ecologies over time.

Over the last few years, **smartwatches** have emerged as a new category of popular information devices. They provide an interesting addition to people's personal device ecologies, as they are the first category of wearable information devices that receives wider adoption. Some early studies have investigated the role of smartwatches in larger personal device ecologies. As a small display that is always worn on the user's body, a smartwatch provides a quick, unobtrusive, and less disrupting way to receive notifications of potentially important events, substituting for other devices in that purpose (Pizza et al. 2016; Cecchinato, Cox, and Bird 2017). Depending on the urgency of the event and the other devices available, the notification may then lead to the continuation of the task on another device (Cecchinato, Cox, and Bird 2017): for example, receiving a message on a smartwatch may lead the user to reply using their smartphone.

## 2.7   Research Gap

While there exists a large body of earlier research on topics related to multi-device interaction, we can identify some gaps in the earlier research. The research

presented in this thesis aims to fulfill these gaps.

First, considering the breadth and diversity of the phenomenon of multi-device use, there have been relatively few studies addressing how people use multiple devices in real life. Many of the earlier studies looking at the everyday practices of multi-device use are rather old as they have been made before the widespread availability and adoption of modern smartphones, tablets, and cloud services. The earlier studies have also primarily focused on technologically advanced users and use cases related to information work. Therefore, in order to update and broaden the understanding of the evolving practices and needs in multi-device use, there is a demand for further studies investigating how people utilize the new extended ecologies of devices and services, covering also more diverse groups of users and exploring beyond work-related use. This gap is addressed by research question RQ1 and by study S1.

Second, despite the large amount of research work that has been invested on studying systems, technologies, and interaction methods for multi-device use, there is still a lack of knowledge that would summarize the key experiential factors that significantly influence the user experience of multi-device interaction. Knowledge of these factors would be valuable in the design of interfaces and applications for multi-device use as these factors should receive particular attention in the design process. This gap is addressed by research question RQ2, based primarily on studies S2, S3, and S4 and supported by studies S1 and S5.

Third, wearable devices are emerging as an important new category of information devices that provide interesting opportunities for multi-user multi-device applications. While a wide range of different methods has been suggested for multi-device interaction in the earlier literature, most of the suggested methods have been designed for conventional devices, such as smartphones, tablets, and computers. As such, these methods may not be suitable or take full advantage of the features of wearable devices which differ from conventional devices in many ways. The development of multi-device interaction methods has also often been driven by technology and security considerations rather than by user experience. Therefore, there is a need for new research on natural and intuitive methods for multi-device interaction on wearable devices. This gap is addressed by research question RQ3 and by study S5.

# Chapter 3

# Research Approach and Methods

This chapter explains the approach and methodology that was followed in the research presented in this thesis. Section 3.1 first presents the general research approach. Section 3.2 then describes in more detail the research methods that were used for collecting and analyzing the data. Finally, Section 3.3 discusses related ethical considerations.

## 3.1  Research Approach

The research presented in this thesis was carried out within the **general framework of design science research** (Hevner and Chatterjee 2010; Vaishnavi and Kuechler 2015). According to the design science research approach, information technology is an artificial phenomenon (Simon 1996). As opposed to natural phenomena, information technology can not only be studied but it can also be created, and scientific research can contribute to each of these activities (March and Smith 1995). Therefore, scientific research on information technology consists of two distinct types of research: behavioral science research that studies information technology and human phenomena that surround its employment as they are, and design science research that creates new and better information technology (March and Smith 1995; Hevner, March, et al. 2004). These two types of research are illustrated in Figure 3.1.

**Behavioral science research** on information technology has its roots in natural science research, but instead of natural phenomena it studies human phenomena related to the application and use of information technology (Hevner, March, et al. 2004). It aims at describing, explaining, and predicting how and why people employ information technology in reality (March and Smith 1995; Hevner, March, et al. 2004). It consists of two kinds of research activities: activities that **develop** new knowledge about human phenomena related to information technology; and activities that **justify** that this knowledge is true (Hevner, March, et al. 2004).

```
┌─────────────────────────────┐         ┌─────────────────────────────┐
│   Behavioral Science        │         │    Design Science           │
│       Research              │         │       Research              │
│  ┌───────────────────────┐  │  Aids   │  ┌───────────────────────┐  │
│  │      Develop          │──┼────────→│  │       Build           │  │
│  └───────────────────────┘  │ Explains│  └───────────────────────┘  │
│         │  ┌────────────┐    │         │         │  ┌────────────┐    │
│         │  │ Descriptive│    │         │         │  │Prescriptive│    │
│         │  │ knowledge  │    │         │         │  │ knowledge  │    │
│         │  └────────────┘    │Provides │         │  └────────────┘    │
│         ↓                    │ targets │         ↓                    │
│  ┌───────────────────────┐  │←────────│  ┌───────────────────────┐  │
│  │      Justify          │  │  Tests  │  │      Evaluate         │  │
│  └───────────────────────┘  │         │  └───────────────────────┘  │
└─────────────────────────────┘         └─────────────────────────────┘
```
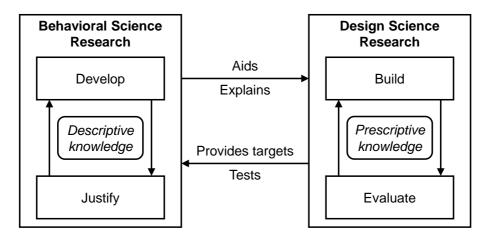
Figure 3.1: Behavioral science and design science in information technology research. Based on March and Smith (1995), Hevner, March, et al. (2004), and Gregor and Hevner (2013).

**Design science research** has its roots in engineering and design (Hevner, March, et al. 2004; Koskinen et al. 2011). Design science aims at improving information technology systems and processes (March and Smith 1995). It consists of two kinds of research activities: activities that **build** new knowledge about effective design and implementation of information technology; and activities that **evaluate** if this knowledge has value and utility over the earlier knowledge (Hevner, March, et al. 2004).

Behavioral science research and design science research produce different types of contributions. These contribution types are illustrated in Figure 3.2. Behavioral science research produces **descriptive knowledge** (Gregor and Hevner 2013). Descriptive knowledge consists of descriptions of human phenomena related to information technology, including simple observations and measurements as well as classifications and catalogs of these observations and measurements. It also contains knowledge, such as patterns, principles, and theories, that make sense of the relationships among the observed phenomena. Design science research produces **prescriptive knowledge** (Gregor and Hevner 2013). The simplest form of prescriptive knowledge is instantiations, for example, situated implementations of novel software products. More mature and abstract knowledge includes constructs (for example, concepts and terms), models (that is, propositions expressing relationships between constructs), methods (for example, algorithms and processes), design principles, and technological rules (March and Smith 1995; Gregor and Hevner 2013). These contributions of design science research are called **artifacts** to indicate that they can be transformed into material existence as artificially made objects (for example, hardware products) or processes (for example, operational software) (Gregor and Hevner 2013). The ultimate contributions of design science research are well-developed design theories that formalize the design knowledge in a specific domain. Note that in both behavioral and design science research, the contributions not only include developing and building new knowledge but also include justifying and evaluating existing knowledge.

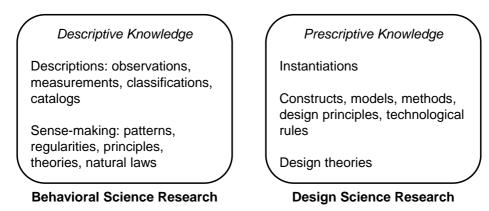Both behavioral science and design science are needed in information technology

| | |
|---|---|
| *Descriptive Knowledge*<br><br>Descriptions: observations, measurements, classifications, catalogs<br><br>Sense-making: patterns, regularities, principles, theories, natural laws | *Prescriptive Knowledge*<br><br>Instantiations<br><br>Constructs, models, methods, design principles, technological rules<br><br>Design theories |
| **Behavioral Science Research** | **Design Science Research** |

Figure 3.2: Contribution types of behavioral science and design science in information technology research. Adapted from Gregor and Hevner (2013).

research and they complement and support each other (March and Smith 1995). Figure 3.1 also illustrates the interactions between behavioral science research and design science research. Behavioral science research produces knowledge about human phenomena and this knowledge may **aid** design science research to build new artifacts, while design science research may **provide targets** for behavioral science research by building artifacts that give rise to new human phenomena that can be studied (March and Smith 1995). Behavioral science research may **explain** why artifacts built by design science research work, while the artifacts built by design science research **test** and may help to justify the knowledge produced by behavioral science research (March and Smith 1995).

This thesis applies the general framework of design science research in the context of **human-computer interaction research**. The ACM SIGCHI Curricula for Human-Computer Interaction (Hewett et al. 1992) defines human-computer interaction as "a discipline concerned with the design, evaluation and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them". As human-computer interaction research is interested both in how to create better interactive computing systems and in how people employ and use such systems, the design science research framework is well suited for research in human-computer interaction.

The framework of design science research also matches well with the goals and research questions of this thesis. As stated in Chapter 1, the overall goal of this thesis is to create new scientific knowledge to inform the design of future multi-device user interfaces, technologies, and applications. The first research question RQ1 focuses on understanding and explaining how people combine and use multiple information devices together in their everyday lives. It belongs to behavioral science research and produces descriptive knowledge about human phenomena related to multi-device use. The second research question RQ2 examines experiential factors that should be considered in the design of multi-device interaction methods and the third research question RQ3 explores natural methods for initiating multi-device interactions between wearable devices. These two research questions belong to design science research and produce prescriptive knowledge that aims at practical improvements in the user experience of multi-device use.

We will discuss the contributions of this thesis in more detail and reflect them on the general contribution types of behavioral and design science research in Subsection 6.1.

## 3.2 Research Methods

This section describes the research methods that were used in the research presented in this thesis. Subsection 3.2.1 first discusses the general research methods that were widely applied in this thesis. Subsection 3.2.2 then presents the methods that were used for collecting data and Subsection 3.2.3 the methods that were used for analyzing data.

### 3.2.1 General Methods

**Grounded theory.** While some of the individual studies in this thesis contain experimental elements, overall the research follows the grounded theory method (Strauss and Corbin 1990). Instead of forming hypotheses and then conducting experiments to test and prove them, the research presented in this thesis begins with the research questions, collects empirical data about them, and then lets the results emerge from the data through multiple rounds of systematic analysis without any preconceived hypotheses or concepts.

**Mixed methods.** All studies that this thesis consists of primarily use qualitative research methods. However, in each study qualitative research methods are complemented with quantitative research methods. By using this kind of mixed methods approach, it is possible to provide a better understanding of the research problem than using qualitative or quantitative research methods alone (J. W. Creswell 2012, p. 535). In the next two subsections, we explain in more detail the qualitative and quantitative research methods that were used for data collection and data analysis in this thesis.

### 3.2.2 Data Collection Methods

**Participant recruitment.** All studies presented in this thesis involved empirical research with human subjects. Overall, a total of 110 people participated in the studies, which necessitated a significant recruitment effort. The primary method of participant recruitment was posting advertisements on local mailing lists and social media groups. People who were interested in participating were prescreened against predefined selection criteria. In each study, we aimed to recruit a diverse and balanced group of participants with different backgrounds. In studies S2, S3, and S5 that involved groups of participants, we utilized snowball sampling and asked the selected participants to recruit their friends to participate in the study. The intention was that the participants would feel more comfortable if they knew some of the other participants in the same evaluation session. No participant knew all the other participants in any session, however.

**Diary studies.** A diary is a document created by an individual who maintains regular recordings about events in their life, at the time that those events occur (Alaszewski 2006). Diaries allow capturing rich data about interesting events and people's underlying feelings and motivations in naturalistic settings over extended periods of time, including also situations, such as private moments at home, where direct observation would be difficult or impossible (Lazar, Feng, and Hochheiser 2010, pp. 127–130). The diary study method was used in study S1 to collect data about the participants' practices of combining devices in everyday activities and tasks. While many diary studies today use online diaries, we decided to use a traditional paper diary as we wanted to avoid a situation that keeping the diary would interfere with the device use that was the subject of the study.

**Laboratory evaluations.** In design science research, it is essential to demonstrate via well-executed evaluations that the new artifacts improve the earlier practice (Hevner, March, et al. 2004). In this thesis, studies S2, S3, and S4 involved evaluations of multi-device interaction method prototypes. These evaluations were conducted in a laboratory environment and followed the tradition of evaluating the user experience of interactive prototypes in human-computer interaction research: after being introduced to the prototype, the study participants were asked to perform a set of given tasks with the prototype while data was collected through observation, measurements, interviews, and questionnaires. These evaluations can be seen as controlled experiments with the hypothesis of the prototype interaction methods being better than the previously known methods (Koskinen et al. 2011, p. 60). On the other hand, methodologically these evaluations resemble usability testing (J. Nielsen 1993, ch. 6). However, there are some important differences between the industrial practice of usability testing and the laboratory evaluations presented in this thesis (Lazar, Feng, and Hochheiser 2010, pp. 254–255). In the industry, usability testing is an engineering technique that is used to improve the user interface of a specific product during the product development process. It aims at identifying the most serious flaws in the product's user interface so that they can be corrected in the next product iteration. Usability testing is driven by pragmatic schedule, resource, and business considerations, and it aims at producing information that is "good enough" in the shortest amount of time and using the fewest resources possible, while optimizing trade-offs (Wixon 2003). In this thesis, however, the laboratory evaluations aim at producing scientific knowledge that is highly reliable and valid and that can be generalized over a broad range of different systems and applications. Therefore, the evaluations in this thesis were done more rigorously than typical usability tests: for example, the number of participants was higher and the data collection and analysis was done more systematically and thoroughly. Evaluations in studies S2 and S3 involved groups of four to six participants which required particularly careful planning and preparations.

**Elicitation studies.** Elicitation studies (M. Nielsen et al. 2004; Wobbrock et al. 2005) aim to gather interaction methods directly from end users, emphasizing the actions that people would naturally do. The participants of an elicitation study are first presented with the end effect of an operation and are then asked to perform the action that caused it. Elicitation studies have been particularly popular in producing sets of touch screen and other types of gestures. In this thesis, the elicitation study methodology was used in study S5 to generate multi-device interaction methods for wearable devices. Also study S3 that explored group

binding strategies in medium-sized groups included elements of the elicitation study methodology. In prior research, elicitation studies have been primarily conducted with individual users. In this thesis, the elicitation studies were conducted with groups of users and the methodology was adapted to multiple participants (Jokela, Rezaei, and Väänänen 2016).

**Interviews.** One of the primary data collection methods that was used in every study of this thesis was participant interviews. Interviews allow gathering deep and detailed data about person's opinions, feelings, and perceptions, encourage reflection and consideration, and enable flexible exploration of new insights and perspectives that may emerge during the interview. All interviews in this thesis were semi-structured and conducted face-to-face. In studies S1 and S4, the interviews were individual interviews, while in studies S2, S3, and S5, the interviews were group interviews with four to six participants. Various techniques, such as diaries, physical artifacts, and scenarios, were used during the interviews to ground the discussion on reality and to improve the quality of the collected data.

**Questionnaires.** In addition to interviews, all studies presented in this thesis included written questionnaires to gather quantitative data. Questionnaires allow efficient collection of well-structured data where each participant is equally represented and which can be analyzed using statistical methods. In studies S2, S3, and S4 that involved laboratory evaluations of interactive prototypes, two validated questionnaires, AttrakDiff and NASA-TLX, were used. **AttrakDiff** (Hassenzahl 2004) measures the attractiveness of interactive products over four dimensions: pragmatic quality, identification, stimulation, and perceived attractiveness. **NASA-TLX** (Hart and Staveland 1988) measures the subjective workload experience when performing a task over six dimensions: mental demand, physical demand, temporal demand, performance, effort, and frustration. In each study, the basic NASA-TLX questionnaire was extended with a few study-specific dimensions, such as learnability or security, to collect richer data. In addition to the generic AttrakDiff and NASA-TLX questionnaires, study-specific questionnaires were developed to address particular research themes as needed. These study-specific questionnaires were validated with pilot participants before the actual study was started.

### 3.2.3   Data Analysis Methods

**Content analysis.** In studies S4 and S5, standard content analysis techniques (Lazar, Feng, and Hochheiser 2010, ch. 11) with emergent coding were used for analyzing qualitative interview data. The analysis was done collaboratively by two or three researchers to improve the reliability and validity of the results (see Section 6.2 for a more detailed discussion about the reliability and validity of the research presented in this thesis). The interview responses were first transcribed based on audio or video recordings. Several researchers then examined the data to develop initial lists of coding categories. These categories were compared, discussed, and tested with subsets of data. This process was iteratively repeated until a consolidated coding hierarchy that all researchers agreed upon was reached. The coding hierarchy was then used for the actual coding of the entire data set. Finally, the resulting categorization of the data was processed into more general results. Microsoft Excel was used as the primary software tool to support content

analysis work (Meyer and Avery 2009). Similar content analysis techniques were also used in study S1 to categorize the user-reported multi-device use cases and in study S5 to categorize the user-suggested multi-device interaction methods for wearable devices.

**Affinity diagramming.** Affinity diagramming (Holtzblatt, Wendell, and Wood 2004; Lucero 2015) is a qualitative data analysis technique where each individual data point is recorded on a tangible paper note and the notes are then iteratively arranged on a real physical wall based on their affinity to each other, resulting in an increasingly hierarchical representation of the data. The method is well suited for collaborative data analysis, as it supports parallel work and creation of a shared interpretation of the data. In this thesis, affinity diagramming was used in studies S1, S2, and S3 to analyze large amounts of little structured interview and observational data. The MixedNotes tool (Jokela and Lucero 2014) was developed during this thesis work for efficient preparation of paper notes for affinity diagramming sessions.

**Quantitative analysis.** Questionnaire responses and other quantitative data was analyzed using statistical methods. Basic quantitative analysis, such as descriptive statistics, was done with Microsoft Excel. For more advanced statistical analysis, such as comparative ANOVA tests, IBM SPSS software was used.

## 3.3   Research Ethics

The term ethics refers to the moral standards or values by which human conduct is judged (Rosnow and Rosenthal 1997, p. 115). Applied to research, ethical guidelines allow us to judge the morality of scientific conduct (Rosnow and Rosenthal 1997, p. 115). In Finland, where the research presented in this thesis was done, ethical guidelines for human-computer interaction research are defined by the Finnish National Board on Research Integrity (TENK). According to the TENK guidelines (Finnish National Board on Research Integrity 2018), formal ethical review of a research plan by the ethical committee should only be done if the planned research may cause physical or mental harm to the research subjects, if the research involves children, or if the research deviates from the principle of informed consent. An ethical review can also be done if it is requested by a publication forum, funder, or international co-operation partner. As the research presented in this thesis did not meet any of these conditions, no formal ethical review of the research plan was done.

In the research presented in this thesis, we followed the generally accepted ethical principles of human-computer interaction research which are also recommended by TENK. We carefully designed each study to be a comfortable experience to the study participants and to avoid causing any potential harm to them. Before the study, we explained every participant what were the purpose and objectives of the study, who was doing the study, what was the study procedure, what data was collected, how the data was processed and stored, and how the study results would be used. We emphasized that participation in the study was voluntary. The participant could then make an informed decision about whether they wanted to participate in the study or not. Finally, we carefully handled the collected data

to make sure it was treated confidentially and reported anonymously.

# Chapter 4

# Research Process

This chapter explains the process that was followed in the research presented in this thesis. Section 4.1 first gives an overview of the research process. Sections 4.2–4.6 then describe in more detail the five empirical studies that form the core of this thesis.

## 4.1 Overview

The research presented in this thesis consists of five empirical studies. Figure 4.1 illustrates the relationships between the studies. Table 4.1 summarizes the key characteristics of the studies.

The first study S1 was a user study that investigated how people currently combine multiple devices in their everyday lives. It aimed to provide a general understanding of the current practices and challenges in multi-device use. The study



Figure 4.1: Relationships of the studies.

| S | P | RQ | N | Data Collection | Data Analysis |
|---|---|---|---|---|---|
| S1 | P1 | RQ1 (RQ2) | 14 | Diary study | Affinity diagramming, content analysis, quantitative analysis |
| S2 | P2 | RQ2 | 24 | Laboratory evaluation | Affinity diagramming, quantitative analysis |
| S3 | P3 | RQ2 | 24 | Laboratory evaluation, (elicitation study) | Affinity diagramming, quantitative analysis |
| S4 | P4 | RQ2 | 18 | Laboratory evaluation | Content analysis, quantitative analysis |
| S5 | P5 | RQ3 (RQ2) | 30 | Elicitation study | Content analysis, quantitative analysis |

Table 4.1: Summary of the key characteristics of the studies (S = Study, P = Publication, RQ = Research Question, N = Number of participants).

primarily contributed to research question RQ1 but it also supported research question RQ2.

The next two studies S2 and S3 addressed the problem of device binding in small and medium-sized groups. Study S2 was a laboratory study that compared three different methods for group binding. Based on the results of study S2, study S3 suggested a novel more flexible method for group binding and evaluated it with a laboratory study, with special emphasis on the patterns and strategies of applying the method in medium-sized groups. The fourth study S4 focused on the challenge of cross-display object movement. It was a laboratory study that compared three different methods for moving virtual objects between displays of personal mobile devices. The studies S2, S3, and S4 contributed to research question RQ2.

Finally, the fifth study S5 looked into a future where wearable devices are widely used and enable collaborative interactions. The study gathered interaction methods for common multi-device interaction tasks on wearable devices from groups of end users using the elicitation study methodology. The study primarily contributed to research question RQ3 but it also supported research question RQ2.

The following subsections summarize the motivation and objectives, the participants, and the data collection and analysis methodology and procedure for each study. The results of the studies are presented in Chapter 5.

## 4.2 Study S1: A Diary Study on Combining Multiple Information Devices in Everyday Life

While a wide variety of different technological solutions have been proposed to support multi-device use, only a few studies have addressed how people actually use multiple information devices together in their everyday lives. Many of these earlier studies have been made before the wide-spread adoption of modern smartphones, tablets, and cloud services and have focused on technologically oriented users and use cases related to information work. Therefore, we decided to conduct a new study on the practices and challenges in using the new extended ecologies of devices and services, aiming to address diverse groups of users and to explore beyond work-related use. The objective of the study was to provide a rich picture of the current and emerging real-life multi-device use and to broaden the understanding of the evolving practices and needs.

The study primarily contributed to research question RQ1. It also supported research question RQ2. The study has been reported in detail in publication P1.

**Participants.** We recruited a total of 14 people (six female, eight male) living in Southern Finland to participate in the study by posting an advertisement on local mailing lists and social media groups. We selected a diverse group of people who owned multiple devices and actively used them for a wide range of different applications and services. Seven of the participants were students, while the other seven were professionals working in different fields. Regarding education and professional background, four participants had their primary background in information technology, while the remaining ten participants represented a wide variety of other professions.

**Data collection.** Each participant filled in a workbook about their multi-device use during a one-week self-reporting period. The workbook included of a diary in which the participant was asked to report in detail all situations and tasks where they used multiple information devices together. The workbook also contained other tasks in which the participant was asked to provide information about their devices, their typical device use, and their ideal multi-device setup. After the self-reporting period, the participant was interviewed for approximately one hour based on the information in the workbook, including detailed discussion of two or three different situations of multi-device use that the participant had reported in the diary. The interviews were audio recorded.

**Data analysis.** The interview data was transcribed based on the audio recordings and analyzed by building an affinity diagram in a series of interpretation sessions. The diary data was separately analyzed and the reported use cases were categorized into different patterns of multi-device use with standard content analysis techniques. Finally, matrices describing relationships between devices, tasks, and contexts of use were constructed based on the data from the other tasks in the participants' workbooks.

Figure 4.2: Touching another device to bind it into a multi-device ecosystem.

## 4.3   Study S2: A Comparative Evaluation of Touch-Based Group Binding Methods

Before a group of users can engage in collaborative interactions with their devices, they must first bind their devices together into a multi-device ecosystem, in order to allow exchange of data and coordination of interactions between the devices. To make spontaneous multi-device interactions attractive, the device binding process should be fast and easy, but it should still provide sufficient security. In this study, we wanted to compare touch and proximity-based methods (see Figure 4.2) for group binding against more conventional methods based on scanning and passwords. We also wanted to explore different ways to divide the group binding task between the participants and to investigate possibilities to define the order of the devices during the group creation process to enable spatial interactions.

The study contributed to research question RQ2. The study has been reported in detail in publication P2.

**Apparatus.**   We designed three different group binding methods called Seek, Ring, and Host (see Figure 4.3). The Seek method represented the conventional approach used, for example, in network games, and was based on scanning for device identification and passwords for authentication. Both the Ring and Host methods used touch for device identification and authentication. The main difference between Ring and Host was that Ring was peer-based, distributing the group creation task equally between all participants, while Host was leader-driven, concentrating the group creation task on a single participant. Additionally, Host utilized device gestures for some interactions. The Host method was based on the EasyGroups method (Lucero, Jokela, et al. 2012). While all the methods were generic, we decided to study them in the context of a simple photo sharing application in order to provide a more realistic goal for the group creation task in the
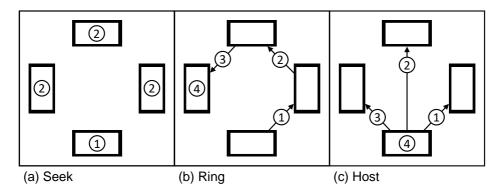
Figure 4.3: Three group binding methods (arrows represent touch actions between devices and numbered circles indicate the order of actions): (a) Seek: a leader creates a group and shares a password (1), which is then entered in parallel by the other participants (2). (b) Ring: one person starts the application and touches the next device to their right to add it to the group (1), then others continue to add the next person to their right (2, 3), and the last person completes the group (4). (c) Host: a leader starts the application, adds people to the group by touching all devices in counter-clockwise order (1–3), and puts the device on the table to complete the group (4).

evaluation. The photo sharing application was a simplified version of the Pass-Them-Around application (Lucero, Holopainen, and Jokela 2011): it allowed the users to browse a collection of photos stored in their own devices and supported spatial interactions of throwing photos between devices. We built prototypes of the three group binding methods on Nokia N9 mobile devices running the MeeGo operating system. Bluetooth-based radio technology was used to detect touch interactions between devices. The prototypes were fully functional with real network communication, except for the security protocols, which were only simulated in the user interface.

**Participants.** We recruited six groups of four participants (a total of 24 participants) for the evaluation by posting an advertisement on a local mailing list. Eight of the participants were female and 16 male. The participants represented a variety of different backgrounds, with eight participants having a software engineering background, 10 other technical background (for example, mechanical engineering), and six non-technical background (for example, administration or linguistics). All participants were active smartphone users and six of the 24 participants had used a Nokia N9 before the study.

**Data collection.** We organized a series of six evaluation sessions in a usability laboratory. In each session, there were four participants and a moderator present. All groups evaluated all methods in a systematically varied order. After being introduced to the first method, the participants were asked to form groups with it and then evaluate it using the AttrakDiff and extended NASA-TLX questionnaires. The same procedure was then repeated for the second and the third method. At the end of the session, the moderator interviewed the participants about their experiences with all the methods. The moderator also showed the participants pictures of three different scenarios and asked them to consider what

would be the most appropriate method for creating a group in each scenario. The objective of using the scenarios was to encourage the participants to think about different situations and environments and their social and physical characteristics. The durations of the evaluation sessions varied between 100 and 120 minutes. The sessions were video recorded.

**Data analysis.** The video recordings of the evaluation sessions were reviewed and notes written about relevant observations. The data recorded on the notes was then analyzed by building an affinity diagram in a series of interpretation sessions. A quantitative analysis of the NASA-TLX and AttrakDiff responses was done separately.

## 4.4 Study S3: An Evaluation of a Touch-Based Group Binding Method

The results of the study S2 suggested that group binding methods should be flexible, allowing people to adopt different group creation strategies in different situations. Still, many of the device binding methods proposed in prior research have been very specific, enforcing a detailed procedure that has to be followed exactly. Therefore, in this study, we wanted to test a group binding method that was more flexible and gave the users a lot of freedom to adapt it to different situations. Further, the earlier research on device binding has focused on individual users or small groups of up to four participants. But as the group size increases, the overall group binding process becomes more parallel and a much wider variety of different approaches and strategies becomes possible. Therefore, in this study, we also wanted to investigate group binding in medium-sized groups of six users in order to better understand the different strategies, group behavior, and communication and collaboration within a larger group

The study contributed to research question RQ2. The study has been reported in detail in publication P3.

**Apparatus.** We designed a novel group binding method called FlexiGroups (see Figure 4.4) that was based on the results of study S2. The FlexiGroups method is highly flexible, enabling the users to apply a broad range of different group formation strategies. While the method itself is generic and can be used in many different applications, we again decided to study the method in the context of a multi-user photo sharing application based on Pass-Them-Around (Lucero, Holopainen, and Jokela 2011). We implemented a prototype of the FlexiGroups group binding method on Nokia N9 mobile devices using similar technologies as in study S2.

**Participants.** We recruited four groups of six participants (a total of 24 participants) for the evaluation by posting an advertisement on a local mailing list. Six of the participants were female and 18 male. The participants represented a variety of different professions, with six participants having a software engineering background, 13 having other technical background (for example, mechanical engineering), and five having a non-technical background (for example, teaching or photography). All participants were experienced smartphone users, but only
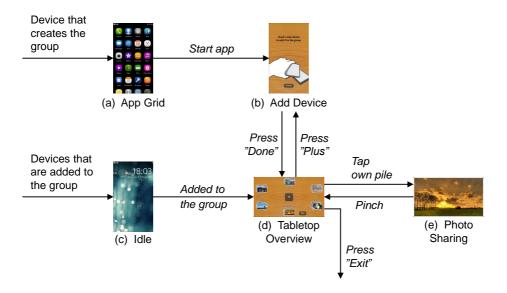
Figure 4.4: FlexiGroups group binding method. One person first creates a new group by initiating the application on their device from the Application Grid view (a). The application starts in the Add Device view (b) which allows the person to use their device to touch other devices and add them to the group. When completed, the person can press the "Done" button to move to the Tabletop Overview view (d) which shows an overview of the devices in the group and allows changing the order of the devices if needed by dragging the device icons to the right positions on the screen. Other persons do not have to start the application manually—their devices can remain in the Idle view (c). When a device is touched and added to the group, it automatically launches the application which starts in the Tabletop Overview view. Any member of the group can press the "Plus" button in the Tabletop Overview view to move to the Add Device view and add new devices to the group. Note that also the added members can add new members to the group and that several persons can add new devices in parallel. When the group is ready, all persons can enter the Photo Sharing view (e) by tapping their device icon in the Tabletop Overview view and start sharing photos between devices. At any time, any person can pinch to zoom out in the Photo Sharing view and return to the Tabletop Overview view and check and manage the current members of the group. To leave the group, a person can press the "Exit" button in the Tabletop Overview view. The group continues to run on the other devices until the last member exits the group.

two of them had used a Nokia N9 device before the study.

**Data collection.** We organized four evaluation sessions in a usability laboratory. In each session, there were six participants and a moderator present. After being introduced to the FlexiGroups method, the participants were asked to set up a group using the method. The participants repeated the task several times. The moderator then asked the participants to evaluate the method using the Attrak-Diff and extended NASA-TLX questionnaires. Finally, the moderator interviewed the participants about their experiences with the method. The average duration of the evaluation sessions was 60 minutes. The sessions were video recorded.

**Data analysis.** The video recordings of the evaluation sessions were reviewed and notes written about relevant observations. Diagrams that recorded the sequences of participant actions in every group creation attempt were also drawn. The data recorded on the notes was analyzed by building an Affinity Diagram in a series of interpretation sessions. A quantitative analysis of the AttrakDiff and NASA-TLX questionnaires was done separately.

## 4.5 Study S4: A Comparative Evaluation of Cross-Display Object Movement Methods for Personal Mobile Devices

In situations where several information devices are combined and used together, there is often a need to move virtual visual objects from the display of one device to the display of another device. In this study, we wanted to compare three different methods for moving objects between device displays. In particular, we wanted to focus on situations where a person wants to move visual objects between the displays of their personal mobile and portable devices (such as smartphones, tablets, and laptops) located within hand's reach. Further, while the earlier comparisons of cross-display object movement methods have concentrated on a single (often abstract) task in a laboratory environment, we wanted to explore the participants' preferences of different methods more broadly over a range of different situations, tasks, and environments. Overall, we were interested in the different factors that influence the users' preferences and selections of the methods in different situations, considering not only the pragmatic aspects (such as efficiency and number of errors) but also the hedonic aspects (such as emotional and social factors).

The study contributed to research question RQ2. The study has been reported in detail in publication P4.

**Apparatus.** To study these questions, we selected three methods, Tray, Transfer Mode, and Device Touch (see Figure 4.5), that featured significantly different approaches to the cross-display object movement problem. The Tray method represented the conventional model of using an intermediate storage, similar to a shared clipboard or a network folder, for moving objects between displays. The Transfer Mode method was based on creating virtual connections between devices. The Device Touch method relied on real-world touch interactions between tangible

(a)           (b)           (c)

Figure 4.5: Different cross-display object movement methods: (a) Tray: a user moves objects by dragging them to and from a virtual tray shared between all devices. (b) Transfer Mode: a user first sets up a connection between devices by making a screen gesture and then shares objects between devices by simply tapping objects on the displays. (c) Device Touch: a user moves objects by physically touching the screen of another device with the corner of the phone.

physical devices. It was inspired by the earlier work of Schmidt et al. (Schmidt, Chehimi, et al. 2010; Schmidt, Seifert, et al. 2012). While all the three methods were generic and could be used to move a wide range of different kinds of objects in different applications, we decided to prototype them in the context of a simple multi-device photo browsing and management application that allowed the user to browse, view, and re-arrange photos stored on the device as well as to move photos between devices using the three cross-display object movement methods.

We built prototypes of the object movement methods on Windows 8 and Windows Phone 8 platforms. The prototype system used in the evaluation consisted of three devices that were carefully selected to represent different device categories used today: a tablet with an 11.6-inch screen, a smartphone with a large 4.7-inch screen, and a smartphone with a small 3.8-inch screen. The implementation of Device Touch followed the general description given by Schmidt et al. (Schmidt, Chehimi, et al. 2010; Schmidt, Seifert, et al. 2012). However, instead of camera-based detection used by Schmidt et al., we used a novel implementation approach of applying capacitive sensing technology to detect the device touch events (see Figure 4.6) which allowed the method to be used with standard commercially available devices featuring capacitive touch screens.

**Participants.** We recruited a total of 18 participants for the evaluation by posting an advertisement on a local mailing list. Nine of the participants were female and nine male. The participants represented a variety of different professional backgrounds, with ten participants having a background related to information technology and eight having other background (for example, a teacher, a civil engineer, and a nurse). All participants were experienced smartphone users and fourteen of them were also active tablet users.

**Data collection.** We organized a series of evaluation sessions in a usability laboratory. In each session, there was a participant and a moderator present. Each participant evaluated all the three methods in a systematically varied order. After being introduced to the first method, the participant was given two different tasks in which they were asked to move groups of photos between the phones and

Figure 4.6: Device Touch implementation using capacitive sensing technology. Stylus heads are attached to the corners of the phone and conductive tape is used on the device sides to connect the user's body to the stylus.

the tablet in both directions and to arrange the photos in a specified order. After completing the tasks, the participant evaluated the method by filling in NASA-TLX and AttrakDiff questionnaires. The same procedure was then repeated for the second and the third method. The evaluation session was concluded with an interview contrasting the methods with each other. The participant was also presented with four scenarios involving cross-display object movement and was asked to rate the applicability of each method to each scenario. The scenarios were based on the results of study S1 and represented real-life situations where people today use multiple devices. The evaluation tasks and the interviews were recorded with a video camera.

**Data analysis.** The video material was transcribed and analyzed using standard content analysis techniques. A quantitative analysis of the extended NASA-TLX and AttrakDiff questionnaires was done separately.

## 4.6 Study S5: An Elicitation Study to Collect Methods for Group Binding and Cross-Display Object Movement on Wearable Devices

As wearable devices become more popular, situations where there are multiple persons present with such devices will become commonplace. In these situations, wearable devices could enable collaborative tasks and shared experiences between co-located persons through multi-user applications. But to make such applications possible, the wearable devices need to provide support for multi-device
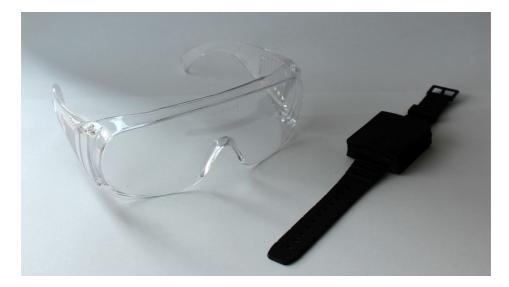
Figure 4.7: Device surrogates used in the study. Smartglasses (left) were simulated with ordinary 3M safety spectacles. Smartwatches (right) were simulated with custom mock-ups built by attaching a Casio watch band to a small block of polystyrene foam.

interactions. While a wide range of solutions has been proposed to multi-device interaction tasks, most of the existing multi-device interaction methods have been designed for conventional devices such as computers, phones, and tablets, and as such they are not necessarily applicable to wearable devices which are far more personal and intimate. The existing multi-device interaction methods may also not take advantage of the unique features of wearable devices that could enable more natural and innovative solutions (Jokela, Chong, et al. 2015). In this study, we wanted to gather interaction methods for two common tasks in multi-device interaction, namely group binding and cross-display object movement, on wearable devices directly from end users. Regarding devices, we wanted to focus on the two most common wearable device categories today: smartglasses and smartwatches. We adopted the elicitation study approach and asked groups of participants to come up with their own techniques for these tasks. Overall, we wanted to approach the problem from a perspective of what pairs and small groups of people would naturally and intuitively do to connect their wearable devices into a group or to move virtual objects between their wearable devices. The study was inspired by similar studies by Chong and Gellersen (Chong and Gellersen 2011; Chong and Gellersen 2013).

The study primary contributed to research question RQ3. It also supported research question RQ2. The study has been reported in detail in publication P5.

**Apparatus.** We provided the participants with simple mock-up devices of smartglasses and smartwatches that acted as surrogates of real devices (see Figure 4.7).

**Participants.** We recruited eight groups of four participants for the study by posting an advertisement on local mailing lists and social media groups. Two persons canceled their participation, which reduced the total number of partic-

ipants to 30. 14 of the participants were female and 16 male. 14 participants had educational backgrounds related to information technology, while the other 16 participants represented a wide variety of different professions (for example, a school teacher, a nurse, and a bus driver). Seven participants had tried smart-glasses and 11 participants had tried a smartwatch or other smart wrist device. 11 participants were wearing eyeglasses continuously and six as needed, while 13 participants did not have eyeglasses. Eight participants were wearing a wristwatch every day and seven sometimes, while 15 participants did not have a wristwatch.

**Data collection.** We organized a total of eight elicitation study sessions in a usability laboratory. In each session, there were three or four participants and a moderator present. Every group generated methods for all the four device and task combinations in a systematically varied order. After being introduced to the first device and the first task, the participants took turns to suggest interaction methods how the task could be achieved with the devices. The participants immediately tried out the method with the mock-up devices, first in pairs and then as a single group of three or four persons. After trying out the method, the participants evaluated the method in terms of practicality and pleasantness by filling in a paper form and could also provide verbal comments about the method. The same procedure was then repeated for the other task with the same device, and after that for both tasks with the other device. At the end of the study session, the moderator briefly interviewed the participants for general comments about the wearable devices and tasks included in the study. The durations of the sessions varied between 75 and 120 minutes. The study sessions were recorded with a video camera.

**Data analysis.** The proposed methods were documented and categorized based on the metaphor and modality of interaction. Participant comments about the methods as well as interview responses were transcribed and analyzed using standard content analysis methods. A quantitative analysis of the evaluation responses was done separately.

# Chapter 5

# Results

This chapter summarizes the research results of this thesis. The results are organized to answer to the three research questions posed in Chapter 1: Section 5.1 addresses the first research question RQ1, Section 5.2 addresses the second research question RQ2, and Section 5.3 addresses the third research question RQ3. As the objective of this chapter is to provide an overview of the most important research results, we refer to the original publications included as an appendix for the complete results.

## 5.1 Combining Multiple Information Devices in Everyday Activities and Tasks

In Section 1.4, the first research question was defined as:

> RQ1 How do people combine multiple information devices in everyday activities and tasks?

This research question was addressed in study S1. In this section, we summarize the main results of the study S1 based on publication P1. We begin by giving an overview of the typical device ecologies that people have and describe the most common device types and their characteristic practices of use. We then discuss the different patterns of combining multiple information devices together, including sequential use and parallel use and their sub-patterns. Finally, we identify three levels of decisions that determine which devices are used in a particular situation: acquiring, making available, and selecting the devices to use.

### 5.1.1 Device Ecologies

People today possess large collections of information devices. The set of devices that a person actively uses can be seen as a **personal device ecology** (Jung
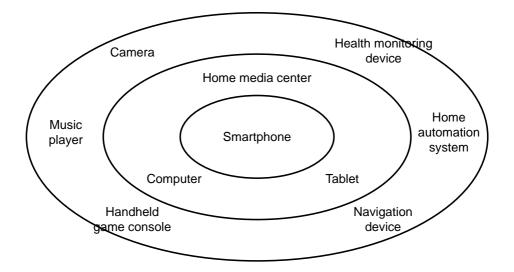
Figure 5.1: Most commonly used device types. The rings illustrate the commonness and importance of the different device types in personal device ecologies, with the more common and important device types located closer to the center of the figure.

et al. 2008). Figure 5.1 illustrates the most commonly used device types today. In most personal device ecologies, the single most important and central device is the smartphone. Beyond the smartphone, the second tier of devices typically consists of computers, tablets, and home media centers. We will next discuss the characteristic usage practices of these four most commonly used device types.

For most people, **smartphone** is the most important device and their everyday lives depend on it. As people carry their smartphones almost always with them, the phones allow them to be reached and keep them up to date anywhere they go. The smartphones are used for a wide variety of different functions, including calendar, e-mail, messaging, phone calls, web browsing, camera, social media, music, and calculator. However, the use of smartphones is focused on retrieving and entering small amounts of information and the smartphones are not considered capable of handling large amounts of information or complex tasks. While many applications and services have smartphone versions, they often lack features or do not work properly.

As laptop and desktop **computers** have been replaced by other devices in many consumption and entertainment oriented tasks, their use is increasingly characterized by more complex tasks, which include entry of long texts, creation of complicated content, detailed work, and handling of large amounts of content. These tasks are often related to work or studies and are considered as important but serious and somewhat boring. Computers represent the most powerful devices available and they provide options beyond what the other devices can offer. Among modern smartphones and tablets, computers feel like legacy devices from the old world but people simply cannot manage without them.

**Tablets** are primarily used for searching for and consuming content and for enter-
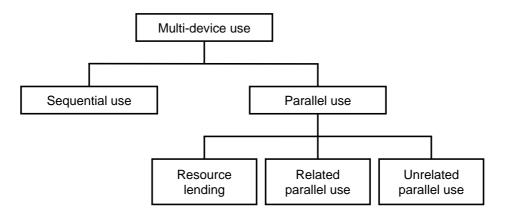
Figure 5.2: High level categorization of main patterns of multi-device use.

tainment, such as for reading, playing games, and watching videos. Some people consider tablets as smartphones with larger screens, while others think that they are like laptops but more lightweight and instantly available for use. Two common patterns of tablet use exist: (1) a shared family tablet at home for multiple purposes; and (2) a personal tablet that partly replaces the laptop in mobile use. However, the tablets are not considered suitable for entering long texts or doing complex or detailed work. In general, the tablets are still looking for their role and their use involves more experimentation than the use of the other devices.

We define a **home media center** to consist of a television or other large screen together with game consoles, home theater PCs, and other media appliances connected to it. The home media center is the heart of social entertainment at home. It provides the largest screen in the household supporting collective viewing and sharing of TV programs, movies, games, music, and personal photos and videos.

In addition to these four most commonly used device types, people also use a wide variety of other devices. However, the use of the other device types is much less frequent and is more focused on specific use cases. Of the other device types, the most often used are digital cameras, which are typically used in pre-planned situations where high quality photographs and videos are desired. Other device types include music players, handheld game consoles, navigation devices, health monitoring devices, and home automation systems.

## 5.1.2   Main Patterns of Multi-Device Use

Figure 5.2 illustrates a categorization of the main patterns of multi-device use. On the highest level, multi-device use can be divided into sequential use and parallel use. In **sequential use**, a person changes the device during the task. In **parallel use**, a person uses several devices at the same time. The cases of parallel use can be further divided into three sub-patterns: resource lending, related parallel use, and unrelated parallel use. In **resource lending**, a person performs the actual task on a single device, but the device borrows some resources from other devices. In **related parallel use**, a person is working on a single task using multiple devices and all devices are involved in the same task. In **unrelated parallel**
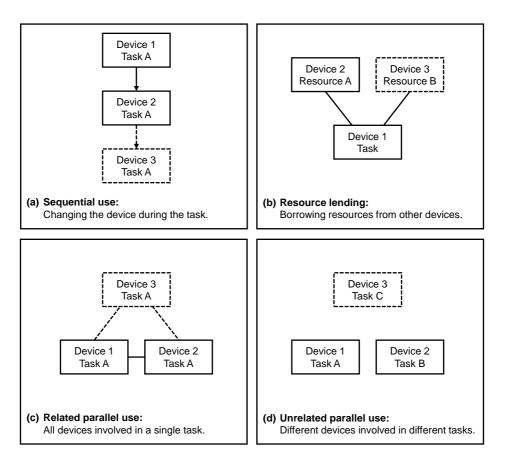
Figure 5.3: Different patterns of multi-device use.

**use**, a person is working on multiple tasks simultaneously and different devices are involved in different tasks.

It is possible to hierarchically combine these main patterns of multi-device use to create more complex **composite patterns** of use. For example, sequential use may involve phases where the person is using a single device but also phases where the person is using multiple devices together, applying hierarchically the pattern of parallel use. Similarly, in parallel use the person may switch one of the devices to another device, applying hierarchically the pattern of sequential use.

In the following subsections, we will discuss each of the main patterns of multi-device use and their sub-patterns in more detail. Figure 5.5 presents a complete view of the different patterns of multi-device use.

### 5.1.3 Sequential Use

In sequential use (see Figure 5.3 (a)), a person changes the device during the task. The person may change the device once or multiple times in a sequence. We can identify several subpatterns of sequential use depending on the reason
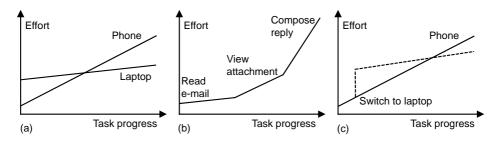
Figure 5.4: (a) Different devices have different setup efforts and task efficiencies. (b) Changing task characteristics may influence device efficiency. (c) Switching devices requires additional effort that may exceed the benefits.

that triggers the person to switch devices:

**Task change.** The character of the task changes, which makes another device better suited for the task and results in a device change. For example, a person may search for a phone number on a tablet and then call to the number with their phone.

**Context change.** As the physical environment or the social situation changes, another device is considered as more appropriate for the task. For example, a person may start reading a page on a smartphone in a bus and then continue reading the same page on a tablet at home.

**Task and context change.** Both the task and the context change at the same time. In these cases, there is typically a workflow consisting of a sequence of clearly separate steps, each done at a different location. For example, a person may search for a recipe with a laptop computer in the living room, and then move to the kitchen, open the recipe on a tablet, and start cooking.

**Task re-evaluation.** The initial evaluation of the task is proven incorrect as more information is learned about the task, ultimately leading to a device change. For example, a person may start writing an e-mail message with a smartphone, but as the message grows longer and more complicated the person may switch to a laptop computer and finish the message with it.

The dilemma of whether to change the device or not can be illustrated with Figure 5.4 (a). The horizontal axis represents the progression of a task, while the vertical axis represents the user effort. Each device requires an initial setup effort, represented by the y-intercept of the line, before it can used to work on the task. Each device also has a certain efficiency, represented by the slope of the line. As illustrated in Figure 5.4 (b), the efficiency of the device may change as the characteristics of the task change. In order to switch the device (see Figure 5.4 (c)), the user has to do additional effort. While the person may gain improved performance after the switch, it may take a long time to outdo the switching effort. The problem may be further complicated by the difficulty of predicting the future evolution of the task.

## 5.1.4   Parallel Use

In parallel use, a person uses two or more devices simultaneously. The pattern of parallel use can be further divided into three sub-patterns: resource lending, related parallel use, and unrelated parallel use.

### Resource Lending

In resource lending (see Figure 5.3 (b)), the actual task is performed on one device, but this device borrows some resources from other devices. We can identify different subpatterns of resource lending depending on the the types of resources that are borrowed:

**Borrowing input resources.** One device is used to give input to another device. For example, a person may use a remote control application on their smartphone to operate a home media center from a distance.

**Borrowing output resources.** One device presents some output through another device. For example, a person may play a video clip on a laptop computer and connect the laptop to a television to display the video on the television screen.

**Borrowing input and output resources.** All interaction with a device occurs through another device. For example, a person may use their laptop computer to organize music on their smartphone.

**Borrowing network connection.** One device connects to the network through another device. For example, a person may use the hotspot feature on their phone to connect their laptop to the Internet.

### Related Parallel Use

In related parallel use (see Figure 5.3 (c)), a person is working on a single task using multiple devices. All the devices are involved in the same task. We can identify several subpatterns of related parallel use based on the motivation to use multiple devices:

**Multiple views to task.** Different devices provide different views to the task. For example, a person may watch a movie on a television and search for information about the movie on a tablet; or a person may install software updates on a tablet and view instructions on a laptop computer; or a person may write a document on a laptop computer and check translations of words with a dictionary application on a smartphone.

**Remote collaboration.** In remote collaboration over a real-time communication link, one device is assigned to handle the communication and other devices are used to access relevant content and applications. For example, a person may call another person with a smartphone to agree a meeting time and use a laptop computer to check their calendar during the call.

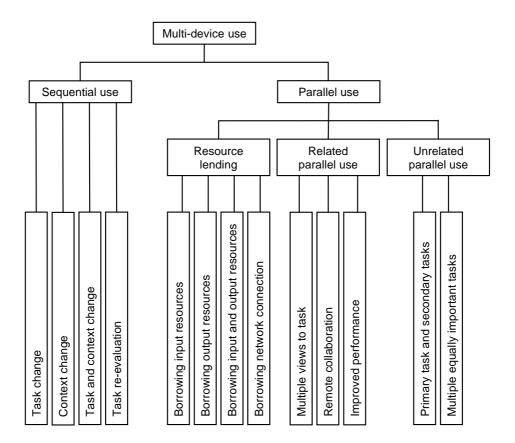**Improved performance.** A complex task is distributed across multiple devices

Figure 5.5: Complete categorization of patterns of multi-device use.

to execute it more efficiently. For example, a person may use multiple desktop computers for parallel rendering of a complex 3D graphics scene in order to complete the rendering task faster.

**Unrelated Parallel Use**

In unrelated parallel use (see Figure 5.3 (d)), a person is working on several tasks simultaneously using multiple devices. Different devices are involved in different tasks. We can identify several subpatterns of unrelated parallel use depending on the priorities of the tasks:

**Primary task and secondary tasks.** There is a primary foreground task on one device and secondary background tasks on other devices. For example, while working on a document on a laptop computer, the user may listen to background music with their smartphone.

**Multiple equally important tasks.** There are several equally important tasks in parallel and each task is carried out on a different device. For example, the user may receive two phone calls simultaneously, one on a personal phone and another on a work phone.

### 5.1.5 Forced Multi-Device Use

So far, we have focused on patterns where people voluntarily decide to use multiple devices, for example, in the hope of improved performance, efficiency, or convenience of use. These patterns cover the majority of the real-life use cases we have investigated. However, in some cases the use of multiple devices may be involuntary as the person is forced to use several devices. This may be because of:

**Technical issue.** Technical limitations or problems may prevent completing a task with a single device. For example, all the required applications may not be available for the device that was used to start the task.

**Content access.** There is a need to access content that is stored on another device. For example, a person may be preparing a presentation on a laptop computer and may want to attach a photograph that they have captured with their smartphone.

**Policy of use.** There exist policies, rules, or legislation that force the use of multiple devices. For example, a corporate security policy may dictate that certain operations can only be performed on authorized devices.

**Device availability.** The device that was used to work on a task is no longer available. For example, the device may have ran out of battery or a person may have forgotten to take the device with them.

These factors may apply broadly to different patterns of multi-device use, including both sequential and parallel use.

### 5.1.6 Deciding Which Devices to Use

Owning multiple devices creates the problem of choosing which devices to use in a specific situation. We can identify three levels of decisions which determine which devices are used: deciding which devices to acquire, deciding which of the devices to make available in a specific context, and deciding which of the available devices to actually use. Each level of decisions narrows down the available options until finally the devices to use are left. The device selections determine the pattern of multi-device use: a decision to change the device results in a sequential use pattern, while a decision to use two or more devices simultaneously results in a parallel use pattern. We will next discuss these different levels of decisions in more detail.

**Acquiring new devices.** Interoperability with the devices that a person already has is a major factor to consider when deciding which new devices to obtain. On the other hand, other factors such as price or curiosity to try and opportunity to learn new devices and systems may override it. The level of influence a person has over the acquisition decisions varies: for example, people have more control over devices they purchase for their personal use than over devices that are provided to them by their employers. Overall, people cannot fully control their device environments as they are constrained by the choices of other people and organizations,

such as their employers, educational institutions, clients, and friends.

**Making devices available.** For fixed devices such as home media centers or desktop computers, decisions on which devices to make available in different contexts primarily relate to the physical locations of the devices, for example, in which room to place the device at home. For mobile devices such as smartphones and tablets, the decisions are related to which devices to take with you in a mobile situation. Many people have a fixed **mobile kit of devices** (Mainwaring, K. Anderson, and Chang 2005; Oulasvirta and Sumari 2007) that they take with them every day. Heavier devices such as laptops and special devices such as USB drives are typically taken based on the expected need. In addition to computing devices, people carry with them a wide variety of auxiliary devices such as chargers, headsets, cables, and memory sticks.

**Selecting the devices to use.** Decision on which of the available devices to use in a specific situation can be based on several criteria. The device characteristics to consider include both user interface capabilities (such as display size, pointing device, and the level of multi-tasking support), technical capabilities (such as processing power, storage capacity, network connection, and camera quality), and physical characteristics (such as the size and weight of the device). For tasks requiring entry of long texts, text entry capability, especially the availability of a physical QWERTY keyboard, strongly influences the device selection. As already discussed in Subsection 5.1.3, another major factor influencing the decision on which device to use is the effort required to start using the device. As discussed in Subsection 5.1.5, a person may be forced to use specific devices, for example, if the required applications or content are available only on certain devices. People also develop habits of using certain devices for certain tasks or in certain contexts of use. Finally, people consider the context of use, including both social aspects (such as the acceptability of using a certain device in a certain social situation) as well as physical aspects (such as the available space).

## 5.1.7   Summary

Instead of a single information device, people today have personal device ecologies consisting of multiple devices. For most people, the most important device is the smartphone which forms the centerpiece of the personal device ecology. It is used for a wide variety of simple information retrieval and entry tasks in everyday life. The second tier of devices consists of computers which are primarily used for complex and large tasks; tablets which are primarily used for consuming content and for entertainment; and home media centers which support social entertainment and sharing of content. In addition to these most common device types, people also use many other device types but their use is not as frequent and is focused on specific use cases.

Situations where people combine and use multiple devices together can be divided into sequential use and parallel use. In sequential use, a person changes the device during the task. In parallel use, a person uses several devices at the same time. The cases of parallel use can be further divided into resource lending, related parallel use, and unrelated parallel use. In resource lending, a person performs the actual task on a single device, but the device borrows some resources from

other devices. In related parallel use, a person is working on a single task using multiple devices and all devices are involved in the same task. In unrelated parallel use, a person is working on multiple tasks simultaneously and different devices are involved in different tasks.

Selecting which devices to use in a specific situation consists of a series of decisions on different levels. Each level of decisions narrows down the available options until the devices to use are left. On the highest level, people decide which devices to acquire for their use. On the next level, people decide which devices to make available in different contexts. On the lowest level, people decide which of the available devices to actually use. These decisions are influenced by many factors, including device attributes, social and physical context, economical considerations, and personal aspirations and preferences.

## 5.2 Design Considerations for Multi-Device Interaction Methods

In Section 1.4, the second research question was defined as:

> RQ2 Which experiential factors should be considered when designing methods for multi-device interaction?

This research question was addressed in studies S2, S3, and S4. In addition, studies S1 and S5 provided supporting results. In this section, we summarize the main results of these studies based on publications P1–P5. We first discuss the experiential factors that should be considered in the design of multi-device interaction methods in general. We then describe in detail a set of experiential factors that are either unique to multi-device interaction or that should receive particular attention in the design of multi-device user interfaces.

### 5.2.1 Introduction

Multi-device user interfaces present a difficult design problem. Compared to traditional single-device user interfaces, distributing the interaction across multiple devices adds significant complexity to the interface. This makes the design of multi-device user interfaces a challenging task. Therefore, guidance to support the designer in the design process of multi-device user interfaces would be useful.

Prior research has identified many experiential factors that should be considered in the design of interactive systems in general. Basic factors to consider include efficiency, learnability, possibility of user errors, and other usability factors, as well as the functions, features, and overall utility of the system. In addition to these pragmatic aspects, non-pragmatic aspects of the system, such as beauty, innovativeness, social value, and reputation, may have a strong impact on the user experience. Beyond the different aspects of the interactive system itself, other important factors to consider include the user characteristics, such their skills,
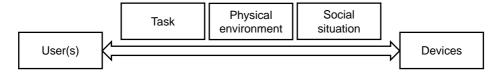
Figure 5.6: High-level categorization of experiential factors to consider in the design of multi-device interaction methods.

prior experiences, motivations, and expectations, as well as the characteristics of the physical, social, and cultural environment of use. All these generic factors should naturally also be considered in the design of user interfaces for multi-device applications. However, as these factors have been thoroughly presented and analyzed in prior literature, we do not elaborate them in more detail in this thesis. A good summary of the generic experiential factors is provided, for example, by Olsson (2012).

Rather, for the rest of this section we focus on experiential factors that are either unique to multi-device user interfaces or that should receive particular attention in the design of multi-device interfaces. These factors have been identified in a series of studies comparing different methods for two common tasks in multi-device interaction, device binding and cross-display object movement, and the factors have been found to influence people's perceptions and preferences of multi-device interaction methods in different situations of use. While the factors are based on studies of device binding and cross-display object movement, we believe that they apply also to many other multi-device interaction tasks and can be helpful in the design of interaction methods for these other tasks as well.

### 5.2.2 Experiential Factors to Consider in Design of Multi-Device Interaction Methods

We can identify several factors which have significant influence on the user experience of multi-device interaction and which should therefore be carefully considered when designing user interfaces for multi-device applications and services. As illustrated in Figure 5.6, these experiential factors can be grouped into a few broader categories, including factors related to the user(s), factors related to the devices, factors related to the task, factors related to the physical environment, and factors related to the social situation. In this subsection, we discuss each of these categories and the related experiential factors in more detail.

As described in Section 5.1, there is a wide range of different situations where people combine and use multiple devices together. The presented experiential factors can greatly vary between different situations of use, and this makes it challenging to design generic multi-device interaction methods that would be optimal for any given situation. Rather, the optimal multi-device interaction methods strongly depend on the situation of use and therefore people tend to prefer different methods in different situations.

**Factors Related to User(s)**

Regarding the users of the multi-device system, the number of users is a factor that should be carefully considered in the user interface design.

**Number of users.** A multi-device application may have a single user, or it may have multiple users (from two or three to dozens or even more). In multi-device applications that have multiple users, special attention should be paid on supporting group work and collaboration between users: coordinating and synchronizing actions, making decisions, and agreeing on and following a common strategy (P2; P3). An essential enabler for efficient group work is awareness of the task status. A multi-device application should provide users feedback about the other users and the overall task status, either directly through the application user interface, or indirectly, for example, through interactions, such as touching or gestures, that can be easily perceived by all users. It is also important to keep all users engaged in the interaction, as people may easily become bored or distracted if they cannot do anything but wait for others to complete the task (P3). Ideally, the multi-device interaction methods should support building a sense of community and positive team morale within the user group (P2; P3; P5). This can be achieved, for example, by providing meaningful ways for all users to contribute to the interaction, by encouraging interactions that several users can do together, and by offering opportunities for self-expression and playfulness.

While multiple users set many new requirements for the design of a multi-device application, a larger group of users may also offer new design opportunities (P3). In a large group, the activity becomes more parallel as several individual users or sub-groups of users may work on different sub-tasks simultaneously. A larger group of users may make possible new approaches and strategies for solving the task. For example, the user interface may support viral or contagion-based interactions (P3; P5; Terrenghi, Quigley, and Dix 2009) that spread from one device to another like an infection and effectively distribute the workload across different devices and users. As a practical example, a group binding method supporting viral interactions could allow User A to first add User B to the group; User B could then add User C while User A could simultaneously add User D to the group; all four Users A, B, C, and D could then continue and add users E, F, G, and H to the group, respectively; and so the group membership would continue to spread from one device to another.

**Factors Related to Devices**

Regarding the devices and the entire multi-device system, the following experiential factors should receive particular attention in the design of the user interface: number of devices, device characteristics and capabilities, perceived device affordances, and system robustness.

**Number of devices.** By definition, multi-device interaction involves at least two devices but the number of devices may also be higher (up to dozens of devices or even more). However, the situation of two devices is very common in many applications and it can often be resolved in a simpler way than situations with more devices. In such cases, a multi-device application should identify if there

are only two devices involved, and if that is the case, offer interactions optimized for two devices (P4). For example, in cross-display object movement between two devices, it may be possible to implicitly determine the target device, so that the user does not have to explicitly indicate it.

If the number of devices is large, the task may easily become time-consuming and physically exhausting if the user has to repeat the same interactions for each device or device pair. For example, if a group binding method allows adding only a single device to the group at a time, the user has to repeat the interaction four times to create a group of five devices. In such situations, it should be considered if the interaction method could enable interactions between multiple devices simultaneously (P3; P5; Chong, Mayrhofer, and Gellersen 2014). For example, a group binding method could allow the users to collect all of their devices together and bind them to a group with a single interaction.

**Device characteristics and capabilities.** Different devices have different physical characteristics and technical capabilities that influence the suitability of different interaction methods (P1; P4). For example, devices with large displays are better suited for methods based on detailed visualizations and touch interactions than devices with small displays, and small and lightweight devices are better suited for methods based on physical device movement and manipulation than large and heavy devices. Important device characteristics to consider include user interface capabilities (such as display size, accuracy of the pointing device, and level of multi-tasking support), technical capabilities (such as processing power, storage capacity, network connection, and camera quality), and physical characteristics (such as the size, weight, and shape of the device). For tasks that require entry of text content, the text input capability, especially the availability of a large physical QWERTY keyboard, is an important consideration.

If the devices constituting a multi-device system have very different characteristics, different methods may be needed for interactions on different devices (P4; P5). For example, in a cross-display object movement method based on device touch interactions, the user could move a photo object from a smartphone to a tablet by taking the smartphone (source device) in their hand and touching the tablet to *put* the object there. To move a photo object from the tablet to the smartphone, the user could similarly take the tablet in their hand and touch the smartphone but that might be inconvenient as the tablet is a larger and heavier device. Instead, the user could take the smartphone (which is now the target device) in their hand and touch the tablet to *pick* the object from there. However, such an asymmetric behavior requires additional user effort and may easily confuse the user as the user actually needs to learn two different methods and needs to select the right method for each interaction.

**Perceived device affordances.** In addition to the actual device characteristics and capabilities, the perceived device affordances also play an important part in the design of multi-device interaction methods (P5). Perceived affordances (Norman 1988, pp. 9–11) refer to actions that the user perceives as possible means to use the device. Ideally, the methods for multi-device interaction should be natural and intuitive, building on people's spontaneous perception of the characteristics and capabilities of the involved devices. In particular, wearable devices are often seen as extensions of the user's body, and the perceived affordances of a wearable

device are strongly influenced by the characteristics of the body part that the device is attached to, including the natural physiological capabilities and social functions of the body part and the technical features that are presumed possible for a device attached to that body part (for example, what kind of body signals can be measured from that location) (P5). As a practical example, the users may anticipate that interaction with smartglasses occurs through gaze pointing or thought commands, as both the eyes and the brain are located in the user's head near the position of the glasses. On the other hand, the users may anticipate that interaction with a smartwatch occurs through actions done with hands, for example, through hand gestures and touch actions, even though it might be equally possible to detect such actions with cameras mounted on the smartglasses.

**System robustness.** When designing interaction methods and user interfaces for multi-device and multi-user systems, it is important to consider robustness in real-life conditions (P3). While many methods can work well in theory or with mock-ups, in reality, multi-user multi-device applications are complex distributed systems. As multiple devices are involved, there is an increased risk of technical issues: the devices may fail to detect each other, the network connections between devices may be lost, and the software may crash on any device. On the other hand, the user groups may not always act in coordinated or systematic ways: some users may not be aware of the procedure they should follow, while others may be unable to do so, for example, because they are momentarily occupied with other tasks such as phone calls. The composition of the user group may also change during the interaction as some users may have to leave and new users may join. Therefore, the interaction methods should not expect an exact procedure to be followed. The methods should be flexible and robust, allowing people to adapt them to the changing needs of the situation and to recover from technical failures.

### Factors Related to Task

Regarding the task that is performed with the multi-device system, the following experiential factors should receive particular attention in the design of the user interface: task size and content sensitivity.

**Task size.** The task performed with the multi-device system may be simple and consist of a single quick action, for example, moving a photo from one device to another. The task may also be very complicated and consist of a long sequence of actions, for example, organizing a large collection of photos between multiple devices. In such a complex task, the same basic actions, for example, moving a photo from one device to another, may be repeated numerous times.

Simple tasks, such as moving a single photo from one device to another, are often a common special case and the multi-device interaction methods should offer efficient and optimized performance in such cases (P4). The user should be able to carry out simple tasks in a straightforward manner without complex initialization procedures such as activating certain technical functions or modes. Changing to another device usually also requires the user to do significant additional effort and therefore it should preferably be possible to perform simple tasks with the device that the user is currently actively using (P1). For example, if the user receives a

message in their smartphone, they should be able to compose a simple reply with the same device.

However, interaction methods optimized for simple tasks may not be efficient for complex tasks. If the task is long and complicated, even very laborious initialization operations can be justified if they improve the performance of a long sequence of subsequent actions needed to solve the actual task as the initialization cost can be allocated over a large number of actions (P4). For example, if the user needs to move a large number of photos between two devices, they might prefer to set up a special photo transfer mode between the two devices even if that requires additional effort, provided that activating the transfer mode then makes it more efficient to move individual photos between the devices, for example, because the user does not have to identify the target device for each photo separately. In a long task, it is also possible to outdo the device change effort if changing the device gives improved performance for the task, thus making the device change a more attractive option to the user (P1). For example, if the user receives a message in their smartphone and needs to compose a long reply, they might prefer to change the device and compose the reply with a laptop.

**Content sensitivity.** Some tasks may involve very sensitive content, such as personal photographs or work-related documents, and have therefore high requirements for privacy and security, while in other tasks, such as in a casual game, such requirements can be minimal (P1). The required level of security should be carefully considered early in the design of multi-device interaction methods as the security requirements may set constraints on the interaction design. The interaction methods should provide an adequate level of security but the user should not to be burdened with unnecessary security procedures (P2). From the security perspective, the step of device binding is very important as that is when the connections between the devices are established. If needed, the interaction methods have to enable proper device authentication, for example, to provide a way to exchange authentication strings over a secure auxiliary communication channel (P2). Some methods, such as viral methods, may make it challenging to implement strict control over access to the content (P3), while other methods, such as leader-driven methods where the interaction is driven by a single user, make it possible for that user to have strong control over the content (P2). Another security-related aspect to consider is the cost of user errors. For example, moving content objects between devices by throwing them with touch screen gestures may be inaccurate and easily result in content objects being sent to incorrect devices, which makes the throw interactions problematic in applications where the security requirements are high (P5).

### Factors Related to Physical Environment

Regarding the physical environment where the multi-device system is used, the users' ability to access the devices that the system consists of is a factor that should be carefully considered in the user interface design.

**Physical access to devices.** The user may have only restricted access to the devices that constitute the multi-device system. The user's access to the devices may be spatially restricted: for example, if the users are sitting around a table,

some devices may be on the other side of the table beyond the user's reach, or there may be some dinnerware on the table blocking the access (P2). The user's access to the devices may also be temporally restricted: for example, the user may start reading a web page with a computer at the workplace, then leave the computer at work, and continue reading the page with a family tablet at home (P4). In such situations, interaction methods that assume physical access to the devices, for example, methods based on device proximity, may make accomplishing the task difficult or even impossible.

### Factors Related to Social Situation

Regarding the social situation where the multi-device system is used, the following experiential factors should receive particular attention in the design of the user interface: social access to devices, social acceptability, and user roles. The social situation is especially important in multi-user applications, but it should also be considered in single-user applications which may be used in situations where there are other people present.

**Social access to devices.** In addition to spatial and temporal constraints, also social constraints may prevent a user from accessing some of the devices. In particular, there is a social barrier of taking another person's device, for example, a personal smartphone, into one's hands and using it, as that can be perceived as an invasion of the device owner's privacy and as a potential security risk (P1; P3; Uzun, Saxena, and Kumar 2011; Hang et al. 2012). Similar although milder concerns apply to interactions that involve using one's device to touch another person's device (for example, NFC interactions), and therefore, such interactions should only be done with the permission of the device owner (P2; P3).

**Social acceptability.** Many other social norms influence the appropriateness of interaction methods, especially in formal situations where people are not familiar with each other (P1). For example, hand gestures, such as pointing another person with a finger, or eye gestures, such as staring at another person, may be considered rude in many social situations (P5). Interaction methods based on physical touching or close proximity may feel socially uncomfortable between unfamiliar persons (P2; P5). In general, people prefer discreet interaction methods in public places with unknown persons present (P4).

**User roles.** In many social situations, people have different roles: for example, in a meeting, there may be a chairperson and a secretary. A group of people may consist of several subgroups that may have their own internal substructures: for example, in a game session, there may be two competing teams and each of the teams may have a captain. In some applications, the order of the devices may be important, for example, if the the people are sitting around a table playing a game and taking turns (P2; P3; Jokela, Chong, et al. 2015).

These social roles, group structures, and device orders should be considered and utilized in the design of multi-device interactions. The multi-device application should be able to capture and become aware of the user roles, preferably automatically or implicitly by monitoring the actions of the users. For example, it might be possible to determine the user roles based on the physical locations and
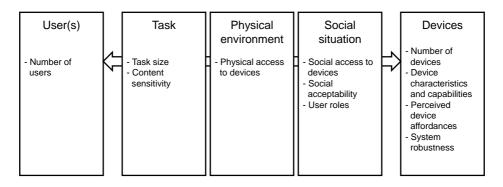
Figure 5.7: Complete categorization of experiential factors to consider in the design of multi-device interaction methods

arrangement of the devices, or based on the sequence of user actions in the device binding phase. Information about the roles could then be used to provide better experiences to the users, for example, to customize the user interface and to suggest relevant actions for each role. The user interface should be flexible and allow the users to change roles (P3): for example, if the chairperson or the team captain has to leave, another person should be able to take their role.

### 5.2.3   Summary

Compared to traditional single-device user interfaces, multi-device user interfaces present a challenging design problem as there are multiple devices involved and the interaction is distributed across them. A wide range of experiential factors should be considered in the design of multi-device interaction methods. In addition to generic factors, we can identify several factors that are either unique to multi-device interfaces or that should receive particular attention in the design of multi-device interfaces. These factors can be divided into a few broader categories: factors related to the user(s), factors related to the devices, factors related to the task, factors related to the physical environment, and factors related to the social situation. Figure 5.7 summarizes these categories and experiential factors.

Important factors to consider in the design of multi-device interfaces include the number of users and the number of devices involved. If the application has multiple users, special attention to support group work and collaboration between users is needed but many new design opportunities may also become possible. Regarding the number of devices, multi-device applications should offer optimized performance for the typically common special case of two devices, but they should also scale efficiently if the number of devices is large. Other device-related considerations include the device characteristics and capabilities, which affect the suitability of different interaction methods, and the perceived device affordances, which influence what kind of interactions the users anticipate and perceive as natural and intuitive. Multi-device interfaces should also be robust and flexible, allowing the users to adapt them to the changing needs of the situation and to recover from errors and technical failures. Regarding the task, the interface should enable the user to perform basic tasks in a simple and straightforward way but

also to carry out large and complicated tasks in an effective manner. The required level of security should be carefully considered early in the design process as the security requirements may set constraints on the user interface, and the level of security should match the actual need. The physical environment and social situation may restrict individual user's access to the devices that constitute the multi-device system and should be considered when selecting the interaction methods. Many other social norms also influence the acceptability and appropriateness of different interaction methods, such as gestures or touching. Finally, the roles and group structures of people in different social situations should be considered and utilized in the design of multi-device interfaces to provide better experiences to the users.

## 5.3  Natural Multi-Device Interaction Methods for Wearable Devices

In Section 1.4, the third research question was defined as:

> RQ3  How would a group of co-located persons naturally initiate multi-device interactions between their wearable devices?

This research question was addressed in study S5. In this section, we summarize the main results of the study S5 based on publication P5. We first discuss multi-device interaction methods on wearable devices in general. We then present a categorization of methods that groups of persons spontaneously suggested for initiating multi-device interactions between their wearable devices. The methods were collected using the elicitation study methodology.

### 5.3.1  Introduction

As wearable devices become more popular, they could enable collaborative tasks and shared experiences between co-located persons through multi-user applications. To make such applications possible on wearable devices, the devices need to provide support for multi-device interactions. While there exists a wide variety of multi-device interaction methods that have been designed for conventional devices, these methods may not be optimal for wearable devices which differ from conventional devices in many ways.

If we consider two common tasks in multi-user multi-device interaction, namely group binding and cross-display object movement, we can see that they share a similar structure (see Figure 5.8). Each task begins with a **preparation phase** which consists of various activities that are needed to initiate the task, including both technical actions, such as activating certain device functions or selecting certain data objects, and social actions, such as agreeing on a common target and a strategy to achieve it with the other persons. In the **target identification phase**, links between the devices are established, that is, it is indicated which of the devices are intended to participate in the new group or to receive the
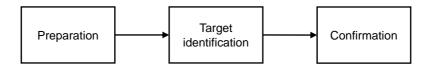
Figure 5.8: Phases of device binding and cross-display object movement interaction tasks.

object to be shared. In the **confirmation phase**, all parties of the interaction make sure that the intended action was successfully completed. Of these phases, the target identification phase constitutes the most characteristic phase of the interaction where the actual connections between the persons and the devices are formed. It is also the phase where the unique features of wearable devices could offer the most new design opportunities. Beyond group binding and cross-display object movement, target identification is also needed in many other multi-device interaction tasks.

As discussed in Section 5.2, the multi-device interaction methods for wearable devices should ideally be natural and intuitive, and they should build on people's spontaneous perception of the characteristics and affordances of the wearable devices. In the following subsection, we report a categorization of methods that groups of persons spontaneously suggested for target identification on wearable devices. The methods were collected using the elicitation study methodology.

### 5.3.2 Categorization of Methods

Figure 5.9 presents a categorization of methods that groups of persons suggested for target identification between their wearable devices in group binding and cross-display object movement tasks. The groups provided suggestions for the two most common wearable device categories today: smartwatches and smartglasses. Both the methods suggested for smartwatches and the methods suggested for smartglasses have been included in the same categorization. While the categorization is based on studies with smartwatches and smartglasses, we believe that it reflects the more general characteristics and perceived affordances of wearable devices, and that most of the suggested methods can also be used with other wearable device types, such as smart clothes, smart footwear, or smart jewelry.

The methods have been categorized primarily based on the metaphor of interaction. Secondarily, the methods have been categorized based on the interaction modality. We will next describe the categories in detail, and also report on the strengths and weaknesses of the methods in each category as evaluated by the groups that suggested them.

#### Spatial

The methods in the spatial category are based on the relative or absolute positions of the persons or the devices. The methods in this category can be further divided
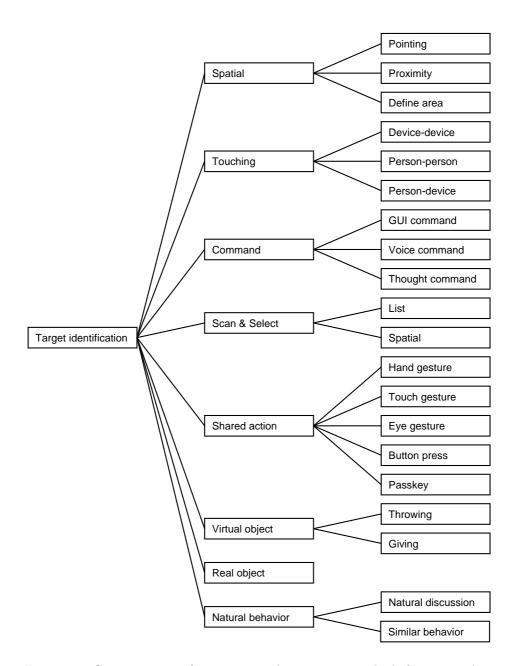
Figure 5.9: Categorization of user-suggested interaction methods for target identification in multi-device applications on wearable devices.

into three subcategories: pointing-based methods, proximity-based methods, and area-based methods.

**Pointing.** The target person or device is selected by pointing, for example, with gaze, with hand, with finger, or with a device. Pointing is a very easy and natural action, but it may be slow if the number of persons or devices is large and it may be considered rude in some social situations.

**Proximity.** The devices are capable of determining the closest persons or devices. The users can make selections by moving themselves or their devices closer to or further away from each other. Proximity-based methods are simple to understand, but they may be error prone if there are many people or devices present and moving very close to unfamiliar persons may feel socially inconvenient.

**Define area.** The user defines an area to select all persons or devices within that area. The area can be defined, for example, by setting the radius for a circle around the user, or by drawing the area boundaries on a map or in the real world. The area method makes it efficient to select a large number of persons or devices, but it may be imprecise in crowded locations.

### Touching

In touching-based methods, connections between persons or devices are created by bringing them into physical contact or very close proximity of each other for a brief period of time. Touch interactions are simple and accurate, but they require that the persons or the devices are within touching distance of each other and may be slow if the number of persons or devices is large. Touch interactions may feel socially uncomfortable between unfamiliar persons but they may also help to create a feeling of togetherness within the user group. The methods in this category can be divided into three subcategories: device-device touch, person-person touch, and person-device touch.

**Device-device.** Touching one device with another device. Device touch interactions (for example, NFC interactions) are familiar to many users, but some users may be concerned that touching may physically damage their devices.

**Person-person.** Two persons touching each other, typically shaking hands. Shaking hands is a familiar and natural gesture but it may feel too formal in casual situations. In a situation which involves a series of multiple interactions, shaking hands several times may feel strange.

**Person-device.** Person touching another person's device, for example, pressing the device touch screen with their finger.

### Command

The methods in the command category are based on giving direct commands to the system. The target persons or devices are referred to by using predefined identifiers such as names or contact cards. This allows users to prepare commands in advance and to refer to persons or devices that are not present, but

is limited to persons or devices that the user knows beforehand. The methods in this category can be divided into three subcategories: GUI-based commands, voice-based commands, and thought commands.

**GUI command.** Giving commands using a conventional Graphical User Interface (GUI). The GUI can be presented, for example, on a physical touch screen or on a virtual screen shown on an augmented reality display. GUI-based commands can offer a familiar, reliable, and discreet solution.

**Voice command.** Giving commands by voice. Voice commands are natural and well-known by many people, but can be annoying in some social situations, sensitive to background noise, and slow, and require the user to remember the names of the other people and devices.

**Thought command.** Giving commands by thought through a Brain-Computer Interface (BCI). Thought commands are novel and exciting and are perceived as effortless and discreet, but they may make it difficult to follow the task progress in multi-user applications as they provide no perceivable indications to others.

### Scan & Select

Scan & Select (see Section 2.4) is currently the most widely used device binding method in commercially available devices. It is a familiar, practical, and anticipated solution for many users, but it may also be perceived as mundane and unexciting. In a crowded location, scanning may take a long time and result in a large number of discovered devices which may be difficult to manage. The methods in this category can be divided into two subcategories: methods using a list representation and methods using a spatial representation for the discovered devices.

**List.** Discovered devices are presented as a list of textual names. It may be difficult to map the names to real-world persons and devices.

**Spatial.** Discovered devices are presented on a graphical map. The spatial representation may make it easier to map between targets on the screen and in the real world.

### Shared Action

The methods in the shared action category are based on a common action that all the target persons do, or that the user does on all the target devices. The actions are typically time constrained, that is, they have to be done simultaneously or nearly simultaneously. The actions can be exactly the same for each person or device, or they can form symmetric pairs, for example, if the first person pushes their hand forward, the second person may respond by pulling their hand backward. The shared action methods are simple and practical, are efficient in large groups, and make it easy to follow the task progress. As people often enjoy making physical actions together, shared action methods may help to strengthen the group cohesion. On the other hand, shared actions may appear as ridiculous

in some social situations and they can be insecure as external persons may easily observe and copy the actions. The methods in this category can be divided into five subcategories based on the type of the action.

**Hand gesture.** The shared action is a hand gesture, for example, waving the hand.

**Touch gesture.** The shared action is a touch screen gesture, for example, drawing a picture on the device screen.

**Eye gesture.** The shared action is an eye gesture, for example, performing a certain movement pattern with the eyes.

**Button press.** The shared action is a button press that is done simultaneously.

**Passkey.** The shared action is entering a passkey, for example, a textual password. Users may consider passwords as old-fashioned and it may be difficult to enter text on wearable devices.

**Virtual Object**

In the methods of the virtual object category, the logical concepts of the user interface, such as a group membership or a file, are represented as virtual objects that can be manipulated like real objects. The virtual objects can be purely imaginary invisible objects, or they can be visualized, for example, as graphical objects on a conventional touch screen or as augmented reality objects on an augmented reality display. As interactions with the virtual objects resemble interactions with real world objects, the methods in the virtual object category are natural and intuitive. They may also feel entertaining and even magical. The methods in this category can be divided into two subcategories based on the action that is used to share the virtual objects: throwing-based methods and giving-based methods.

**Throwing.** User grabs the virtual object and throws it towards the target person or device. While practical, throwing may be inaccurate and therefore also insecure. It may also be inefficient if there is a large number of target persons or devices.

**Giving.** User takes the virtual object in their hand and makes it available to others, for example, by offering it on their palm. Other persons can then grab the virtual object if they want. Compared to throwing, the giving-based methods can be more accurate, secure, and efficient.

**Real Object**

The methods in the real object category are similar to the methods in the virtual object category, but instead of virtual objects the logical user interface concepts are mapped to real physical objects. For example, the group membership may be represented by a t-shirt and the user may join in the group by dressing up the t-shirt. Like virtual object methods, real object methods are natural and intuitive. However, they require that suitable physical objects are available for

all the relevant user interface concepts.

### Natural Behavior

While the methods in the other categories are based on explicit user actions, in the methods of the natural behavior category the devices monitor the natural behavior of people and pro-actively execute commands on their behalf. Such methods may be very easy to use and intuitive, but they may also result in unintended actions and raise privacy issues related to continuous monitoring of people. The methods in this category can be divided into two subcategories: methods based on monitoring natural discussion and methods based on detecting similar behavior.

**Natural discussion.** Devices monitor people's natural discussion, identify persons based on their voices, and automatically execute commands based on the discussion.

**Similar behavior.** Devices monitor people's behavior in general and cluster into the same group everybody who behaves in a similar way, for example, moves together or does similar actions at the same time.

## 5.3.3   Summary

Wearable devices offer new opportunities to support collaborative tasks and shared experiences between groups of co-located persons through multi-user applications. To enable such applications, the wearable devices need to support multi-device interactions. While there exists a wide variety of multi-device interaction methods for conventional devices, these methods may not be optimal for wearable devices which differ from conventional devices in many ways. This creates a need for novel multi-device interaction methods for wearable devices. Ideally, these methods should be natural and intuitive, building on people's spontaneous perception of the characteristics and affordances of wearable devices.

Target identification is a key action that is needed in many multi-device interaction tasks. Elicitation studies allow gathering suggestions for target identification methods on wearable devices directly from end users. While some of the methods suggested by the users may appear futuristic, they may still be valuable as design inspiration. Analysis of the user-suggested methods for target identification on wearable devices shows that there is a large space of potential design approaches. The main approaches, categorized based on the metaphor of interaction, are spatial methods, touch-based methods, command methods, different variants of the Scan & Select method, shared action based methods, virtual object methods, real object methods, and methods based on natural behavior. Each of the approaches has different strengths and weaknesses. As discussed in Section 5.2, many factors related to users, devices, tasks, physical environment, and social situation should be considered when designing target identification methods for specific applications.

# Chapter 6

# Discussion

This chapter discusses the contributions and limitations of the research presented in this thesis as well as potential directions for future research. Section 6.1 first summarizes the key contributions, contrasts them to prior research, and discusses their practical utility. Section 6.2 then considers the reliability, validity, generalizability, and limitations of the research presented in this thesis. Finally, Section 6.3 presents future directions to continue and expand the research presented in this thesis.

## 6.1  Contributions

As articulated in Chapter 1, the research presented in this thesis aims to create new scientific knowledge to inform the design of future multi-device user interfaces, technologies, and applications. This thesis produces many contributions supporting that goal. In this section, we summarize the contributions of this thesis, addressing each of the research questions one at a time. We reflect the contributions of this thesis on the general discussion of the contributions of behavioral and design science research presented in Section 3.1. We also contrast the contributions of this thesis on the related earlier research described in Section 2 and discuss their practical utility.

The first research question RQ1 focuses on understanding and explaining how people combine and use multiple information devices together in their everyday lives. This research question belongs to behavioral science research and it aims to produce descriptive knowledge about human phenomena related to multi-device use. As discussed in Section 3.1, behavioral science research can produce two kinds of contributions: knowledge that describes human phenomena and knowledge that makes sense of the phenomena. Considering the knowledge of the first kind, this thesis contributes detailed descriptions of the current and emerging practices and challenges in multi-device use. It provides a rich picture of people's personal device ecologies and describes the most common device types and their characteristic conventions of use. It also presents a novel three-level catego-

rization of the patterns of combining multiple information devices together. The descriptive knowledge contributed in this thesis provides a solid foundation for further research on more specific topics related to multi-device use as well as for design and development of practical interfaces and applications that support using multiple devices together. While some earlier studies, such as the well-known studies by Dearman and Pierce (2008) and Oulasvirta and Sumari (2007), have also provided descriptions of the practices of multi-device use, most of the earlier studies are several years old and many significant changes have occurred since they were published: for example, new device categories such as smartphones and tablets have become popular, a wide variety of cloud-based services supporting multi-device use has been introduced, and the overall number of devices and the level of multi-device use have increased. Therefore, it is valuable to provide an updated view on the current practices of multi-device use. The more recent study of Santosa and Wigdor (2013) was done in parallel with this thesis: it provides many similar findings and the the two studies support and justify each other. While the study of Santosa and Wigdor focuses on information work, the research presented in this thesis explores also beyond work-related use.

In addition to descriptions and classifications of the current human practices of multi-device use, this thesis also provides theoretical contributions that aim to make sense and explain some of the observed practices. This thesis contributes a novel model of the user's decision to change the device, a key decision point in multi-device use. It also presents a three-level model of how users decide which devices to use in a specific situation. Earlier the problem of device selection has been studied in the mobile context by Oulasvirta and Sumari (2007). The work presented in this thesis validates and justifies the earlier work of Oulasvirta and Sumari, and extends it by addressing also device selection in non-mobile contexts and considering the acquisition of new devices. Overall, these theoretical contributions increase the understanding of some of the fundamental aspects of multi-device use and consequently also support the development of practical applications.

The second research question RQ2 examines experiential factors that should be considered in the design of multi-device interaction methods. This research question belongs to design science research and it aims to produce prescriptive knowledge that improves the user experience of multi-device use. Towards that goal, this thesis contributes a set of design principles in the form of experiential factors that should receive special consideration when designing methods for multi-device interaction. These design principles summarize the learnings from a series of studies on multi-device interaction methods. They provide practical guidance to designers and developers working with interfaces and applications that involve the use of multiple devices. In earlier related research, Chong and Gellersen (2012) present a categorization of factors that influence the usability of device binding methods. They identify many of the same factors that have been identified in this thesis, and in this way, the two studies support each other. However, there are also important differences between the studies: while Chong and Gellersen focus on device binding, the research presented in this thesis covers multi-device interaction in general, including also other tasks than device binding; and while Chong and Gellersen provide a detailed technical analysis of the factors influencing usability, the research presented in this thesis addresses broader user experience, including also other aspects than usability.

In addition to the design principles, the studies addressing the second research question also provide several secondary contributions. First, they contribute novel methods for multi-device interaction, such as the FlexiGroups group binding method. Second, they contribute novel instantiations of known multi-device interaction methods, such as the implementation of the DeviceTouch cross-display object movement method with capacitive sensing technology. Third, they contribute comparative evaluations of several group binding and cross-display object movement methods, assessing the strengths and weaknesses of the different methods against each other and providing recommendations on their use. Few such comparative evaluations have been published in the earlier research and the methods included in the studies presented in this thesis have not been evaluated against each other before. These secondary results can be directly utilized in practical applications.

The third research question RQ3 explores natural methods for initiating multi-device interactions on wearable devices. This research question belongs to design science research and it aims to produce prescriptive knowledge that improves the user experience of multi-device interactions on wearable devices. Regarding this goal, this thesis contributes a categorization of user-generated methods for initiating multi-device interactions between wearable devices. It also provides initial evaluations of the strengths and weaknesses of the different methods. The categorization gives a comprehensive overview of the design opportunities for solving the target identification problem on wearable devices, and it provides a strong basis for further research, design, and development of multi-device interaction methods for such devices. Beyond wearable devices, most of the methods included in the categorization can also be applied to other more conventional device types, such as smartphones, tablets, or computers, and the categorization may be useful in the research and development of multi-device interaction methods for these other device types as well. Considering the earlier related research, Chong and Gellersen (2011; 2013) present similar categorizations of user-generated device binding methods for various device configurations. This thesis extends the work of Chong and Gellersen by examining new device configurations of multiple smartwatches and smartglasses and by exploring the new use case of cross-display object movement.

## 6.2 Reliability and Validity

In this section, we consider the reliability, validity, and generalizability of the research presented in this thesis, and describe measures that were taken to improve them. We also discuss the limitations of the presented research.

**Reliability** refers to the consistency of the research (J. W. Creswell and J. D. Creswell 2018, p. 199). Reliable research can be replicated and it produces consistent and stable results. A prerequisite for reliable research is that the used research procedures are well documented (Yin 2003, p. 38). In the research presented in this thesis, we prepared detailed research plans that described the envisaged research process before each study. During data collection and analysis, notes about the deviations from the research plans were taken and the plans were updated to reflect the actual research procedures. The published papers carefully

report the used research processes and methods so that it is possible for other researchers to replicate the studies.

This thesis uses primarily qualitative research methods. Analysis of qualitative data presents particular reliability challenges because of the ambiguous nature of the data: different researchers may interpret words, expressions, gestures, and body language in different ways, and the overall analysis process may be vulnerable to biases and inconsistencies (Lazar, Feng, and Hochheiser 2010, p. 296). Several measures were taken to improve the reliability of qualitative research in this thesis. In all studies of this thesis, qualitative data was analyzed by two or three researchers, applying the technique of investigator triangulation to reduce personal bias. However, involving multiple researchers may introduce inconsistencies in the analysis as different researchers interpret and code the data differently. Therefore, in studies S1, S4, and S5 where content analysis was used to analyze qualitative data, detailed coding instructions were defined in the beginning of the coding process. Researchers then independently test coded a subset of the data, the coded data was compared, any inconsistencies between the researchers were resolved, and the coding instructions were improved (Lazar, Feng, and Hochheiser 2010, p. 297). Only when agreement on the coding principles was reached between the researchers, the coding of the actual data was started. No statistical reliability measures were used to estimate agreement between researchers, though. In studies where particular accuracy and consistency was desired, such as in the categorization of multi-device use cases in study S1 and in the categorization of user-suggested interaction methods in study S5, the full data set was independently coded by three researchers and all cases where any of the researchers disagreed were commonly discussed and resolved. In studies S1, S2, and S3 where affinity diagramming was used to analyze qualitative data, the entire analysis process was done collaboratively and any inconsistencies and disagreements were immediately resolved during the analysis sessions.

**Validity** refers to the accuracy of the research (J. W. Creswell and J. D. Creswell 2018, p. 199). Research results have high validity if they are well-founded, precisely reflect the real world, and are relevant to the problem being studied. The use of high-quality data collection and analysis methods and procedures is essential for establishing validity (Lazar, Feng, and Hochheiser 2010, p. 295). As described in Section 3.2, the research presented in this thesis rigorously uses well-established research methods. All collected data was systematically organized in a research database. All individual data points, such as interview notes, observations, and use cases, were assigned unique identifiers, so that the analytic results can be traced back to the raw data to establish a chain of evidence (Yin 2003, p. 105).

Different triangulation techniques were used to improve the quality of the data collection and analysis and to counterbalance the deficiencies of any single approach. Data sources triangulation involved collecting data from a total of 110 participants over a series of five different studies. Methodological triangulation involved combining qualitative and quantitative research approaches as well as combining different research methods within the same approach: for example, qualitative data was collected using diaries, interviews, and observation, and quantitative data was collected using several different questionnaires. In addition to increasing reliability, investigator triangulation also improved validity of the research

as different researchers could provide alternative interpretations of the data and complement each other's skills.

**Generalizability** (also known as external validity) refers to the extent the research results can be applied beyond the immediate research setting. The research presented in this thesis primarily uses qualitative research approach, emphasizing detailed analysis of a limited number of cases and subjects. As a relatively small number of subjects is studied, the issue of generalizability is particularly relevant, which is often the case in qualitative research (Yin 2003, p. 37). An essential question regarding the generalizability of the results is how well the studied participants and cases represent the broader population. As described in Subsection 3.2.2, a significant effort was made to recruit a diverse and balanced group of participants for each study. Still, all the studies were made in Finland and the participants were relatively advanced users of technology with a high level of education. While it is likely that the study participants represent well typical user populations also in other western countries, we cannot know it for certain. Further, it is possible that different user groups in other parts of the world might display significantly different behavior. Therefore, additional studies addressing different and larger user groups would be useful to further justify the results in the broader population.

Regarding the second research question RQ2, the presented set of experiential factors to consider in the design of multi-device interaction methods is based on studies of interaction methods for two specific tasks: device binding and cross-display object movement. It can be argued how well the factors derived from studies of these two tasks generalize to other tasks in multi-device interaction. The tasks of device binding and cross-display object movement were selected because they are common and important tasks in multi-device use, and they appear in numerous variations in different applications and domains. They represent core problems in multi-device interaction, and by studying these two tasks, it is possible to reveal fundamental issues about multi-device use. Therefore, we believe that the experiential factors derived from studies of device binding and cross-display object movement apply to many other multi-device interaction tasks as well.

## 6.3   Future Work

While this thesis has presented multiple important new research contributions related to multi-device use, multi-device use is a complex phenomenon and it has been possible to study only a few aspects of it in this thesis. We can identify many possible directions for future research to continue and extend the research presented in this thesis.

As already discussed in the previous section, the results presented in this thesis are primarily based on qualitative research of a limited number of subjects and cases. While this approach has provided us with rich and detailed results, it is difficult to estimate how well the results generalize to the broader population. Therefore, it would be useful to have studies with larger numbers of participants to further validate the results. These studies could be, for example, questionnaire surveys or logging studies.

While we aimed to recruit diverse and balanced groups of participants in the studies presented in this thesis, the participants were still relatively advanced users of technology in a western country. Studies addressing different user groups in other parts of the world might find different behavior. Another possible future direction could be more focused studies targeting specific user groups, applications, or contexts of use, for example, in a professional setting.

The phenomenon of multi-device use is constantly evolving and changing. New information technologies and device types, such as wearable devices, smart cars, and the Internet of Things, are introduced to the market at a rapid pace, further increasing the number and diversity of devices that people have in their personal device ecologies. Over time, people may develop new practices and behaviors of using their extended ecologies of devices, and this may also lead to emergence of new challenges and needs. Identifying and understanding these new practices, challenges, and needs in multi-device use provides an interesting direction for future research.

Wearable devices represent a promising new category of devices for multi-device interaction in the future. The categorization of user-suggested target identification methods that was contributed in this thesis provides a strong basis for further research on natural and intuitive multi-device interaction methods on wearable devices. This new research could aim to build concrete novel multi-device interaction methods for wearable devices and then systematically evaluate them with functional prototypes, addressing also technical aspects that were not included in the elicitation study presented in this thesis. The work on the technical aspects could also address the underlying software platforms and tools to make the design and development of such interaction methods easier and more efficient. Further, the future studies could explore other wearable device configurations beyond smartwatches and smartglasses that were considered in this thesis. These new configurations could include configurations where different users have different wearable devices, configurations where individual users are wearing multiple devices, and configurations which include other wearable device types, such as smart clothes, smart footwear, or smart jewelry.

Considering the user experience evaluation of multi-device interaction methods, the evaluations presented in this thesis, like the vast majority of the other evaluations of multi-device interaction methods that we are aware of, were done in a controlled environment under ideal conditions. The evaluation tasks were defined by the researchers and given to the participants. While a lot has been learned from these evaluations, in real life many situational factors influence the user experience of multi-device interaction and the tasks are much more diverse. Therefore, to gain a deeper understanding of the strengths and weaknesses of multi-device interaction methods in different situations and tasks, it would be important to study them also in more realistic settings and over extended periods of time, for example, with longitudinal field trials.

# Chapter 7

# Conclusion

Instead of a single information device, people today have personal device ecologies that consist of multiple devices. As people have many information devices, situations and workflows where several devices are combined and used together to accomplish a task have become common. Groups of co-located persons may also join their information devices together for collaborative activities and experiences. While these developments towards computing with multiple devices offer many opportunities, they also create a need for interfaces and applications that support using multiple devices together. Compared to traditional single-device interfaces, multi-device interfaces present a challenging design problem as there are multiple devices involved and the interaction is distributed across them.

This thesis has provided contributions in three areas related to multi-device use. First, this thesis has investigated how and why people actually combine and use multiple information devices together in their everyday lives. It has contributed detailed descriptions of people's personal device ecologies, including the most common device types and their characteristic practices of use. It has presented a three-level categorization of patterns of combining multiple information devices together. It has also provided theoretical contributions that explain some of the observed practices, including how users decide which devices to use in a specific situation and how users decide to change the device that they are currently using. The descriptive knowledge presented in this thesis gives a detailed picture of the current and emerging practices and challenges in multi-device use and increases the understanding of some of the key aspects of multi-device use. It provides a solid foundation for further research on more specific topics of multi-device use and also supports the design and development of practical multi-device interfaces and applications.

Second, this thesis has examined the factors that influence the user experience of multi-device use. It has contributed a set of experiential factors related to users, devices, tasks, physical environments, and social situations that have significant effect on the user experience of multi-device interaction and that should therefore be carefully considered when designing multi-device interaction methods. This set of experiential factors provides practical guidance to designers and developers

working with multi-device interfaces and applications. The factors are based on a series of studies evaluating interaction methods for two common tasks in multi-device interaction: device binding and cross-display object movement. This series of studies has also provided a number of secondary contributions, including novel methods for multi-device interaction and comparisons of different interaction methods against each other.

Third, this thesis has explored multi-device interaction methods for wearable devices, focusing on the two most popular categories of wearable devices today: smartwatches and smartglasses. Wearable devices offer new opportunities to support collaborative tasks and shared experiences between groups of co-located persons through multi-user multi-device applications. While there exists a wide variety of multi-device interaction methods for conventional devices, these methods may not be optimal for wearable devices which differ from conventional devices in many ways. This creates a need for novel multi-device interaction methods for wearable devices. This thesis has contributed a categorization of user-generated methods for initiating multi-device interactions between wearable devices based on elicitation studies with groups of participants. It has also provided initial evaluations of the strengths and weaknesses of the different methods. The categorization gives a comprehensive overview of the broad space of potential design opportunities for solving the target identification problem on wearable devices. It provides a strong basis for further research on natural and intuitive multi-device interaction methods for wearable devices as well as for design and development of practical interaction solutions.

Overall, this thesis has contributed new research-based descriptive knowledge about the human phenomena of multi-device use and new research-based prescriptive knowledge about improving the user experience of multi-device use. Altogether, this knowledge advances the scientific understanding of multi-device use in the domain of human-computer interaction research and informs the design and development of future interfaces, applications, and technologies that better support multi-device use.

# Bibliography

Alaszewski, Andy (2006). *Using Diaries for Social Research*. Sage Publications.

Ayatsuka, Yuji and Jun Rekimoto (2005). "tranSticks: Physically Manipulatable Virtual Connections". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '05. Portland, Oregon, USA: ACM, pp. 251–260. DOI: 10.1145/1054972.1055008.

Bachl, Stefan et al. (2011). "The Effects of Personal Displays and Transfer Techniques on Collaboration Strategies in Multi-Touch Based Multi-Display Environments". In: *Proceedings of the IFIP TC International Conference on Human-Computer Interaction*. INTERACT '11. Berlin, Heidelberg, Germany: Springer, pp. 373–390.

Ballagas, Rafael et al. (2003). "iStuff: A Physical User Interface Toolkit for Ubiquitous Computing Environments". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '03. Ft. Lauderdale, Florida, USA: ACM, pp. 537–544. DOI: 10.1145/642611.642705.

Ballendat, Till, Nicolai Marquardt, and Saul Greenberg (2010). "Proxemic Interaction: Designing for a Proximity and Orientation-Aware Environment". In: *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*. ITS '10. Saarbrücken, Germany: ACM, pp. 121–130. DOI: 10.1145/1936652.1936676.

Bardram, Jakob et al. (2012). "ReticularSpaces: Activity-Based Computing Support for Physically Distributed and Collaborative Smart Spaces". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '12. Austin, Texas, USA: ACM, pp. 2845–2854. DOI: 10.1145/2207676.2208689.

Basapur, Santosh et al. (2012). "FANFEEDS: Evaluation of Socially Generated Information Feed on Second Screen as a TV Show Companion". In: *Proceedings of the European Conference on Interactive TV and Video*. EuroITV '12. Berlin, Germany: ACM, pp. 87–96. DOI: 10.1145/2325616.2325636.

Baudisch, Patrick et al. (2003). "Drag-and-Pop and Drag-and-Pick: Techniques for Accessing Remote Screen Content on Touch- and Pen-Operated Systems". In: *Proceedings of the IFIP TC International Conference on Human-Computer Interaction*. INTERACT '03, pp. 57–64.

Benko, Hrvoje, Edward W. Ishak, and Steven Feiner (2005). "Cross-Dimensional Gestural Interaction Techniques for Hybrid Immersive Environments". In: *Proceedings of IEEE Virtual Reality*. IEEE VR '05, pp. 209–216. DOI: 10.1109/VR.2005.1492776.

Bharat, Krishna A. and Luca Cardelli (1995). "Migratory Applications". In: *Proceedings of the Annual ACM Symposium on User Interface Software and Technology*. UIST '95. Pittsburgh, Pennsylvania, USA: ACM, pp. 132–142. DOI: 10.1145/215585.215711.

Biehl, Jacob T. and Brian P. Bailey (2004). "ARIS: An Interface for Application Relocation in an Interactive Space". In: *Proceedings of Graphics Interface*. GI '04. London, Ontario, Canada: Canadian Human-Computer Communications Society, pp. 107–116.

Biehl, Jacob T., William T. Baker, et al. (2008). "Impromptu: A New Interaction Framework for Supporting Collaboration in Multiple Display Environments and Its Field Evaluation for Co-Located Software Development". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '08. Florence, Italy: ACM, pp. 939–948. DOI: 10.1145/1357054.1357200.

Budhiraja, Rahul, Gun A. Lee, and Mark Billinghurst (2013). "Using a HHD with a HMD for Mobile AR Interaction". In: *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. ISMAR '13, pp. 1–6. DOI: 10.1109/ISMAR.2013.6671837.

Cecchinato, Marta E., Anna L. Cox, and Jon Bird (2017). "Always On(Line)?: User Experience of Smartwatches and Their Role Within Multi-Device Ecologies". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '17. Denver, Colorado, USA: ACM, pp. 3557–3568. DOI: 10.1145/3025453.3025538.

Chen, Nicholas, Francois Guimbretiere, and Abigail Sellen (2012). "Designing a Multi-Slate Reading Environment to Support Active Reading Activities". In: *ACM Transactions on Computer-Human Interaction* 19.3, 18:1–18:35. DOI: 10.1145/2362364.2362366.

Chen, Xiang Anthony et al. (2014). "Duet: Exploring Joint Interactions on a Smart Phone and a Smart Watch". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '14. Toronto, Ontario, Canada: ACM, pp. 159–168. DOI: 10.1145/2556288.2556955.

Chi, Pei-Yu (Peggy) and Yang Li (2015). "Weave: Scripting Cross-Device Wearable Interaction". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '15. Seoul, Republic of Korea: ACM, pp. 3923–3932. DOI: 10.1145/2702123.2702451.

Chi, Pei-Yu (Peggy), Yang Li, and Björn Hartmann (2016). "Enhancing Cross-Device Interaction Scripting with Interactive Illustrations". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '16. San Jose, California, USA: ACM, pp. 5482–5493. DOI: 10.1145/2858036.2858382.

Chong, Ming Ki and Hans Gellersen (2011). "How Users Associate Wireless Devices". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '11. Vancouver, British Columbia, Canada: ACM, pp. 1909–1918. DOI: 10.1145/1978942.1979219.

Chong, Ming Ki and Hans Gellersen (2012). "Usability Classification for Spontaneous Device Association". In: *Personal and Ubiquitous Computing* 16.1, pp. 77–89. DOI: 10.1007/s00779-011-0421-1.

Chong, Ming Ki and Hans Gellersen (2013). "How Groups of Users Associate Wireless Devices". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '13. Paris, France: ACM, pp. 1559–1568. DOI: 10.1145/2470654.2466207.

Chong, Ming Ki, Rene Mayrhofer, and Hans Gellersen (2014). "A Survey of User Interaction for Spontaneous Device Association". In: *ACM Computing Surveys* 47.1, 8:1–8:40. DOI: 10.1145/2597768.

Creswell, John W. (2012). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. 4th Edition. Pearson.

Creswell, John W. and J. David Creswell (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. 5th Edition. Sage Publications.

Cruickshank, Leon et al. (2007). "Making Interactive TV Easier to Use: Interface Design for a Second Screen Approach". In: *The Design Journal* 10.3, pp. 41–53. DOI: 10.2752/146069207789271920.

Dearman, David and Jeffery S. Pierce (2008). "It's on My Other Computer!: Computing with Multiple Devices". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '08. Florence, Italy: ACM, pp. 767–776. DOI: 10.1145/1357054.1357177.

Denis, Charles and Laurent Karsenty (2004). "Inter-Usability of Multi-Device Systems: A Conceptual Framework". In: *Multiple User Interfaces: Cross-Platform Applications and Context-Aware Interfaces*. John Wiley & Sons, pp. 373–385. DOI: 10.1002/0470091703.ch17.

Dix, Alan et al. (2004). *Human-Computer Interaction*. 3rd Edition. Pearson Education.

Elmqvist, Niklas (2011). "Distributed User Interfaces: State of the Art". In: *Distributed User Interfaces*. London, United Kingdom: Springer, pp. 1–12. DOI: 10.1007/978-1-4471-2271-5_1.

Finnish National Board on Research Integrity (2018). *Ethical Review in Finland*. http://www.tenk.fi/en/ethical-review-in-finland, last accessed on 2018-Aug-01.

Foley, James D. et al. (1990). *Computer Graphics: Principles and Practice*. 2nd Edition. Addison-Wesley.

Geißler, Jörg (1998). "Shuffle, Throw or Take It! Working Efficiently with an Interactive Wall". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '98. Los Angeles, California, USA: ACM, pp. 265–266. DOI: 10.1145/286498.286745.

Girouard, Audrey, Aneesh Tarun, and Roel Vertegaal (2012). "DisplayStacks: Interaction Techniques for Stacks of Flexible Thin-Film Displays". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '12. Austin, Texas, USA: ACM, pp. 2431–2440. DOI: 10.1145/2207676.2208406.

Google (2012). *The New Multi-Screen World: Understanding Cross-Platform Consumer Behavior*. https://www.thinkwithgoogle.com/research-studies/the-new-multi-screen-world-study.html, last accessed on 2018-Aug-01.

Gregor, Shirley and Alan R. Hevner (2013). "Positioning and Presenting Design Science Research for Maximum Impact". In: *MIS Quarterly* 37.2, pp. 337–356. DOI: 10.25300/MISQ/2013/37.2.01.

Grolaux, Donatien, Peter Van Roy, and Jean Vanderdonckt (2004). "Migratable User Interfaces: Beyond Migratory Interfaces". In: *Proceedings of the International Conference on Mobile and Ubiquitous Systems*. MOBIQUITOUS '04, pp. 422–430. DOI: 10.1109/MOBIQ.2004.1331749.

Grubert, Jens et al. (2015). "MultiFi: Multi Fidelity Interaction with Displays On and Around the Body". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '15. Seoul, Republic of Korea: ACM, pp. 3933–3942. DOI: 10.1145/2702123.2702331.

Grudin, Jonathan (2001). "Partitioning Digital Worlds: Focal and Peripheral Awareness in Multiple Monitor Use". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '01. Seattle, Washington, USA: ACM, pp. 458–465. DOI: 10.1145/365024.365312.

Hamilton, Peter and Daniel J. Wigdor (2014). "Conductor: Enabling and Understanding Cross-Device Interaction". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '14. Toronto, Ontario, Canada: ACM, pp. 2773–2782. DOI: 10.1145/2556288.2557170.

Hang, Alina et al. (2012). "Too Much Information!: User Attitudes Towards Smartphone Sharing". In: *Proceedings of the Nordic Conference on Human-Computer Interaction*. NordiCHI '12. Copenhagen, Denmark: ACM, pp. 284–287. DOI: 10.1145/2399016.2399061.

Hardy, Robert and Enrico Rukzio (2008). "Touch & Interact: Touch-Based Interaction of Mobile Phones with Displays". In: *Proceedings of the International Conference on Human Computer Interaction with Mobile Devices and Services*. MobileHCI '08. Amsterdam, The Netherlands: ACM, pp. 245–254. DOI: 10.1145/1409240.1409267.

Hart, Sandra G. and Lowell E. Staveland (1988). "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research". In: *Human Mental Workload*. Ed. by Peter A. Hancock and Najmedin Meshkati. North-Holland, pp. 139–183. DOI: https://doi.org/10.1016/S0166-4115(08)62386-9.

Hascoët, Mountaz (2003). "Throwing Models for Large Displays". In: *Proceedings of the International Conference on Human-Computer Interaction*. HCI '03, pp. 73–77.

Hassenzahl, Marc (2004). "The Interplay of Beauty, Goodness, and Usability in Interactive Products". In: *Human-Computer Interaction* 19.4, pp. 319–349. DOI: 10.1207/s15327051hci1904\_2.

Hevner, Alan R. and Samir Chatterjee (2010). *Design Research in Information Systems*. Springer.

Hevner, Alan R., Salvatore T. March, et al. (2004). "Design Science in Information Systems Research". In: *MIS Quarterly* 28.1, pp. 75–105.

Hewett, Thomas T. et al. (1992). *ACM SIGCHI Curricula for Human-Computer Interaction*. Technical Report. New York, New York, USA.

Hinckley, Ken (2003). "Synchronous Gestures for Multiple Persons and Computers". In: *Proceedings of the Annual ACM Symposium on User Interface Software and Technology*. UIST '03. Vancouver, Canada: ACM, pp. 149–158. DOI: 10.1145/964696.964713.

Hinckley, Ken, Morgan Dixon, et al. (2009). "Codex: A Dual Screen Tablet Computer". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '09. Boston, Massachusetts, USA: ACM, pp. 1933–1942. DOI: 10.1145/1518701.1518996.

Hinckley, Ken, Gonzalo Ramos, et al. (2004). "Stitching: Pen Gestures That Span Multiple Displays". In: *Proceedings of the International Conference on Advanced Visual Interfaces*. AVI '04. Gallipoli, Italy: ACM, pp. 23–31. DOI: 10.1145/989863.989866.

Holman, David et al. (2005). "Paper Windows: Interaction Techniques for Digital Paper". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '05. Portland, Oregon, USA: ACM, pp. 591–599. DOI: 10.1145/1054972.1055054.

Holmquist, Lars Erik et al. (2001). "Smart-Its Friends: A Technique for Users to Easily Establish Connections Between Smart Artefacts". In: *Proceedings of the International Conference on Ubiquitous Computing*. UBICOMP '01. Atlanta, Georgia, USA: Springer, pp. 116–122.

Holtzblatt, Karen, Jessamyn Burns Wendell, and Shelley Wood (2004). *Rapid Contextual Design*. Morgan Kaufmann.

Houben, Steven and Nicolai Marquardt (2015). "WatchConnect: A Toolkit for Prototyping Smartwatch-Centric Cross-Device Applications". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '15. Seoul, Republic of Korea: ACM, pp. 1247–1256. DOI: 10.1145/2702123.2702215.

Ion, Iulia et al. (2010). "Influence of User Perception, Security Needs, and Social Factors on Device Pairing Method Choices". In: *Proceedings of the Symposium on Usable Privacy and Security*. SOUPS '10. Redmond, Washington, USA: ACM, 6:1–6:13. DOI: 10.1145/1837110.1837118.

Johanson, Brad, Armando Fox, and Terry Winograd (2002). "The Interactive Workspaces Project: Experiences with Ubiquitous Computing Rooms". In: *IEEE Pervasive Computing* 1.2, pp. 67–74. DOI: 10.1109/MPRV.2002.1012339.

Jokela, Tero, Ming Ki Chong, et al. (2015). "Connecting Devices for Collaborative Interactions". In: *interactions* 22.4, pp. 39–43. DOI: 10.1145/2776887.

Jokela, Tero and Andrés Lucero (2014). "MixedNotes: A Digital Tool to Prepare Physical Notes for Affinity Diagramming". In: *Proceedings of the International Academic MindTrek Conference*. AcademicMindTrek '14. Tampere, Finland: ACM, pp. 3–6. DOI: 10.1145/2676467.2676478.

Jokela, Tero, Parisa Pour Rezaei, and Kaisa Väänänen (2016). "Using Elicitation Studies to Generate Collocated Interaction Methods". In: *Proceedings of the International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*. MobileHCI Adjunct '16. Florence, Italy: ACM, pp. 1129–1133. DOI: 10.1145/2957265.2962654.

Jung, Heekyoung et al. (2008). "Toward a Framework for Ecologies of Artifacts: How Are Digital Artifacts Interconnected Within a Personal Life?" In: *Proceedings of the Nordic Conference on Human-Computer Interaction*. NordiCHI '08. Lund, Sweden: ACM, pp. 201–210. DOI: 10.1145/1463160.1463182.

Kainda, Ronald, Ivan Flechais, and A. W. Roscoe (2010). "Two Heads Are Better Than One: Security and Usability of Device Associations in Group Scenarios". In: *Proceedings of the Symposium on Usable Privacy and Security*. SOUPS '10. Redmond, Washington, USA: ACM, 5:1–5:13. DOI: 10.1145/1837110.1837117.

Kane, Shaun K. et al. (2009). "Exploring Cross-Device Web Use on PCs and Mobile Devices". In: *Proceedings of the IFIP TC International Conference on Human-Computer Interaction*. INTERACT '99. Berlin, Heidelberg, Germany: Springer, pp. 722–735.

Karlson, Amy K. et al. (2010). "Mobile Taskflow in Context: A Screenshot Study of Smartphone Usage". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '10. Atlanta, Georgia, USA: ACM, pp. 2009–2018. DOI: 10.1145/1753326.1753631.

Koskinen, Ilpo et al. (2011). *Design Research Through Practice*. Morgan Kaufmann.

Kray, Christian et al. (2008). "Group Coordination and Negotiation Through Spatial Proximity Regions Around Mobile Devices on Augmented Tabletops". In: *Proceedings of the IEEE International Workshop on Horizontal Interactive*

*Human Computer Systems*. TABLETOP '08, pp. 1–8. DOI: 10.1109/TABLETOP.2008.4660176.

Lazar, Jonathan, Jinjuan Heidi Feng, and Harry Hochheiser (2010). *Research Methods in Human-Computer Interaction*. John Wiley & Sons.

Lucero, Andrés (2015). "Using Affinity Diagrams to Evaluate Interactive Prototypes". In: *Proceedings of the IFIP TC International Conference on Human-Computer Interaction*. INTERACT '15. Springer International Publishing, pp. 231–248.

Lucero, Andrés, Jussi Holopainen, and Tero Jokela (2011). "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. Vancouver, British Columbia, Canada: ACM, pp. 1787–1796. DOI: 10.1145/1978942.1979201.

Lucero, Andrés, Jussi Holopainen, and Tero Jokela (2012). "MobiComics: Collaborative Use of Mobile Phones and Large Displays for Public Expression". In: *Proceedings of the International Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI '12. San Francisco, California, USA: ACM, pp. 383–392. DOI: 10.1145/2371574.2371634.

Lucero, Andrés, Tero Jokela, et al. (2012). "EasyGroups: Binding Mobile Devices for Collaborative Interactions". In: *Extended Abstracts on Human Factors in Computing Systems*. CHI EA '12. Austin, Texas, USA: ACM, pp. 2189–2194. DOI: 10.1145/2212776.2223774.

Mainwaring, Scott D., Ken Anderson, and Michele F. Chang (2005). "Living for the Global City: Mobile Kits, Urban Interfaces, and Ubicomp". In: *Proceedings of the International Conference on Ubiquitous Computing*. UBICOMP '05. Tokyo, Japan: Springer-Verlag, pp. 269–286. DOI: 10.1007/11551201_16.

Mäkelä, Ville et al. (2018). "Pocket Transfers: Interaction Techniques for Transferring Content from Situated Displays to Mobile Devices". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '18. Montreal, Quebec, Canada: ACM, 135:1–135:13. DOI: 10.1145/3173574.3173709.

March, Salvatore T. and Gerald F. Smith (1995). "Design and Natural Science Research on Information Technology". In: *Decision Support Systems* 15.4, pp. 251–266. DOI: 10.1016/0167-9236(94)00041-2.

Marcotte, Ethan (2011). *Responsive Web Design*. A Book Apart.

Marquardt, Nicolai, Frederik Brudy, et al. (2018). "SurfaceConstellations: A Modular Hardware Platform for Ad-Hoc Reconfigurable Cross-Device Workspaces". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '18. Montreal, Quebec, Canada: ACM, 354:1–354:14. DOI: 10.1145/3173574.3173928.

Marquardt, Nicolai, Ken Hinckley, and Saul Greenberg (2012). "Cross-Device Interaction via Micro-Mobility and F-Formations". In: *Proceedings of the Annual ACM Symposium on User Interface Software and Technology*. UIST '12. Cambridge, Massachusetts, USA: ACM, pp. 13–22. DOI: 10.1145/2380116.2380121.

Mate, Sujeet, Umesh Chandra, and Igor D. D. Curcio (2006). "Movable-Multimedia: Session Mobility in Ubiquitous Computing Ecosystem". In: *Proceedings of the International Conference on Mobile and Ubiquitous Multimedia*. MUM '06. Stanford, California, USA: ACM. DOI: 10.1145/1186655.1186663.

Mayrhofer, Rene and Martyn Welch (2007). "A Human-Verifiable Authentication Protocol Using Visible Laser Light". In: *Proceedings of the International Con-

*ference on Availability, Reliability and Security.* ARES '07, pp. 1143–1148. DOI: `10.1109/ARES.2007.5`.

McCune, Jonathan M., Adrian Perrig, and Michael K. Reiter (2005). "Seeing-Is-Believing: Using Camera Phones for Human-Verifiable Authentication". In: *Proceedings of the IEEE Symposium on Security and Privacy.* S&P '05, pp. 110–124. DOI: `10.1109/SP.2005.19`.

Merrill, David, Jeevan Kalanithi, and Pattie Maes (2007). "Siftables: Towards Sensor Network User Interfaces". In: *Proceedings of the International Conference on Tangible and Embedded Interaction.* TEI '07. Baton Rouge, Louisiana: ACM, pp. 75–78. DOI: `10.1145/1226969.1226984`.

Meyer, Daniel Z. and Leanne M. Avery (2009). "Excel as a Qualitative Data Analysis Tool". In: *Field Methods* 21.1, pp. 91–112.

Miller, Robert C. and Brad A. Myers (1999). "Synchronizing Clipboards of Multiple Computers". In: *Proceedings of the Annual ACM Symposium on User Interface Software and Technology.* UIST '99. Asheville, North Carolina, USA: ACM, pp. 65–66. DOI: `10.1145/320719.322584`.

Mori, Giulio, Fabio Paternò, and Carmen Santoro (2004). "Design and Development of Multidevice User Interfaces Through Multiple Logical Descriptions". In: *IEEE Transactions on Software Engineering* 30.8, pp. 507–520. DOI: `10.1109/TSE.2004.40`.

Müller, Hendrik, Jennifer Gove, and John Webb (2012). "Understanding Tablet Use: A Multi-Method Exploration". In: *Proceedings of the International Conference on Human-Computer Interaction with Mobile Devices and Services.* MobileHCI '12. San Francisco, California, USA: ACM, pp. 1–10. DOI: `10.1145/2371574.2371576`.

Myers, Brad A., Herb Stiel, and Robert Gargiulo (1998). "Collaboration Using Multiple PDAs Connected to a PC". In: *Proceedings of the ACM Conference on Computer-Supported Cooperative Work.* CSCW '98. Seattle, Washington, USA: ACM, pp. 285–294. DOI: `10.1145/289444.289503`.

Nacenta, Miguel A., Dzmitry Aliakseyeu, et al. (2005). "A Comparison of Techniques for Multi-Display Reaching". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* CHI '05. Portland, Oregon, USA: ACM, pp. 371–380. DOI: `10.1145/1054972.1055024`.

Nacenta, Miguel A., Carl Gutwin, et al. (2009). "There and Back Again: Cross-Display Object Movement in Multi-Display Environments". In: *Human-Computer Interaction* 24.1-2, pp. 170–229. DOI: `10.1080/07370020902819882`.

Nandakumar, Abhishek and Janet Murray (2014). "Companion Apps for Long Arc TV Series: Supporting New Viewers in Complex Storyworlds with Tightly Synchronized Context-Sensitive Annotations". In: *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video.* TVX '14. Newcastle Upon Tyne, United Kingdom: ACM, pp. 3–10. DOI: `10.1145/2602299.2602317`.

Nebeling, Michael (2017). "XDBrowser 2.0: Semi-Automatic Generation of Cross-Device Interfaces". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* CHI '17. Denver, Colorado, USA: ACM, pp. 4574–4584. DOI: `10.1145/3025453.3025547`.

Nebeling, Michael and Anind K. Dey (2016). "XDBrowser: User-Defined Cross-Device Web Page Designs". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* CHI '16. San Jose, California, USA: ACM, pp. 5494–5505. DOI: `10.1145/2858036.2858048`.

Nebeling, Michael, Theano Mintsi, et al. (2014). "Interactive Development of Cross-Device User Interfaces". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '14. Toronto, Ontario, Canada: ACM, pp. 2793–2802. DOI: 10.1145/2556288.2556980.

Nielsen, Jakob (1993). *Usability Engineering*. Morgan Kaufmann.

Nielsen, Michael et al. (2004). "A Procedure for Developing Intuitive and Ergonomic Gesture Interfaces for HCI". In: *Gesture-Based Communication in Human-Computer Interaction*. Ed. by Antonio Camurri and Gualtiero Volpe. Berlin, Heidelberg, Germany: Springer, pp. 409–420.

Norman, Donald (1988). *The Design of Everyday Things*. Basic Books.

Norman, Donald (1999). *The Invisible Computer*. MIT Press.

Olsen Jr., Dan R. and Travis Nielsen (2001). "Laser Pointer Interaction". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '01. Seattle, Washington, USA: ACM, pp. 17–22. DOI: 10.1145/365024.365030.

Olsson, Thomas (2012). "User Expectations and Experiences of Mobile Augmented Reality Services". D.Sc. Thesis. Tampere University of Technology.

Oulasvirta, Antti and Lauri Sumari (2007). "Mobile Kits and Laptop Trays: Managing Multiple Devices in Mobile Information Work". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '07. San Jose, California, USA: ACM, pp. 1127–1136. DOI: 10.1145/1240624.1240795.

Paay, Jeni et al. (2017). "Investigating Cross-Device Interaction Between a Handheld Device and a Large Display". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '17. Denver, Colorado, USA: ACM, pp. 6608–6619. DOI: 10.1145/3025453.3025724.

Park, Seonwook et al. (2018). "AdaM: Adapting Multi-User Interfaces for Collaborative Environments in Real-Time". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '18. Montreal, Quebec, Canada: ACM, 184:1–184:14. DOI: 10.1145/3173574.3173758.

Paternò, Fabio (2000). *Model-Based Design and Evaluation of Interactive Applications*. Springer.

Paternò, Fabio and Carmen Santoro (2012). "A Logical Framework for Multi-Device User Interfaces". In: *Proceedings of the ACM SIGCHI Symposium on Engineering Interactive Computing Systems*. EICS '12. Copenhagen, Denmark: ACM, pp. 45–50. DOI: 10.1145/2305484.2305494.

Paternò, Fabio, Carmen Santoro, and Lucio Davide Spano (2009). "MARIA: A Universal, Declarative, Multiple Abstraction-Level Language for Service-Oriented Applications in Ubiquitous Environments". In: *ACM Transactions on Computer-Human Interaction* 16.4, 19:1–19:30. DOI: 10.1145/1614390.1614394.

Pizza, Stefania et al. (2016). "Smartwatch in Vivo". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '16. San Jose, California, USA: ACM, pp. 5456–5469. DOI: 10.1145/2858036.2858522.

Rashid, Umar and Aaron Quigley (2009). "Interaction Techniques for Binding Smartphones: A Desirability Evaluation". In: *Human Centered Design*. Ed. by Masaaki Kurosu. Berlin, Heidelberg, Germany: Springer, pp. 120–128.

Rekimoto, Jun (1997). "Pick-and-Drop: A Direct Manipulation Technique for Multiple Computer Environments". In: *Proceedings of the Annual ACM Symposium on User Interface Software and Technology*. UIST '97. Banff, Alberta, Canada: ACM, pp. 31–39. DOI: 10.1145/263407.263505.

Rekimoto, Jun (2004). "SyncTap: Synchronous User Operation for Spontaneous Network Connection". In: *Personal and Ubiquitous Computing* 8.2, pp. 126–134. DOI: `10.1007/s00779-004-0262-2`.

Rekimoto, Jun et al. (2003). "Proximal Interactions: A Direct Manipulation Technique for Wireless Networking". In: *Proceedings of the IFIP TC International Conference on Human-Computer Interaction*. INTERACT '03, pp. 511–518.

Robertson, Scott et al. (1996). "Dual Device User Interface Design: PDAs and Interactive Television". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '96. Vancouver, British Columbia, Canada: ACM, pp. 79–86. DOI: `10.1145/238386.238408`.

Rolland, Jannick P., Larry D. Davis, and Yohan Baillot (2001). "A Survey of Tracking Technology for Virtual Environments". In: *Fundamentals of Wearable Computers and Augmented Reality*. Lawrence Erlbaum Associates, pp. 67–112.

Román, Manuel et al. (2002). "A Middleware Infrastructure for Active Spaces". In: *IEEE Pervasive Computing* 1.4, pp. 74–83. DOI: `10.1109/MPRV.2002.1158281`.

Rosnow, Ralph L. and Robert Rosenthal (1997). *People Studying People: Artifacts and Ethics in Behavioral Research*. W. H. Freeman.

Rukzio, Enrico et al. (2006). "An Experimental Comparison of Physical Mobile Interaction Techniques: Touching, Pointing and Scanning". In: *Proceedings of the International Conference on Ubiquitous Computing*. UBICOMP '06. Orange County, California, USA: Springer-Verlag, pp. 87–104. DOI: `10.1007/11853565_6`.

Santosa, Stephanie and Daniel J. Wigdor (2013). "A Field Study of Multi-Device Workflows in Distributed Workspaces". In: *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing*. UBICOMP '13. Zurich, Switzerland: ACM, pp. 63–72. DOI: `10.1145/2493432.2493476`.

Schmidt, Dominik, Fadi Chehimi, et al. (2010). "PhoneTouch: A Technique for Direct Phone Interaction on Surfaces". In: *Proceedings of the Annual ACM Symposium on User Interface Software and Technology*. UIST '10. New York, New York, USA: ACM, pp. 13–16. DOI: `10.1145/1866029.1866034`.

Schmidt, Dominik, Julian Seifert, et al. (2012). "A Cross-Device Interaction Style for Mobiles and Surfaces". In: *Proceedings of the Designing Interactive Systems Conference*. DIS '12. Newcastle Upon Tyne, United Kingdom: ACM, pp. 318–327. DOI: `10.1145/2317956.2318005`.

Scott, Stacey D., Guillaume Besacier, and Phillip J. McClelland (2014). "Cross-Device Transfer in a Collaborative Multi-Surface Environment Without User Identification". In: *Proceedings of the International Conference on Collaboration Technologies and Systems*. CTS '14, pp. 219–226. DOI: `10.1109/CTS.2014.6867568`.

Simon, Herbert A. (1996). *The Sciences of the Artificial*. 3rd Edition. MIT Press.

Stajano, Frank (2000). "The Resurrecting Duckling". In: *Security Protocols*. Ed. by Bruce Christianson et al. Berlin, Heidelberg, Germany: Springer, pp. 183–194.

Stajano, Frank and Ross Anderson (2002). "The Resurrecting Duckling: Security Issues for Ubiquitous Computing". In: *Computer* 35.4, pp. 22–26. DOI: `10.1109/MC.2002.1012427`.

Stefik, M. et al. (1986). "WYSIWIS Revised: Early Experiences with Multi-User Interfaces". In: *Proceedings of the ACM Conference on Computer-Supported Cooperative Work*. CSCW '86. Austin, Texas: ACM, pp. 276–290. DOI: `10.1145/637069.637107`.

Strauss, Anselm and Juliet Corbin (1990). *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Sage Publications.

Streitz, Norbert A. et al. (1999). "i-LAND: An Interactive Landscape for Creativity and Innovation". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '99. Pittsburgh, Pennsylvania, USA: ACM, pp. 120–127. DOI: 10.1145/302979.303010.

Suomalainen, Jani, Jukka Valkonen, and N. Asokan (2009). "Standards for Security Associations in Personal Networks: A Comparative Analysis". In: *International Journal of Security and Networks* 4.1-2, pp. 87–100. DOI: 10.1504/IJSN.2009.023428.

Tandler, Peter et al. (2001). "Connectables: Dynamic Coupling of Displays for the Flexible Creation of Shared Workspaces". In: *Proceedings of the Annual ACM Symposium on User Interface Software and Technology*. UIST '01. Orlando, Florida, USA: ACM, pp. 11–20. DOI: 10.1145/502348.502351.

Terrenghi, Lucia, Aaron Quigley, and Alan Dix (2009). "A Taxonomy for and Analysis of Multi-Person-Display Ecosystems". In: *Personal and Ubiquitous Computing* 13.8, pp. 583–598. DOI: 10.1007/s00779-009-0244-5.

Thevenin, David and Joëlle Coutaz (1999). "Plasticity of User Interfaces: Framework and Research Agenda". In: *Proceedings of the IFIP TC International Conference on Human-Computer Interaction*. INTERACT '99, pp. 110–117.

Ullmer, Brygg, Hiroshi Ishii, and Dylan Glas (1998). "mediaBlocks: Physical Containers, Transports, and Controls for Online Media". In: *Proceedings of the Annual Conference on Computer Graphics and Interactive Techniques*. SIGGRAPH '98. New York, New York, USA: ACM, pp. 379–386. DOI: 10.1145/280814.280940.

Uzun, Ersin, Nitesh Saxena, and Arun Kumar (2011). "Pairing Devices for Social Interactions: A Comparative Usability Evaluation". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '11. Vancouver, British Columbia, Canada: ACM, pp. 2315–2324. DOI: 10.1145/1978942.1979282.

Vaishnavi, Vijay K. and William Jr. Kuechler (2015). *Design Science Research Methods and Patterns: Innovating Information and Communication Technology*. 2nd Edition. CRC Press.

Wäljas, Minna et al. (2010). "Cross-Platform Service User Experience: A Field Study and an Initial Framework". In: *Proceedings of the International Conference on Human Computer Interaction with Mobile Devices and Services*. MobileHCI '10. Lisbon, Portugal: ACM, pp. 219–228. DOI: 10.1145/1851600.1851637.

Weiser, Mark (1991). "The Computer for the 21st Century". In: *Scientific American* 265.3, pp. 66–75.

Weiser, Mark and John Seely Brown (1997). "The Coming Age of Calm Technology". In: *Beyond Calculation: The Next Fifty Years of Computing*. New York, New York, USA: Springer, pp. 75–85. DOI: 10.1007/978-1-4612-0685-9_6.

Wixon, Dennis (2003). "Evaluating Usability Methods: Why the Current Literature Fails the Practitioner". In: *interactions* 10.4, pp. 28–34. DOI: 10.1145/838830.838870.

Wobbrock, Jacob O. et al. (2005). "Maximizing the Guessability of Symbolic Input". In: *Extended Abstracts on Human Factors in Computing Systems*. CHI EA '05. Portland, Oregon, USA: ACM, pp. 1869–1872. DOI: 10.1145/1056808.1057043.

Yang, Jishuo and Daniel J. Wigdor (2014). "Panelrama: Enabling Easy Specification of Cross-Device Web Applications". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '14. Toronto, Ontario, Canada: ACM, pp. 2783–2792. DOI: 10.1145/2556288.2557199.

Yin, Robert K. (2003). *Case Study Research: Design and Methods*. 3rd Edition. Sage Publications.

Zimmerman, Thomas (1995). "Personal Area Networks (PAN): Near-Field Intra-Body Communication". M.Sc. Thesis. Massachusetts Institute of Technology.

# Original Publications

**Publication P1**

Tero Jokela, Jarno Ojala, and Thomas Olsson. 2015. A Diary Study on Combining Multiple Information Devices in Everyday Activities and Tasks. In *Proceedings of the 33rd SIGCHI Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 3903–3912. DOI: https://doi.org/10.1145/2702123.2702211

# A Diary Study on Combining Multiple Information Devices in Everyday Activities and Tasks

**Tero Jokela**
Nokia Technologies
P.O. Box 1000, FI-33721
Tampere, Finland
tero.jokela@nokia.com

**Jarno Ojala**
Tampere Univ. of Technology
P.O. Box 589, FI-33101
Tampere, Finland
jarno.ojala@tut.fi

**Thomas Olsson**
Tampere Univ. of Technology
P.O. Box 589, FI-33101
Tampere, Finland
thomas.olsson@tut.fi

## ABSTRACT

As people possess increasing numbers of information devices, situations where several devices are combined and used together have become more common. We present a user study on people's current practices in combining multiple information devices in their everyday lives, ranging from pragmatic tasks to leisure activities. Based on diaries and interviews of 14 participants, we characterize the usage practices of the most common devices, including smartphones, computers, tablets, and home media centers. We analyze 123 real-life multi-device use cases and identify the main usage patterns, including Sequential Use, Resource Lending, Related Parallel Use, and Unrelated Parallel Use. We discuss the practical challenges of using several information devices together. Finally, we identify three levels of decisions that determine which devices are used in a particular situation, including acquiring, making available, and selecting the devices for use.

## Author Keywords

Information devices; smartphones; tablets; multi-device; device ecologies; mobile use; user study.

## ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

## INTRODUCTION

Visions of ubiquitous computing have long predicted an evolution from single-device computing towards computing with multiple devices. Weiser [19] envisioned a future of multiple computers of different scales per user, all interconnected by a ubiquitous network. Norman [14] proposed families of information appliances, specialized to perform specific activities and capable of sharing information among themselves, as a solution to the

complexity of personal computers. In line with these predictions, people today own and use increasing numbers of interconnected devices with diverse form factors. First smartphones and then tablets have established themselves as new device categories alongside personal computers. Some conventional devices such as televisions and cameras have become connected to the Internet and gained capabilities to provide access to many of the same applications and services as computers. Emerging industry trends, including wearable devices, connected cars, and the Internet of Things, suggest continuing increase in the number and diversity of devices in the future.

While this development towards computing with multiple devices offers many new opportunities, it has also created a need for interfaces, applications, and services that better support multi-device use. Today, many popular applications and services can be accessed with a range of devices with different screen sizes and form factors. A session of use can be saved and closed on one device and re-opened and continued on another device. Cloud services support centralized management of device families, including device settings and installed applications, and file hosting services allow accessing and synchronizing content between devices. Recent web browsers provide a common usage history across all devices and allow moving browser tabs between devices. A multiplicity of technologies has been developed for presenting visual or audio information through other devices, for example, using one device to show pictures on the screen or to play music through the speakers of another device. Attempts to harmonize software and hardware platforms and tools across device categories aim to make it easier to develop applications that support multiple devices.

While a wide variety of different technologies have been proposed to support multi-device use, only a few studies have addressed how people actually use multiple information devices together in their everyday lives. Of these studies, many are relatively old, pre-dating the emergence of modern smartphones, tablets, and cloud services. In this paper, we present a recent diary study on how people today combine multiple information devices in everyday activities and tasks. Based on diaries and interviews of 14 participants, we characterize the evolving usage practices of the most common devices, including

smartphones, computers, tablets, and home media centers. We analyze 123 user-reported cases of multi-device use and identify the main usage patterns, including Sequential Use, Resource Lending, Related Parallel Use, and Unrelated Parallel Use. We discuss the practical challenges of owning and operating several information devices together. While most of the earlier studies have been based solely on user interviews, our study is additionally grounded on user diaries reporting real-life use cases recorded in real contexts of use over a time period of one week. Also, while most of the earlier studies have focused on information work related use, our study explores a range of different use cases from pragmatic tasks to leisure activities in various mundane contexts of use. The objective of our study is to provide qualitative insights into real-life practices of multi-device use. The results inform the design of future interfaces, technologies, and applications that better support multi-device use.

## RELATED WORK

A wide variety of systems, user interfaces, and interaction methods have been proposed to support computing with multiple devices. These include device binding methods that allow connecting several devices to operate together, ranging from virtual methods (such as scanning for available devices in the proximity) to physical methods (such as synchronous user actions, spatial alignment, and use of auxiliary devices) [18]. Similar techniques have also been developed to support transferring content objects or application windows between screens, including transfers between devices within hand's reach and between devices at longer distances from each other [13]. Migratory interfaces [1] provide techniques for moving application windows between devices. Solutions have also been developed for managing and switching tasks in multi-device computing environments [2]. Beyond single-user systems, a broad range of multi-user systems that support collaboration with multiple devices have been developed, addressing different physical scales and types of social interaction [17].

In addition to the development of methods and solutions to support multi-device use, a number of studies have addressed how people actually use multiple devices in real life. Several reasons and motivations for using multiple devices have been identified [15, 3]. Different devices suit different tasks, social situations, and physical environments. Estimated effort to set up the device is an important consideration in selecting which device to use, especially for short tasks. No single device may have all the functions or data needed, forcing the use of multiple devices. Additional devices can also be used as data or battery backups. Sometimes personal preferences and habits may influence the decision to use different devices for different tasks.

Practices and workflows in using multiple devices have been found to vary between different individuals and professional groups [3, 9, 16]. People tend to divide tasks between devices, assigning each device a specific role within the workflow [5, 3, 16]. In addition to serial patterns of multi-device use [9], also parallel patterns have recently become more common [16, 4, 12].

Accessing and managing content across devices has been observed to be one of the key concerns in multi-device environments [15, 3, 16, 12]. People have been found to assemble their own personal patchworks of solutions by combining multiple different tools and approaches, often in creative ways. Conventional solutions include manual synchronization and mirroring between devices, dedicating a certain device for certain kind of information, portable storage devices such as USB memory sticks, e-mailing content items to oneself, and network drives. More recently, various cloud-based storage and synchronization solutions have increased their popularity. In addition to content, other kinds of data such as interaction histories should also be synchronized between devices [15, 3, 8].

In mobile use, managing different device configurations may require significant physical effort and planning [15]. People have been observed to address this problem by adopting different strategies, including development of stable habits, making just-in-case preparations for potential situations, and doing careful advance planning. A Mobile Kit refers to a stable set of multiple personal devices kept together and carried while traveling [11]. Jung, et al. [7] suggest examining the set of digital devices that a person uses as an ecology of interactive artifacts, in order to understand how people experience and strategize the use of interactive artifacts and the development of their ecologies over time.

While a number of studies have addressed how people use multiple devices in real life, they are still relatively few given the broadness of the topic. Further, several of the earlier studies have been made before the availability and widespread adoption of modern smartphones, tablets, and cloud services. Therefore, there is room for an updated view into how the new extended ecologies of devices and services are managed by users. Most of the earlier studies have also focused on technologically advanced users and use cases related to information work. We believe our study addressing diverse groups of users and exploring beyond work-related use is useful in broadening the understanding of the evolving practices and needs in multi-device use.

## OUR STUDY

### Objectives

In this study, we were interested in how people today combine multiple information devices in their everyday activities and tasks. For the purposes of this study, we defined an information device as any device that can be used to create or consume digital information, including personal computers, smartphones, tablets, televisions, game devices, cameras, music players, navigation devices, and

wristwatch computers. In particular, we were interested in the following three research questions related to multi-device use: (1) What devices do the participants have and what are the roles of the different devices in their personal device ecologies? (2) In which situations do they combine multiple information devices and what kinds of practices do they have for multi-device use? (3) What kind of challenges and problems they face in multi-device use? Overall, based on a qualitative analysis, we aimed at creating a rich picture of the current and emerging practices and challenges in using multiple information devices together, in order to support the development of future multi-device interfaces, technologies, and applications.

### Participants

We recruited a total of 14 people living in Southern Finland to participate in the study by posting an advertisement on local mailing lists and social media groups. We selected a diverse group of people who owned multiple devices and actively used them for a wide range of different applications and services. Six of the participants were female and eight male. Seven of the participants were students (age M=24.4, SD=2.8 years), while the other seven were professionals working in different fields (age M=35.6, SD=6.7 years). Regarding education and professional background, four participants had their primary background in information technology. The remaining ten participants represented a wide variety of other professions, including graphic and industrial designers, a medical doctor, a goldsmith, a teacher, and a management assistant. Despite not being pure technologists, the participants were fairly advanced users of information technology: on a scale between 1 and 7 (1=novice, 7=expert), the participants rated their familiarity with information technology above average (M=5.6, SD=0.8). The participants received small rewards for their participation.

### Method

The study consisted of three phases: an initial briefing session, a one-week self-reporting period, and a final interview. All the briefing sessions and interviews were individual with the participant and one or two researchers present. Five briefing sessions were held over the phone – other briefing sessions and all interviews were made face-to-face, either in a meeting room in our laboratory or in the participant's place of study or work. The purpose of the briefing session was to provide the participant with necessary information and materials for the self-reporting period. After introducing the participant to the scope and the objectives of the study, the researcher gave the participant a workbook for the self-reporting period and explained the contents of the book.

The main part of the workbook was a diary where the participant was asked to report all situations and tasks where they used multiple information devices together. We asked the participant to keep a complete diary for at least

three days (two working days and one non-working day) during the study period, but we also encouraged them to report any interesting situations that occurred on the other days. For each situation, the participant was asked to report the context of use, overall course of events, devices that were used, practices for and motivations of using multiple devices, frequency and importance of the situation, as well as satisfaction with the course of events and possible problems. The researcher emphasized the importance of filling in the book as soon as possible after the event occurred. In addition to the diary, the workbook included other tasks, where the participant was asked to provide information about their devices, the tasks they were used for, and the contexts they were used in. The participant was also asked to describe the typical set of devices that they carried with them in mobile situations, as well as their ideal multi-device setup. The purpose of these tasks was to help the participant to prepare for the interview by collecting information about their device ecology and practices, and reflecting them in real contexts of use (for example, at home or at the workplace).

The participant then started the one-week self-reporting period during which they filled in the workbook. At the end of the period, the researchers tentatively analyzed the workbooks to prepare for the semi-structured interviews. We asked the participant to bring with them the devices they carried on a typical work day. The interview started by asking about the participant's devices and their use, followed by a detailed discussion of two or three different situations that the participant had reported. Specific questions were asked about various topics, including sharing of content between devices, maintaining multiple devices, and obtaining new devices. At the end of the approximately one-hour interview, the participant was asked to summarize the main benefits and drawbacks of using multiple devices. The interviews were audio recorded and photographs of the participant's devices and their use were taken when considered relevant.

For each interview, the researcher who made the interview wrote notes about it based on the audio recordings. Three researchers then analyzed the data and built an Affinity Diagram [6] in a series of interpretation sessions. Based on interpretative content analysis, the notes were grouped based on similarity. The groups were then further clustered to broader categories that were identified from the data and jointly revisited, discussed, and refined. The diary data was separately analyzed and the reported use cases were categorized into different patterns of multi-device use based on a bottom-up data-driven categorization. This was done to develop an understanding of different patterns of multi-device use based on the collected data rather than to prove any a priori hypothesis. The categorization scheme was primarily based on the roles and the configuration of the devices as well as the user's motivations for multi-device use. Finally, based on the data from the other tasks in the participants' workbooks, we constructed matrices

describing relationships between devices, tasks, and contexts of use.

**RESULTS**

We begin by giving an overview of the participants' device ecologies, describing the most common devices and their practices of use. We then present an analysis of the collected multi-device use cases and identify the main usage patterns, including Sequential Use, Resource Lending, Related Parallel Use, and Unrelated Parallel Use. Finally, we highlight various practical challenges related to owning and using several devices together and discuss the problem of selecting which devices to use.

**Device Ecologies**

On the average, each participant was actively using 7.9 information devices in their everyday life. Four participants had mainly Apple devices, while the remaining 10 had a more balanced mixture of devices from different manufacturers. In addition to their personal devices, seven participants also had devices provided by their employers. Not surprisingly, the most commonly used devices by the participants were smartphones, computers, tablets, and home media centers (Fig. 1).

All participants (14/14) had a smartphone; two participants had two smartphones, a work phone and a personal phone. The smartphone was the only device that every participant said they used every day. Several participants commented that smartphone was their most important device and their everyday life depended on it. *"[P7] I am hooked on my [smartphone]. There is no life without it. I use it for everything. ... It is my most important device that I use the most."* As the participants carried their smartphones almost always with them, the phones allowed them to be reached and kept them up to date anywhere they went. The smartphones were used for a wide variety of different tasks, including calendar, e-mail, messaging, phone calls, web browsing, social media, music, and calculator. Smartphones were also commonly used for taking pictures and for several participants, the phone was their primary camera. However, the use of smartphones was focused on retrieving and entering small amounts of information and the smartphones were not considered capable of handling large amounts of information or complex tasks. While many applications and services had smartphone versions, they often lacked features or did not work properly.

All participants (14/14) also had laptop computers and more than half of them (8/14) had several laptops in their use. Nearly half of the participants (6/14) also had a desktop computer. As computers were replaced by other devices in consumption and entertainment oriented tasks, their use was increasingly characterized by more complex tasks, which included entry of long texts, creation of complicated content, detailed work, and handling large amounts of content. These tasks were often related to work or studies, and they were described as important but serious and
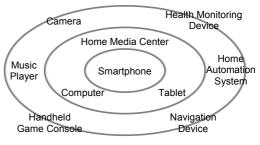


**Figure 1. Most commonly used devices.**

somewhat boring. *"[P13] My laptop is my trusted companion, which I use for all important things: school and work."* The participants considered their computers as their most powerful devices that provided options beyond what their other devices could offer. Computers were also regarded as fallback devices that the participants turned to when their other devices failed. Sometimes computers formed a link between devices that could not otherwise communicate with each other. Among modern smartphones and tablets, the participants felt that their computers were like legacy devices from the old world but they simply could not manage without them. *"[P6] After I got the tablet, my laptop feels really ancient, somehow it is awfully old-fashioned. But certain things you just cannot do without it."*

Almost all participants (12/14) actively used tablets [12] and two participants had several tablets. The two participants who did not use tablets had tested them but considered that they did not have use for a tablet as other devices (especially smartphones) served the same purposes. The tablets were primarily used for searching for and consuming content and for entertainment, such as reading, playing games, and watching videos. Some participants considered tablets as smartphones with larger screens. Compared to laptops, the tablets were considered to be more lightweight and instantly available for use. Two common patterns of use were observed: 1) a shared family tablet at home for multiple purposes; 2) a personal tablet that was partly replacing a laptop in mobile use. However, tablets were not considered suitable for entering long texts or doing complex or detailed work. One participant described this limitation also as an advantage since it forced him to focus on the essential. In general, tablets were still looking for their role and their use involved more experimentation than the use of the other devices.

For the purposes of this study, we defined Home Media Center to consist of a TV or other large screen together with game consoles, Home Theater PCs (HTPCs), and other media appliances connected to it. All participants had some kind of a Home Media Center system. The Home Media Center was the heart of social entertainment at home. It provided the largest screen in the household supporting

collective viewing of and sharing of TV programs, movies, games, music, and personal photos and videos. Video consumption was dominated by on-demand Internet services and local content providers. Lack of ready-made solutions supporting all content sources led the participants to build custom solutions, which were often PC based.

In addition to these four most commonly used devices, the participants reported using a variety of other devices. Of the other devices, the most common were digital cameras, which were used in pre-planned situations where high quality photographs and videos were desired. Other devices included music players, handheld game consoles, navigation devices, health monitoring devices, and home automation systems.

The smartphones were considered the most personal devices and were rarely shared, with only a few exceptions such as a child playing games on a parent's phone or a wife taking a photo with her husband's phone. Sharing tablets and computers was more common, and within families, all devices were in principle shared. Still, most devices had a clear primary user and others used the device only randomly. The primary user had their services and applications configured and often always running on the device. *"[P5] All of our devices are shared. But I use primarily some devices and my husband uses other devices, because on those you always have the web pages, files, and apps you need already opened."* Some participants also expressed a desire to keep some of their devices private. *"[P6] I have tried to keep the tablet to myself, but of course [my daughter] can use it as well if she wants. But some devices you just want to keep to yourself."*

**Patterns of Multi-Device Use**

The participants reported a total of 111 situations and tasks in which they had used multiple information devices together during the one-week diary period. Initial analysis of the data indicated that 23 of the situations involved several separate instances of multi-device use and therefore these situations were split into multiple cases. On the other hand, 12 situations were rejected, either because the description of the situation was too unclear or because the situation did not involve use of multiple information devices (for example, there was only one information device used). Eventually, the preliminary analysis resulted in a total of 123 cases of multi-device use.

Fig. 2 illustrates the categorization of these multi-device use cases into the main patterns of use. On the highest level, multi-device use can be divided into Sequential Use and Parallel Use. In Sequential Use, the participant changed the device during the task. In Parallel Use, the participant used several devices simultaneously. Of the multi-device use cases analyzed in the study, 37% were Sequential Use while the remaining 63% were Parallel Use.

The cases of Parallel Use can be further divided into three subtypes: Resource Lending, Related Parallel Use, and
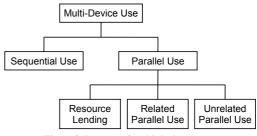


**Figure 2. Patterns of multi-device use.**

Unrelated Parallel Use. In Resource Lending, the participant's task focused on a single device, but the device borrowed some resources from other devices. In Related and Unrelated Parallel Use, the participant used several devices simultaneously. The difference is that in Related Parallel Use, the participant was working on a single task and all devices were involved in this task, while in Unrelated Parallel Use, the participant was working on multiple tasks simultaneously and different devices were involved in different tasks. Of the cases of Parallel Use analyzed in the study, 43% were classified as Resource Lending, 44% as Related Parallel Use, and 13% as Unrelated Parallel Use.

In general, the participants' motivations for using multiple devices could be divided in two broad categories. First, in 83% of the cases, the participant voluntarily decided to use multiple devices, for example, in the hope of improved performance, efficiency, or convenience of use. Second, in the remaining 17% of the cases, the participant was forced to use several devices, for example, because of technical limitations or errors, or because the original device used to start the task had become unavailable.
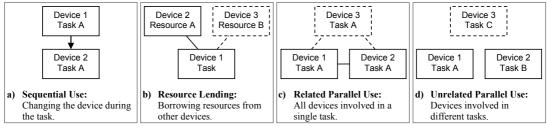
In the following subsections, we discuss each pattern of multi-device use as well as related motivations and behaviors in more detail.

**Sequential Use**

In Sequential Use (Fig. 3.a), the participant changed the device during the task. The participant could change the device once or several times in a sequence.

In 68% of the Sequential Use cases analyzed in the study, the user voluntarily changed the device. We observed several reasons that triggered the participant to switch devices. A common reason was a transformation in the character of the task, which made the participant to consider another device better suited for the task and to change the device. *"[P7] I googled the phone number of my physiotherapist with the tablet and called with my phone."* Another common reason was a change in the physical environment or the social context, which resulted in another device deemed more appropriate for the task. *"[P1] In the bus, I was browsing with my phone and found an interesting page about teacher's copyrights. At work, I*

Figure 3. Different patterns of multi-device use.

*continued reading the page with my laptop."* Often, both the task and the context changed at the same time. In these cases, there was typically a workflow consisting of a sequence of clearly separate steps, each done at a different location. *"[P13] I searched for a recipe with my laptop. Then I moved to the kitchen, opened the recipe with the tablet, and started cooking."* There were also a few cases in which the task and the physical and social context remained unchanged, but the participant still decided to change the device. In these cases, the participant's initial evaluation of the task proved incorrect as more information was learned about the task, ultimately resulting in a device change. *"[P13] When meeting the other members of the project group, I tried to browse the course materials with my phone, but finally switched to my laptop because the phone was too slow and the display too small."*

The dilemma of whether to change the device or not can be illustrated with Fig. 4.a. The horizontal axis represents the progression of a task, while the vertical axis represents the user effort. Each device requires an initial setup effort, represented by the y-intercept of the line, before it can used to work on the task. Each device also has certain efficiency, represented by the slope of the line. As illustrated in Fig 4.b, the efficiency of the device may change as the characteristics of the task change. In order to switch the device (Fig. 4.c), the user has to do additional effort. While the person may gain improved performance after the switch, it may take a long time to outdo the switching effort. The problem may be further complicated by the difficulty of predicting the future evolution of the task.

In 32% of the Sequential Use cases, the user was forced to change the device. The most common reason was technical problems that prevented continuing with the original device [9]. *"[P6] I opened the electronic messaging system of my daughter's school with my tablet and tried to reply to a message, but a scrolling text field did not work with the tablet. I mailed the text I had entered to myself and continued with my laptop."* Overall, technical problems were the reason for using multiple devices in 25% of the Sequential Use cases. Other reasons that forced the user to change the device included the original device running out of battery or the need to access content that was stored on another device.

We observed a variety of different methods of moving a task from one device to another. It was common that the devices and applications did not support any way of moving the task from one device to another, or the user was unaware of or unwilling to use them. In these cases, the user just started the task from the beginning with the new device and manually copied information from the old device to the new one as necessary. However, in many cases it was possible to save the task on a network resource, for example, on a cloud service or a network disk, and to reopen it with another device. Other less common approaches were to send a message, typically an e-mail, to oneself, or to use a physical medium such as a USB stick or memory card to transfer the task.

**Parallel Use**

In Parallel Use, the participant used two or more devices simultaneously. The cases of Parallel Use can be further divided into three subtypes: Resource Lending, Related Parallel Use, and Unrelated Parallel Use.

*Resource Lending*

In Resource Lending (Fig. 3.b) use cases, the participant's activity primarily focused on a single device, but this device borrowed some resources from other devices.

Common examples of Resource Lending included borrowing the input and output capabilities of another device. For example, the screen of another device could be used to display visual information or speakers to play audio. *"[P6] I connected my laptop to my TV to watch an episode of a TV series."* Alternatively, the input devices of one device could be borrowed to provide input for another device. *"[P8] I controlled my home theatre system with an app installed on my phone."* In some cases, both input and output capabilities of another device were borrowed and all interaction took place through another device. *"[P2] I used my laptop to organize messages and music on my phone."* In addition to input and output capabilities, another common resource shared between devices was the network connection. *"[P2] In the bus, I used the hotspot feature on my phone to connect my laptop to the Internet."* Resource Lending was almost exclusively done using direct wireless connections between devices. Other approaches included traditional cables and docking stations as well as lending resources through a cloud service.
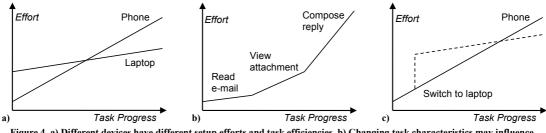
**Figure 4. a) Different devices have different setup efforts and task efficiencies. b) Changing task characteristics may influence device efficiency. c) Switching devices requires additional effort that may exceed the benefits.**

*Related Parallel Use*

In Related Parallel Use (Fig. 3.c), the participant was working on a single task using two or more devices. All devices were involved in the same task.

In most cases, the motivation for using multiple devices in parallel was to have multiple views to the task, dedicating different devices to different content or applications. A common situation was that a participant was using another device to view additional information when watching video content with a home media center. *"[P5] While watching a movie on TV, I opened the IMDB page with my phone."* In other cases, a participant was viewing instructions on one device while working on a task on another device. *"[P1] I installed software on my phone and looked at instructions on my computer."* The participant could also work on a new document on one device, while having the source documents open on the other devices. *"[P10] While making a project presentation with my laptop, I had my notes open on my tablet."* Sometimes, there simply were several documents that needed to be viewed in parallel. *"[P10] I displayed CAD drawings on my laptop and images on my tablet."* A second device could also be used for simple auxiliary tasks, for example, to access the dictionary. *"[P8] While watching TV, I checked the translation of a word with my phone."*

Another common category of cases of Related Parallel Use was situations where the participant was collaborating with remote persons over a real-time communication link. One device, typically a phone, was then assigned for handling the communication, while other devices were used for accessing related content and applications. *"[P8] My friend called me and asked me to the movies. I checked my calendar and information about the movie with my laptop while talking on the phone."* Other less common motivations for Related Parallel Use included speeding up a computationally complex task like 3D graphics rendering by using multiple devices as well as various technical reasons which forced the use of several devices.

In most cases, the participants used two devices together. However, in a few cases the participants had built more sophisticated device configurations which involved Parallel Use of three or more devices. *"[P4] While watching sports on TV, I have live statistics open on a laptop. At the same*

*time, I can do real-time betting on my tablet or phone."* In practically all cases, the devices provided no technical support of any kind for Related Parallel Use. The participants had to manually connect and transfer information between the devices.

*Unrelated Parallel Use*

In Unrelated Parallel Use (Fig. 3.d), the participant was working on several tasks simultaneously using several devices. Different devices were involved in different tasks.

In cases of Unrelated Parallel Use, there was typically a primary foreground task and a secondary background task. Typical examples of background tasks included watching videos or listening to music. *"[P10] I listen to music with my phone while doing homework with my laptop or tablet."* In a few cases, the participant was working on two equal tasks in parallel, for example, when the participant received two simultaneous phone calls. *"[P1] I was talking with my mother on my personal phone, when my girlfriend called to my work phone."* There were also a few cases where technical restrictions required the participant to use different devices for different tasks. While different devices were involved in different tasks, in some situations coordination between the devices would have been beneficial. *"[P6] I was listening to a net radio with my laptop, when I received a phone call. I would have wanted to turn down the laptop volume when the phone rang."*

The participants reported relatively few cases of Unrelated Parallel Use. It is possible that the participants did not recognize many common situations as Parallel Use of multiple devices, as the unrelated parallel tasks may not always be conscious. For example, carrying a phone in order to be reachable by voice calls or messages when working on other tasks with other devices (undoubtedly a very common case) could be classified as Unrelated Parallel Use.

**Deciding Which Devices to Use**

Owning multiple devices creates the problem of choosing which devices to use in a specific situation. In our analysis of multi-device use cases, we identified three levels of decisions which determined which devices to use: 1) deciding which devices to acquire, 2) deciding which of your devices to make available in a specific context, and 3)

deciding which of the available devices to actually use. Each level of decisions narrowed the available options until finally the devices to use were left. The participants' device selections determined the pattern of multi-device use: the decision to change device resulted in a Sequential Use pattern, while the decision to use two or more devices simultaneously resulted in a Parallel Use pattern.

Regarding acquisition of new devices, most participants said that interoperability with their current devices was a major factor when considering which new devices to obtain. On the other hand, other factors such as price or curiosity to try and opportunity to learn new devices and systems often overrode it. The level of influence the participants had over the acquisition decisions varied: the participants had more control over devices they purchased for their personal use than over devices that were provided to them by their employers. Overall, the participants could not fully control their device environments as they were constrained by the choices of other people and organizations, such as their employers, educational institutions, clients, and friends.

For fixed devices, the decisions on which devices to make available in specific contexts primarily related to the physical locations of the devices, for example, in which room to place the device at home. For mobile devices, the decisions related to which devices to take with you in mobile situations. Regarding the mobile devices, we observed similar practices as Oulasvirta, et al. [15]. Most participants had a fixed Mobile Kit [11] of devices they took with them every day. Heavier devices such as laptops and special devices such as USB drives were taken based on expected need. In addition to computing devices, the participants carried with them a wide variety of auxiliary devices such as chargers, headsets, cables, and memory sticks.

Decisions on which of the available devices to use in specific situations were based on several criteria. The device characteristics that were considered included both user interface capabilities (such as display size, pointing device, and the level of multi-tasking support), technical capabilities (such as processing power, storage capacity, network connection, and camera quality), and physical characteristics (such as the size and weight of the device). For tasks requiring entry of long texts, the text entry capability, especially availability of a physical QWERTY keyboard, strongly influenced the device selection. *"[P4] I won't write any long texts with a virtual keyboard. I want the good old physical QWERTY."* As already discussed in section Sequential Use, another major factor influencing the decision on which device to use was the easiness of starting to use the device. If the required content was available only on certain devices, the participant was forced to use those devices. People had also developed different habits of using certain devices for certain tasks or in certain contexts of use. Finally, the participants also considered the context of

use, including both social aspects (such as acceptability of using a certain device in a certain social situation) as well as physical aspects (such as the available space). *"[P13] As I met new persons, I did not want to hide behind the [laptop] screen, as we were supposed to work together and be social."*

**Accessing Content Between Devices**
Being able to access any content on any device was the most commonly requested feature to support multi-device use by the participants. For transferring individual content items between devices, sending e-mail to oneself was still common, but also direct network transfers between devices over WiFi or Bluetooth were used. Traditional physical methods of memory sticks, cables, and DVDs had largely been replaced by wireless or cloud-based solutions, but were still used as fallbacks when more advanced solutions failed to work, especially between systems of different owners.

The participants utilized cloud storage services (for example, Dropbox, Google Drive, and iCloud) for a wide variety of purposes, including storing and accessing personal content from any device, and synchronizing and making backups between devices. Cloud storage services were also used for other purposes, such as sharing and collaborating between people and for temporary storage. Using several services in parallel was common as this allowed the participants to grab free storage space from every service and reduced the risk of being tied to a single service provider. It was also common to dedicate different services to different purposes or content types, for example, using one service for personal content, another for work, and a third for collaboration with others.

While the participants utilized cloud storage services in many ways, they also raised many concerns about cloud storage and did not trust the cloud as the only storage solution [16], especially for important content. Accessing content in the cloud was considered slow and sometimes unreliable, particularly over wireless networks. *"[P5] Accessing [cellular] networks sometimes causes problems in certain areas. It is really infuriating."* To address this problem, one participant used different wireless network operators on different devices to be able to always select the best working network. A few participants were concerned about privacy and security risks, for example, of storing personal photographs or work-related content in the cloud. Another source of concern was the persistence and possible discontinuation of cloud services. Finally, many current cloud services were considered complicated and poorly integrated with native applications.

**Device Maintenance and Energy Management**
Participants with large collections of information devices had to do significant amounts of maintenance work to keep their devices up to date and running. However, half of the participants (7/14) said that this was not a major problem as

the software updates were so easy to do. *"[P8] All [devices] have automatic updates and I just have to accept them."* Other participants (6/14) considered the need to maintain a large amount of devices as troublesome and would have preferred more automated solutions. Maintenance responsibility was often concentrated on a particular person in a social network, for example, the husband in a family. These people typically enjoyed the maintenance task and considered it as a positive challenge. *"[P4] Device maintenance in our family is my responsibility. ... I like to do it. I take it as a challenge. Updating the devices of my parents and my mother-in-law are my responsibility, too."*

Another challenge related to the use of multiple devices was managing the energy and charging the batteries of portable devices. The battery concerns were mainly related to smartphones and most participants (9/14) said they were frustrated with the battery lives of the current smartphones. Other devices were considered to have reasonably good battery lives – tablets, in particular, were praised for their long battery lives.

## DISCUSSION

In general, the participants wanted all their devices to seamlessly work with each other. However, in practice, they continuously encountered problems in multi-device use, especially between devices from different ecosystems. Common problems included connecting and transferring information between devices, incompatible content formats, web pages that did not work on all devices, and applications and services that were not available for all devices. The participants often had to resort to common core functions, such as e-mail, to work around interoperability problems. Plenty of work still remains to be done to realize the visions of smooth and effortless multi-device computing.

Today, most of the commercial support for multi-device use is aimed towards supporting patterns of Sequential Use, for example, moving sessions between devices. Also, traditional forms of Resource Lending, such as displaying pictures on the screen of another device, are well supported. However, the results of our study indicate that also patterns of Related Parallel Use were common among the participants. This finding is supported by other recent studies [16, 4, 12]. Still, the devices and systems used by the participants provided practically no support for Related Parallel Use.

Being able to access any content with any device was the most commonly requested feature to support multi-device use by the study participants. While the participants utilized cloud storage services in many ways, it was obvious that the current services were not adequate as the only storage solutions. A wide range of user concerns should be addressed in future storage solutions, including capacity, cost, privacy, security, performance, reliability, persistence, and complexity. As suggested by one of our participants,

one potential direction to explore might be more direct sharing of content between devices, in order to provide improved reliability and performance and lower cost: *"[P10] Information between devices currently flows through the cloud, but when the devices are close to each other, why it couldn't pass directly between the devices?"*

The participants described that in the modern world of smartphones and tablets, the PC felt like a legacy device from the old world. Some participants had tried to completely replace their computers with smartphones and tablets. However, it was clear that while smartphones and tablets had taken over many simple tasks traditionally done with computers, in their current form they could not fully replace computers in more demanding tasks. The main problems were related to limited text-entry capability, inaccurate pointing devices, and restricted multi-tasking capability. This calls for "reinventing the PC" in the age of multi-device computing, that is, developing multi-device solutions that are capable of handling the large, complex, and detailed tasks that currently require the use of a personal computer. Interestingly, the limited multi-tasking capabilities of current tablets and smartphones already seem to encourage the use of multiple devices to overcome them.

As the participants purchased new devices, they sometimes recycled their old devices for others to use. However, often the old devices were gradually left unused, partly because the limitations of current systems to support multi-device use promote the use of a single device. Over time, this resulted in large collections of unused devices. Multi-device ecologies might provide opportunities to extend the life of old devices. The old devices could adopt more specialized or supporting roles in the device ecology, for example, an old tablet could be permanently attached to the kitchen wall to support cooking activities, or an old computer could be used as a home server.

### Generalizability of Results and Future Work

As we wanted to gain insights into the current practices and behaviors in combining multiple information devices, we decided to approach the topic though a detailed analysis of a limited number of subjects and cases, emphasizing qualitative research methods. While this approach provided us with a rich picture of the current practices and the underlying motivations and needs, it is not possible to statistically estimate the frequency of the observed behaviors in the overall population. In order to validate the results, a quantitative study, such as a survey or a logging study, with a larger sample would be useful.

It should also be noted that the study results reflect current practices – future devices and technologies may enable new practices, as can also be seen when comparing these results with the earlier similar studies. Also, while the participants represented a rather diverse sample in terms of occupations, they were still relatively advanced users of technology with generally high level of education living in a western

country. Different user groups in different parts of the world might display different behavior. Another possible future direction could be more focused studies addressing specific user groups, applications, or contexts of use.

## CONCLUSION

We have presented a user study on people's practices in combining multiple information devices in their everyday lives, ranging from pragmatic tasks to leisure and entertainment activities. Based on diaries and interviews of 14 participants, we have characterized the usage practices of the most common devices, including smartphones, computers, tablets, and home media centers. We have analyzed 123 real-life cases of multi-device use and identified the main usage patterns, including Sequential Use, Resource Lending, Related Parallel Use, and Unrelated Parallel Use. Additionally, we have observed three levels of decisions that determine which devices are used in a particular situation, including acquiring, making available, and selecting the devices for use. We have also discussed the practical challenges related to owning and operating several information devices together, including content access, maintenance, and energy management.

While the participants wanted all their devices to seamlessly work with each other, in practice they continuously encountered problems in multi-device use. Of the multi-device use patterns, Sequential Use and Resource Lending were relative well supported by current devices and systems, but there was little technical support for Related Parallel Use even though it was found to be common among the study participants. Current cloud-based storage solutions were found to have several weaknesses in supporting multi-device use. Finally, improved support for multi-device use might also provide opportunities to extend the life of old devices by allowing the old devices to take more specialized or supporting roles in the device ecology.

## ACKNOWLEDGMENTS

## REFERENCES

1. Bandelloni, R., and Paternò, F. Migratory user interfaces able to adapt to various interaction platforms. *Int. Journal of Human-Computer Studies* 60, 5-6 (2004), 621-639.

2. Bardram, J. Activity-based computing for medical work in hospitals. *ACM Transactions on Computer-Human Interaction* 16, 2 (2009).

3. Dearman, D. and Pierce, J. It's on my other computer!: computing with multiple devices. In *Proc. CHI '08*, 767-776.

4. Google. The New Multi-Screen World Study. https://www.thinkwithgoogle.com/research-studies/the-new-multi-screen-world-study.html

5. Grudin, J. Partitioning digital worlds: focal and peripheral awareness in multiple monitor use. In *Proc. CHI '01*, 458-465.

6. Holtzblatt, K., Wendell, J. B. and Wood, S. *Rapid Contextual Design.* Morgan Kaufmann, 2004.

7. Jung, H., Stolterman, E., Ryan, W., Thompson, T. and Siegel, M. Toward a framework for ecologies of artifacts: how are digital artifacts interconnected within a personal life? In *Proc. NordiCHI '08*, 201-210.

8. Kane, S., Karlson, A., Meyers, B., Johns, P., Jacobs, A. and Smith, G. Exploring cross-device web use on PCs and mobile devices. In *Proc. INTERACT '09*, 722-735.

9. Karlson, A., Iqbal, S., Meyers, B., Ramos, G., Lee, K. and Tang, J. Mobile taskflow in context: a screenshot study of smartphone usage. In *Proc. CHI '10*, 2009-2018.

10. Karlson, A., Meyers, B., Jacobs, A., Johns, P. and Kane, S. Working overtime: patterns of smartphone and PC usage in the day of an information worker. In *Proc. Pervasive '09*, 398-405.

11. Mainwaring, S., Anderson, K. and Chang, M. Living for the global city: mobile kits, urban interfaces, and Ubicomp. In *Proc. UbiComp '05*, 269-286.

12. Müller, H., Gove, J. and Webb, J. Understanding tablet use: a multi-method exploration. In Proc. *MobileHCI '12*, 1-10.

13. Nacenta, M. *Cross-display object movement in multi-display environments*. PhD thesis, University of Saskatchewan, 2009.

14. Norman, D. *The Invisible Computer.* MIT Press, 1999.

15. Oulasvirta, A. and Sumari, L. Mobile kits and laptop trays: managing multiple devices in mobile information work. In *Proc. CHI '07*, 1127-1136.

16. Santosa, S., and Wigdor, D. A field study of multi-device workflows in distributed workspaces. In *Proc. UbiComp '13*, 63-72.

17. Terrenghi, L., Quigley, A. and Dix, A. A taxonomy for and analysis of multi-person-display ecosystems. *Personal and Ubiquitous Computing* 13, 8 (2009), 583-598.

18. Uzun, E., Saxena, N. and Kumar, A. Pairing devices for social interactions: a comparative usability evaluation. In *Proc. CHI '11*, 2315-2324.

19. Weiser, M. The computer for the 21st century. *Scientific American* 265 (1991), 94-104.

**Publication P2**

Tero Jokela and Andrés Lucero. 2013. A Comparative Evaluation of Touch-Based Methods to Bind Mobile Devices for Collaborative Interactions. In *Proceedings of the 31st SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 3355–3364. DOI: https://doi.org/10.1145/2470654.2466459

# A Comparative Evaluation of Touch-Based Methods to Bind Mobile Devices for Collaborative Interactions

**Tero Jokela, Andrés Lucero**
Nokia Research Center
P.O. Box 1000, FI-33721 Tampere, Finland
{tero.jokela, andres.lucero}@nokia.com

## ABSTRACT

We present a comparative evaluation of two touch-based group-binding methods, a leader-driven method and a peer-based method, against a more conventional group-binding method based on scanning and passwords. The results indicate that the participants strongly preferred the touch-based methods in both pragmatic and hedonic qualities as well as in the overall attractiveness. While the leader-driven method allowed better control over the group and required only one participant to be able to form a group, the peer-based method helped to create a greater sense of community and scaled better for larger group sizes and distances. As the optimal group-binding method depends on the social situation and physical environment, the binding methods should be flexible, allowing the users to adapt them to different contexts of use. For determining the order of the devices, manual arrangement was preferred over defining the order by touching.

## Author Keywords

Collocated interaction; mobile phones; user interfaces; device ecosystem binding; group association; pairing.

## ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

## INTRODUCTION

Mobile devices were originally conceived as, and have traditionally been, very personal devices targeted at individual use. Recent advances in sensor and short-range communication technologies offer new opportunities for collaborative use of mobile devices. Groups of collocated users can couple their devices together and create ecosystems of interaction [21]. This allows the users to engage in collaborative activities and experiences with their mobile devices, thus shifting from *personal-individual* towards *shared-multi-user* interactions. Examples of applications that would benefit from such collaborative use of mobile devices include sharing of digital content,

collaborative creation and editing of content, and different kinds of games. In many of these applications, it would be natural to utilize spatial interactions in the shared space, for example, throwing virtual objects such as files between devices. However, finding the positions of the devices has presented a challenging problem, requiring the use of special tracking equipment or dedicated infrastructure.

Before a group of users can engage in collaborative interactions with their mobile devices, the multi-device ecosystem must first be set up. This involves initiating the necessary system and application software in all devices. The devices must become aware of the other devices existing in the proximity, and the devices intended to participate in the ecosystem must be identified. A communication channel then needs to be established between the devices participating in the ecosystem, in order to allow exchange of data and coordination of the interactions. Wireless short-range communication technologies such as WLAN or Bluetooth are typically used to exchange data between devices. The process of setting up the ecosystem is generally known as device binding or ecosystem binding [21] (also known as device association, pairing, or coupling [3]). As the intention is to enable spontaneous interactions, it should be possible to bind devices having no prior knowledge of each other in a fast and easy way. If the process of binding devices is too complicated or tedious, the users might lose interest in using multi-device interactions in the first place. As the wireless connections provide no physical indications (for example, cables) of which devices are actually connected, the binding process should provide sufficient security and cues so that the users can ensure that the right devices are connected.

In this paper, we are concerned with device-binding methods for establishing an ecosystem of mobile devices to support collaborative interactions within small-to-medium-sized groups of collocated users. While the problem of a single user pairing two devices has been extensively studied in prior research, researchers have started to address more complex scenarios involving multiple users and devices only recently. In particular, we focus on methods based on device proximity and touch interactions, which have been found to be intuitive and easy to explain, but which have been little explored in the literature [2]. We present a comparative evaluation of two touch-based group-binding methods, a leader-driven method called Host and a peer-

based method called Ring, against a more conventional method called Seek, which is based on scanning the available devices in the proximity and passwords for security. While most earlier studies on device binding have focused on pragmatic aspects such as security and usability, we approach the problem from a broader user experience perspective, covering also hedonic aspects such as social and emotional factors, which have been shown to be important considerations when users select binding methods in real-life situations [10, 18]. We consider the complete group creation process in a realistic application context, including identification of the devices to participate in the group, initiation of the application software in all devices, and authentication of the connection. We also explore options to determine the device order during the group creation phase, in order to allow spatial interactions without dedicated tracking equipment. The evaluation results indicate that the participants strongly preferred touch-based methods over Seek. Several important differences were identified between leader-driven and peer-based methods. The optimal group-binding method was found to depend on various social and environmental factors, suggesting that the binding methods should be flexible to allow users to adopt different group creation strategies in different contexts of use. For determining the order of the devices, manual arrangement was preferred over defining the order by touching.

The rest of this paper is structured as follows. First, we provide a brief overview of the related work. We then give a detailed description of the three group-binding methods and the evaluation procedure. Finally, we present the results of the evaluation, followed by conclusions.

## RELATED WORK

The problem of device binding has been extensively studied in the fields of human-computer interaction and security research. A wide range of methods for device binding has been proposed – in security research alone, over 20 different methods have been identified [17]. These methods vary in terms of device hardware requirements, amount of user involvement, and level of provided security.

The problem of device binding can be divided into two subproblems: device *identification* and *authentication*. Device identification involves selecting which of the devices available in the proximity should be bound with each other. The need for device authentication originates from the invisibility of wireless communications. As the users cannot see the wireless communication channels, they cannot be sure that they are really connecting to the other devices intended to, opening the possibility for so-called Man-in-the-Middle attacks. To counter this threat, a wide variety of methods have been proposed that authenticate the wireless connection over auxiliary communication channels (also known as Out-of-Band Channels), which can be perceived and managed by human users.

The most common device-binding methods today, such as those typically used in Bluetooth and WLAN networks, are based on scanning the environment for available devices and then presenting a list of the found devices to the user for selecting the other device to bind with. The authentication is based on short strings (also known as PIN codes) that the user is expected to copy or compare between devices. The authentication strings can be represented as numbers, words, graphical images, or audio signals in the user interface.

The proposed alternative methods include a variety of techniques based on synchronous user actions, for example, pressing buttons on both devices [19] or touching both devices [23] simultaneously, shaking the devices together [8], or bumping the devices together [6]. Bumping is also used in the popular commercial service Bump[1]. Further, device binding can be based on continuous gestures spanning from one device display to another [7]. Methods based on spatial alignment of the devices include pointing, for example, with laser light [15], touching [20], or placing the devices in close proximity of each other [12]. It is also possible to bind devices with various auxiliary devices, for example, tokens [1] or cameras [16]. Some of the proposed methods cover only device identification or authentication, while others combine both identification and authentication into a single user action.

The development of new binding methods has been largely technology-driven with little user involvement. As an example of a more user-centered approach, Chong and Gellersen [2] present a study on users' spontaneous actions for device binding. In the study, the users' were asked to invent methods for binding together low-fidelity acrylic prototypes of different devices. Device proximity and touch based methods were found to be among the most commonly proposed methods, and the physical contact of devices was also considered as the easiest method to describe and teach to another person. Still, there has been little work exploring such techniques in the literature.

Binding methods are not just means for connecting devices – they have strong social and emotional aspects. In real-life situations, the users do not always use the easiest or fastest method available, nor the one they like best. Many factors influence their choice of binding method, including the place, the social setting, the other people present, and the sensitivity of data [10, 18]. Users are willing to take security risks to comply with social norms [10].

The vast majority of prior research has focused on scenarios of a single user binding two devices with each other (for example, binding a headset with a mobile device). Only recently have researchers started to consider more complex scenarios involving multiple users and devices. Such multi-user scenarios differ in many respects from single-user
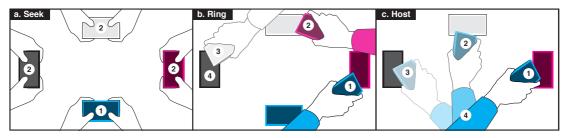
---

[1] http://bu.mp/

**Figure 1. Three group-binding methods. a) Seek: a leader creates a group and shares a password (1), which is then entered in parallel by the other participants (2). b) Ring: one person starts the app and touches the next device to their right to add it (1), then others continue adding the next person to their right (2,3), and the last person completes the group (4). c) Host: a leader starts the app, adds people by touching all devices in counter-clockwise order (1-3), and puts the device on the table to complete the group (4).**

scenarios, making the single-user device-binding methods not necessarily applicable to multi-user scenarios. In multi-user scenarios, communication between group members provides an additional source for potential errors. On the other hand, the users are typically willing to help each other and make decisions by mutual agreement, which reduces the amount of errors [11]. Methods that involve physical exchange of devices have been found to be unacceptable unless the users know each other very well, as the users are unwilling to hand in their devices to strangers [22].

Chong and Gellersen [3] present a framework that summarizes and classifies the different factors that influence the usability of spontaneous device binding, identifying technology, user interaction, and application context as the three most important criteria.

**EVALUATION OF GROUP-BINDING METHODS**

**Objectives**
In this study, we were primarily interested in three research questions. First, we wanted to compare touch and proximity-based methods for group binding against more conventional methods based on scanning and passwords. Second, we wanted to explore different ways to divide the group-binding task between the participants – in particular, we were interested in differences between leader-driven and peer-based methods. Third, we wanted to investigate possibilities to define the device order as a part of the group creation process, in order to allow implementation of spatial interactions without extensive tracking equipment.

**Group-Binding Methods**
To study these research questions in practice, we designed three different group-binding methods called Seek, Ring, and Host. The Seek method represented the conventional approach used, for example, in network games and was based on scanning for device identification and passwords for authentication. Both the Ring and Host methods used touch for device identification and authentication. The main difference between Ring and Host was that Ring was peer-based, distributing the group creation task between all participants, while Host was leader-driven, concentrating

the group creation task on one participant. Additionally, Host utilized device gestures for some interactions. The Host method was based on the EasyGroups method [14] reported earlier. While all the methods were generic, we decided to study them in the context of a simple photo sharing application in order to provide a more realistic goal for the group creation task during the evaluation. The photo sharing application was a simplified version of Pass-Them-Around [13] and it allowed the users to browse a collection of photos stored in their own devices and supported spatial interactions of throwing photos between devices.

*Seek*
To set up a new group, one person (the leader) should start the Seek application on their device and create a new group (Fig. 1a). The application prompts the leader to join a WLAN network and enter a name for the new group. The application automatically generates a six-digit password for the group. The application then moves to the Table Overview (Fig. 2) showing all devices that are currently part of the group and their order as well as the group name and password. As new devices join the group, an animation shows how the device order changes on the table. To enable the users to identify the devices, the color of each device is indicated on the screen. The other persons can then join the group in parallel by starting the Seek application, joining the same network as the leader, and selecting the existing group from the list. The application then prompts the user to enter the password. If the password is correct, the device joins the group and moves to the Table Overview. If the order of the devices presented on the screen is different from the order of the devices on the table, the leader can correct it by dragging the devices to the right positions on the screen. The users can move to the Photo Sharing Mode by tapping their own piles of photos on the screen.

If a new person wishes to join an existing group, the person should start the Seek application and join the group in the same way as during the initial group creation phase. The leader can check the order of the devices on the screen and correct it if necessary. To leave the group, the person should press the "Exit" button on the screen.

**Figure 2. The Table Overview during Seek.**

*Ring*

To begin group formation, one person should start the Ring application on their device (Fig. 1b). This device automatically enters Discovery Mode and visual feedback is shown in portrait view to suggest holding the device vertically for a more comfortable grip. The person holding the device is instructed to touch the next device to their right. When the person moves their device close to the next device, the device detects the new device and the person holding the device is asked to hold their device still while the new device is added to the group. When the new device has been added to the group, the device exits Discovery Mode and moves to the Table Overview, which shows all devices that are currently part of the group and their order. The new device that was just added to the group now automatically starts the application and enters the Discovery Mode. The owner of that device is instructed to continue in the same way and touch the next device to their right. By asking the user always to connect to the next device to their right, we are able to define the order of the devices on the table based on the touching order. When all the devices around the table have been added to the group, the owner of the last device can complete the group by pressing the "Complete" button on screen. The users can move to the Photo Sharing Mode and start sharing pictures.

If a new person wishes to join an existing group, the person on the left side of the new person can press the "Add Device" button on the screen and touch the incoming person's device. The new person is then added to the right side of the person who just added them. The new group member can then continue adding new devices, or press the "Complete" button if there are no more devices to add. To leave the group, the person should press the "Exit" button.

*Host*

To set up a new group, one person (the leader) should start the Host application on their device (Fig. 1c). When the device is picked up from the table, it detects the pick-up gesture and enters Discovery Mode, which allows the leader to add new people to the group. The leader should then touch the other devices one by one in counter-clockwise order around the table (Fig. 3). The order of the

devices on the table is automatically defined based on the touching order. A similar procedure (and visual feedback) as the one described for Ring is used to detect, connect, start the application, and join the group. When all the other devices have been added to the group, the leader should put their device back on the table. The device detects the gesture and exits the Discovery Mode and completes the group set-up. The persons can then start sharing photos between devices.

If a new person wishes to join an existing group, the person on the left side of the new person should pick up their device to enter Discovery Mode and touch the new person's device. The new person is then added next after the person who just added them. To leave the group, the person should pick their device up from the table and flip it upside down. The device detects the gesture and exits.

**Prototype Implementation**

We built prototypes of the three group-binding methods on Nokia N9[2] mobile devices running the MeeGo operating system. The prototypes were implemented in C++ on top of the Qt 4.7 software framework. QML and Qt Quick with OpenGL ES hardware acceleration were used for fluent animated user interface graphics. The N9's internal accelerometer was used for gesture detection in Host.

In all methods, the objective was to establish a WLAN connection between the devices. In Seek, each device was manually connected to the WLAN network. The device then scanned the network for available groups and presented a list to the user to choose from. In Ring and Host, touching was detected with Bluetooth-based radio technology, which was able to detect other devices at ranges closer than 20 cm in approximately 5 seconds. While the technology generally worked reliably, there were occasionally longer delays before the other devices were detected or detections of devices further away. The necessary connectivity and initialization information was then sent to the discovered device over Bluetooth. A daemon, which listened to a Bluetooth socket, received the



**Figure 3. The Host method. The user holding the cyan device has connected the black (right) and magenta devices (top).**

---

[2] http://swipe.nokia.com/

connectivity information on the discovered device and started the actual application, which connected to the correct WLAN network and joined the group. The prototypes were fully functional with real network communication, except for the security protocols, which were only simulated in the user interface.

## Participants

We recruited a total of 24 participants for the evaluation by posting an advertisement on a local mailing list. Of the 24 participants, 20 were pairs of users, while the remaining four were individual participants. We preferred to recruit pairs of people who knew each other, so that the participants would feel more comfortable during the evaluation session. We assigned the participants into six groups of four users in the order they registered for the study. Each participant typically knew one other participant in the group, while the two others were strangers. Eight of the participants were female and 16 male. The ages of the participants varied between 23 and 45 years (M=33.6, SD=6.0). Three of the participants were left-handed and 21 right-handed. The participants represented a variety of different backgrounds, with eight participants having a software engineering background, 10 other technical background (for example, mechanical engineering), and six non-technical background (for example, administration or linguistics). The participants were fairly advanced users of technology: on a scale between 1 and 7 (1=novice, 7=expert), the participants rated their familiarity with technology above average (M=5.1, SD=1.2). All participants were active smartphone users and six of the 24 participants had used a Nokia N9 before the study.

## Procedure

We organized a series of six evaluation sessions. The evaluation sessions were arranged in a usability laboratory of approximately 40 $m^2$ (430 sq ft) in size. Fig. 4 shows the evaluation setup. In each session, there were four participants and a moderator present. We used devices of four different colors (black, white, magenta, and cyan) and each participant was assigned a device with a different color. This provided a practical method of identifying the devices of the different participants during the evaluation session. The participants were given seats around a rectangular table of approximately 150x70 cm (60x27 inches) in size, one on each side of the table. The table was carefully selected so that there would be different distances between the participants and that the participants sitting on the short edges would have some difficulty reaching each other. The total durations of the evaluation sessions varied between 100 and 120 minutes.

As the participants arrived in the laboratory, the moderator guided them to their seats around the table and asked them to fill in a background questionnaire form. When all the participants had completed the forms, the moderator introduced the participants to the idea of collaborative use



**Figure 4. Evaluation setup with four participants.**

of mobile devices and demonstrated it with the photo sharing application. The participants were then given their own devices and they were encouraged to try throwing photos between devices. This small introductory task provided the users with an opportunity to become familiar with their devices. The moderator then explained to the participants that before they could share photos between devices by throwing, they first had to bind their mobile devices together into a group and the objective of the session was to evaluate different methods for that task. Before the actual evaluation started, the moderator informed the participants that some of the methods might require touching other devices and demonstrated how to do it in practice. The participants were then asked to practice touching with their own devices. We saw this training step necessary, because while many of the participants were aware of touching as an interaction technique, few had tried it in practice.

To begin the actual evaluation, the moderator showed a short video clip demonstrating the first group-binding method. The videos were prepared so that they simulated a situation of a participant observing another group of users using the method to create a group. We used video recordings to minimize the variations between the instructions that the different groups received. After the participants had watched the video, the moderator gave them the following task: "By using the method that was just demonstrated to you, create a group so that you can throw photos between your devices." The moderator then observed as the participants tried out the method and only intervened if the participants clearly could not proceed with the method or there were some technical problems with the devices. The task was considered complete, when the participants could successfully throw photos between all devices. The moderator then asked everybody to leave the group and create another group with a different participant initiating the group creation. The moderator also asked at least one person to leave the group and rejoin it. Overall, each group tried each method two to four times.

After testing the method, the moderator asked the participants to fill in two validated questionnaires. The first questionnaire was NASA-TLX [4], which measures the subjective workload experience when performing a task. To

gain a broader view of the methods, we extended the questionnaire with four additional scales: learnability, quickness, security, and overall preference. The second questionnaire was AttrakDiff [5], which measures the attractiveness of interactive products.

The same procedure was then repeated for the second and the third methods. We systematically varied the order in which the six groups were exposed to the three methods to counter-balance any learning effects.

After the participants had tested all the methods, the moderator interviewed the participants about their experiences with the methods. The interview was semi-structured and covered a variety of themes including general feedback about the different methods, perceptions about their learnability and security, as well as specific interaction techniques like touching and hand gestures. The moderator also showed the participants three pictures representing different scenarios and asked them to consider what would be the most appropriate method for creating a group in each scenario. The scenarios were: 1) meeting other family members in the living room at home, 2) meeting representatives of another company in a meeting room at the office, and 3) meeting friends in a busy café. The objective was to encourage the participants to think about different situations and environments and their social and physical characteristics. After the interview was completed, the moderator thanked the participants and gave them a movie ticket each to compensate them for their time.

All sessions were video recorded and interaction with the devices was logged. Two researchers independently analyzed the video recordings and wrote notes about their observations. The same two researchers then analyzed the data and built an Affinity Diagram [9] in a series of interpretation sessions. Each researcher individually studied the notes and grouped them into clusters of related items. The clusters then evolved to broader categories that were naturally revealed and were jointly revisited, discussed, and refined. In the end, the categories were processed into more general findings that form the core of the Results section.

## RESULTS
We first give an overview of the quantitative results. We then present the qualitative results and contrast them with the quantitative results when relevant.

### Extended NASA-TLX
Fig. 5 illustrates the results of the extended NASA-TLX questionnaire [4]. The main bars indicate the means for each subscale, while the error bars indicate standard errors. The original six subscales of NASA-TLX are presented on the left and the four subscales that we added for the purposes of this study (learnability, quickness, security, and overall preference) are on the right. As the participants were observed in groups, the responses of each participant were influenced by the other participants in the same group.
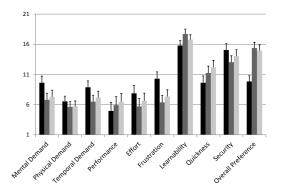


**Figure 5. Extended NASA-TLX results.**

Therefore, we used mixed model techniques to analyze the data with the binding method as a fixed factor and the groups and the participants nested in the groups as a random component. The results indicate that the binding method had a significant effect on mental demand ($F_{(2, 44.54)}$ = 8.39, p = .001), frustration ($F_{(2, 36.92)}$ = 9.54, p < .001), and overall preference ($F_{(2, 37.18)}$ = 22.16, p < .001). Pair-wise comparisons with Bonferroni correction show that the levels of mental demand (p = .001) and frustration (p < .001) for Seek were significantly higher compared to Ring, and that the level of overall preference was significantly higher for both Ring (p < .001) and Host is (p < .001) compared to Seek. There were no significant differences between Ring and Host on any of the subscales.

### AttrakDiff
Fig. 6 illustrates the results of the AttrakDiff questionnaire [5] for the three group-binding methods. Pragmatic quality (PQ) refers to the product's ability to support the achievement of behavioral goals (usability). Hedonic quality refers to the users' self: stimulation (HQ-S) is the product's ability to stimulate and enable personal growth, while identification (HQ-I) is the product's ability to address the need of expressing one's self through objects one owns. Perceived attractiveness (ATT) describes a global value of the product based on the quality perception.

We analyzed the AttrakDiff data with the same methodology as the extended NASA-TLX data. The results
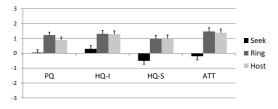


**Figure 6. AttrakDiff results.**

indicate that the binding method had a significant effect on all dimensions PQ ($F_{(2, 35.63)}$ = 15.74, $p < .001$), HQ-I ($F_{(2, 32.70)}$ = 60.37, $p < .001$), HQ-S ($F_{(2, 41.98)}$ = 52.66, $p < .001$), and ATT ($F_{(2, 31.52)}$ = 55.79, $p < .001$). Pairwise comparisons with Bonferroni correction show that the levels of PQ for Ring ($p < .001$) and Host ($p = .002$) were significantly higher compared to Seek, and that the levels of HQ-I, HQ-S, and ATT for both Ring and Host were significantly higher compared to Seek (all $p < .001$). There were no significant differences between Ring and Host on any of the dimensions.

### Performance

Seek was the most reliable method with all group creation attempts succeeding without moderator assistance. With Ring, two of the six groups failed their initial attempts because several participants started the application simultaneously. With Host, two groups failed their initial attempts because of multiple participants starting the application and two groups because of incorrect touching order. After solving these initial difficulties, all groups were able to successfully create groups with all the methods.
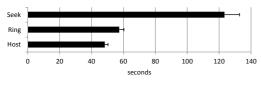


**Figure 7. Completion times.**

We measured the fastest completion time for the group creation task for each method in each of the sessions from the video recordings and device logs. The participants were instructed to create a group as they would in a real-life situation. If the participants clearly performed the group creation task in a non-optimal way, for example, encountered problems or started to explore different features, the moderator asked them to repeat the task until the process was completed smoothly. Fig. 7 illustrates the mean completion times, with the error bars showing the standard errors. The fastest method was Host, followed by Ring. Seek was clearly slower than the two touch-based methods. While the questionnaire results on perceived quickness show similar order, the distinctions are smaller with no statistically significant differences.

### Attractiveness

Half of the participants (12/24) commented that Seek was old-fashioned and boring. *"[P23] Seek was so 90's, engineering style."* Further, many participants (10/24), especially the ones that were less technologically oriented, commented that Seek was far too technical for them. They felt that Seek had too many steps and it was too complicated to use. *"[P16] Seek is too technical. Predictable but not intuitive. Not fun to use."* Compared to Seek, the touch based methods, especially Ring, were

considered to be novel, intuitive, and simple to use. *"[P20] I think Ring is very stylish. It is new… I am not a very technical person, but Ring was simple to use and understand what was happening."*

These qualitative findings are supported by the quantitative results. In NASA-TLX (Fig. 5), Seek was rated significantly higher in mental effort and frustration compared to Ring. On the AttrakDiff questionnaire (Fig. 6), both Host and Ring were rated significantly more attractive (ATT) than Seek.

### Acting as a Group

During the group creation task, the participants clearly acted together as a group instead of individuals. There was rich interaction between the participants, suggesting and agreeing the next actions and confirming the results. The participants were eager to help each other, if they noticed that some other participant was experiencing problems with the system. This contributed to the high success rate of the group creation tasks. The attention of the participants was divided between their own devices, and the devices and actions of the other participants. In touch-based methods, the touching actions were clearly visible to everybody, making it easier to follow the situation already before the participants' own devices joined the group and started to provide feedback about the system status. In Seek, the participants were forced to check the status by asking verbally or by peeking on the other users' screens.

### Leader-Driven vs. Peer-Based Group Creation

One of the main differences between Ring and Host was that in Host, the group-binding process was driven by one person (the leader) who did most of the work, while in Ring, all participants contributed to the group-binding process as equal peers. This had several interesting effects.

Host provided the leader with control over who could join the group. Many participants (13/24) considered that this would be an important feature in some situations. *"[P20] If there were people [around] that I didn't know so well, like at my workplace, Host would be the best [method] because I could control with whom I share."* Some participants (5/24) suggested that the leader should also be able to force participants to leave the group. Further, Host allowed one person to create a group for everybody, so that the others did not have to do anything. Some participants (9/24) commented that it was good that only one person who was able to create a group was required, for example, if some of the participants were less technologically oriented than the others, or if some of the participants were not fully able to use their devices because of some situational factors (for example, because they had children sitting on their knees).

During the evaluation sessions, the participants were very polite towards each other in selecting the leader. However, as commented by one of the participants, selecting the leader might be more challenging in real-life situations,

involving complex group dynamics and cultural factors. *"[P12] How can this guy be the leader, if [another person] is the senior? Or if the oldest guy is the leader, he might not know much about technology. Or with youngsters, if there is one who is the leader of the group, how does the group creation go?"* On the other hand, Host was considered as natural in situations where there was a clear leader, for example, in official meetings.

Almost half of the participants (11/24) felt that Ring brought people more together and helped to create a greater sense of community, because everybody was equally involved in creating the group and was forced to interact with the others by touching their devices. *"[P5] Ring makes a spiritual chain between participants. It makes you feel better."* Participants compared Ring *to "[P14] passing the torch"* or *"[P12] shaking hands"*, and commented that it helped to *"[P10] break the ice"* and *"[P9] take down the barriers."* Ring was considered to be particularly suitable for informal situations where there was no strict hierarchy, for example, when meeting a group of friends.

While Host worked well for small groups with all participants located near each other, most participants (16/24) commented that it would not work for larger groups because it would be tiring for the leader to touch a large number of devices, nor longer distances because the leader could not reach all other devices without moving around. Also other factors, for example, having dinnerware on the table, might make it difficult for the leader to touch the other devices. Some participants (10/24) commented that Ring would scale better to larger groups and distances. One participant contrasted the difference between Host and Ring with distributing handouts in meetings. *"[P6] In large meetings, there is no time to give handouts to everybody one at a time. You circulate them."* On the other hand, in a large scattered group, it might be difficult to know who is the last person and should complete the group.

### Touching

In Ring and Host, identification of the devices intended to participate in the group was based on touching. Almost half of the participants (11/24) commented that touching was an easy and intuitive way to add participants to the group. *"[P19] Touching to join was a clear, physical, easy, natural way to bring someone into the group."* On the other hand, some participants (9/24) commented that touching could be socially awkward, for example, in formal situations, and brought about privacy issues. *"[P13] Touching is the same as using the other person's phone myself."* In the case of group creation, however, there was a clear reason to touch the other person's phone, so it did not feel like an invasion of privacy. Many participants (10/24) spontaneously pushed their devices forward when another user approached to touch it. This might simply have been a polite gesture to make it easier for the other participant to reach the device, but it could also have indicated giving a permission to touch one's personal device. Finally, some participants (6/24) stressed that to be useful, touching should be detected fast and work very reliably.

### Other Gestures

In addition to touching, Host also used gestures for two other purposes. The first gesture allowed the participants to leave the group by flipping their devices upside down. Most participants (19/24) flipped their devices by putting them upside down on the table – only a few flipped their devices in their hands. Some participants (8/24) commented that flipping was a novel, simple, and entertaining way to leave the group. On the other hand, some participants (9/24) raised concerns that it was difficult to know and remember the gesture and it was easy to do it accidentally.

The second gesture enabled the participants to move between Photo Sharing and Discovery Modes by putting their devices on the table and picking them up. Half of the participants (12/24) commented that they did not like this feature because holding the device in their hands was the natural way to use the device and allowed them to control the privacy and viewing angle of their screens and because there might not always be a table available to put the device on. *"[P21] Keeping the device on the table is not something I usually do. I usually hold the device in my hand."*

### Ordering

In order to allow throwing of photos between devices, the participants had to define the order of the devices on the table. In Seek, this was done manually by the leader, while in Ring and Host, the participants were expected to touch the devices in counter-clockwise order and the order of the devices was automatically determined based on the touching order. Almost all participants (20/24) considered the requirement to touch the devices in a specific order too restrictive, difficult to remember, and unforgiving to errors. *"[P17] I did not like that you had to go in [counter-]clockwise order. Why not the other way? It should work both ways. It is difficult to remember and learn."* Instead, the participants liked the flexibility and robustness that the manual reordering provided to them. *"[P12] Being able to easily change the order would be the number one feature for me."* The participants pointed out several cases, where manual reordering would be beneficial, for example, if there was a human or technical error in the initial group creation phase, or if the participants moved or changed places. Almost half of the participants (11/24) considered the colored dots, which identified the devices on the screen inadequate, and proposed that textual names should be used in addition to the color.

### Perceived Security

In Seek, security was based on six-digit authentication strings that were automatically generated by the system. The participants who wanted to join the group had to manually copy and enter the authentication string into their devices. The participants considered the authentication

strings as passwords that they were familiar with in other systems. The dominant way of sharing the password was that the leader read the password aloud. Typically, the password had to be repeated many times as not all the participants were ready to enter it at the same time, or some of the participants missed parts of it. In only one of the six sessions, the participants shared the password by putting the device of the leader at the center of the table, so that everybody could read the password from the screen. However, also in this case some of the participants sitting further away from the leader had difficulties in obtaining the password because they could not clearly see the screen. Most participants (14/24) considered the passwords awkward and would have preferred some other security mechanism. *"[P18] If you need that security level, there must be a better way than [passwords]."* Some participants proposed improvements to the passwords used, for example, making the passwords shorter, using common words, or allowing the participants to define the passwords. Half of the participants (12/24) considered sharing the password verbally as a security risk as anybody in the proximity could hear it. In that sense, the passwords were thought to provide a false sense of security. *"[P14] Password is a complication without any security element."*

In Seek and Ring, security was based on physical proximity enforced by the short range of the touch detection technology. Compared to passwords, which were familiar to all, this was a new concept to the participants. Most participants (13/24) considered that touching provided adequate security for scenarios like sharing photos, provided that the detection technology works reliably and the range is not too long. *"[P22] If phones have to touch, it is quite safe. If somebody I don't know comes so close, I would be alert anyway."* This finding is also supported by the extended NASA-TLX results, which indicate no significant differences in perceived security between Seek and the touch-based methods. Still, some participants (8/24) raised concerns over unauthorized persons accessing their devices by touching, for example, when they had their devices in their pockets in a crowded bar or in a queue.

## DISCUSSION

### Seek vs. Touch-Based Methods
Both quantitative and qualitative results of the user evaluation show that touch-based methods provide a promising alternative to dominant scanning and password based group-binding methods. While Seek was familiar and reliable in practice, it was considered to be technical, complicated, old-fashioned, and boring. Overall, the participants clearly preferred the touch-based methods and considered them to be simple and intuitive as well as novel and enjoyable to use. The touch-based methods were also faster and they allowed the participants to better maintain awareness of the status of the group formation task as the touching actions could easily be perceived by everybody. Regarding security, touching was considered to be equally

secure to passwords. However, to work well in practice, touch detection should be fast and it should work reliably only within the defined distance.

### Leader-Driven vs. Peer-Based Methods
The group-binding task can be divided in different ways between the participants. The study results show that different approaches have different strengths and weaknesses. The leader-driven methods, which concentrate the task on a single participant, enable the leader to have strong control over the group and require only one person who is able to create a group. On the other hand, selecting the leader may add more complexity to the group creation process. The peer-based methods, which distribute the work between all participants, help to create a stronger sense of community and scale better to larger numbers of participants and distances. The study results indicate that there is no single optimal method, but the best method depends on the application, social situation, and physical environment. Therefore, the group-binding methods should not strongly enforce a single group creation procedure, but allow for flexibility, so that the participants could adapt the method to the particular needs of each situation.

### Device Ordering
The group-binding methods also allowed the determination of the device order using two different approaches: arranging the devices manually or defining the order by touching. The study results indicate that the participants found the requirement to touch the devices in a specific order too restrictive and preferred to touch the devices in a free order and then arrange the devices manually. Again, the optimal touching order depends on social and environmental factors and the group-binding methods should allow the participants to adapt the touching order to each situation. Also, flexible touching order allows the participants to better recover from human and technical errors that may occur during group creation. A well-defined relationship between the touching order and the initial positions of the participants might still be useful for advanced users who want to optimize the group creation process for efficiency.

### Supporting Self-Expression and Playfulness
We observed an overall positive mood where participants collaborated and helped each other during group creation. On top of that, we also noticed participants were often laughing, making jokes by creating funny group names, celebrating their collective successes by cheering when they had successfully created a group, and describing the touch-based methods as *"[P8] this is like some Enterprise stuff from Star Trek."* These situations bring to our attention that we are not purely dealing with connecting devices together, but that people are looking for an overall experience that allows them to express themselves and be playful. Therefore, the group-binding methods should look beyond the purely functional task of connecting devices and sharing

information, and aim to also engage users on other aspects such as supporting self-expression and playfulness.

## CONCLUSION

We have presented a comparative evaluation of two touch-based group-binding methods, a leader-driven method called Host and a peer-based method called Ring, against a more conventional method called Seek, which was based on scanning the available devices in the proximity and passwords for security. The results indicate that the participants strongly preferred the touch-based methods in both pragmatic and hedonic qualities as well as in overall attractiveness. In terms of perceived security, touching was considered equally secure to passwords. While Host allowed better control over the group and required only one participant to be able to form a group, Ring helped to create a greater sense of community and scaled better for larger group sizes and distances. As the optimal group-binding method depends on the social and physical environment, the binding methods should be flexible, allowing the users to adapt them to different contexts of use. For determining the order of the devices, manual arrangement was preferred over defining the order by touching.

## ACKNOWLEDGMENTS

## REFERENCES

1. Ayatsuka, Y. and Rekimoto, J. tranSticks: physically manipulatable virtual connections. In *Proc. CHI '05*, 251-260.

2. Chong, M. and Gellersen, H. How users associate wireless devices. In *Proc. CHI '11*, 1909-1918.

3. Chong, M. and Gellersen, H. Usability classification for spontaneous device association. *Personal and Ubiquitous Computing*, published online (2011).

4. Hart, S. G. and Staveland, L. E. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In Hancock and Meshkati (eds.) *Human Mental Workload*, North Holland Press, 1988.

5. Hassenzahl, M. The Interplay of Beauty, Goodness, and Usability in Interactive Products. *Human-Computer Interaction* 19, 4 (2004), 319-349.

6. Hinckley, K. Synchronous gestures for multiple persons and computers. In *Proc. UIST '03*, 149-158.

7. Hinckley, K., Ramos, G., Guimbretiere, F., Baudisch, P. and Smith, M. Stitching: pen gestures that span multiple displays. In *Proc. AVI '04*, 23-31

8. Holmquist, L.E., Mattern, F., Schiele, B., Alahuhta, P., Beigl, M. and Gellersen, H. Smart-Its friends: a

9. Holtzblatt, K., Wendell, J. B. and Wood, S. *Rapid Contextual Design*. Morgan Kaufmann, 2004.

10. Ion, I., Langheinrich, M., Kumaraguru, P. and Čapkun, S. Influence of user perception, security needs, and social factors on device pairing method choices. In *Proc. SOUPS '10*.

11. Kainda, R., Flechais, I. and Roscoe, A. Two heads are better than one: security and usability of device associations in group scenarios. In *Proc. SOUPS '10*.

12. Kray, C., Rohs, M., Hook, J. and Kratz, S. Group coordination and negotiation through spatial proximity regions around mobile devices on augmented tabletops. In *Proc. TABLETOP 2008*, 1-8.

13. Lucero, A., Holopainen, J. and Jokela, T. Pass-Them-Around: collaborative use of mobile phones for photo sharing. In *Proc. CHI '11*, 1787-1796.

14. Lucero, A., Jokela, T., Palin, A., Aaltonen, V. and Nikara, J. EasyGroups: binding mobile devices for collaborative interactions. In *CHI EA '12*, 2189-2194.

15. Mayrhofer, R. and Welch, M. A human-verifiable authentication protocol using visible laser light. In *Proc. ARES '07*, 1143-1148.

16. McCune, J.M., Perrig, A. and Reiter, M.K. Seeing-is-believing: using camera phones for human-verifiable authentication. In *Proc. SOUPS '05*, 110-124.

17. Nithyanand, R., Saxena, N., Tsudik, G. and Uzun, E. Groupthink: usability of secure group association for wireless devices. In *Proc. Ubicomp '10*, 331-340.

18. Rashid, U. and Quigley, A. Interaction techniques for binding smartphones: a desirability evaluation. In *Proc. HCD '09*, 120-128.

19. Rekimoto, J. SyncTap: synchronous user operation for spontaneous network connection. *Personal and Ubiquitous Computing* 8, 2 (2004), 126-134.

20. Rekimoto, J., Ayatsuka, Y., Kohno, M. and Oba, H. Proximal interactions: a direct manipulation technique for wireless networking. In *Proc. INTERACT '03*, 511-518.

21. Terrenghi, L., Quigley, A. and Dix, A. A taxonomy for and analysis of multi-person-display ecosystems. *Personal and Ubiquitous Computing* 13, 8 (2009), 583-598.

22. Uzun, E., Saxena, N. and Kumar, A. Pairing devices for social interactions: a comparative usability evaluation. In *Proc. CHI '11*, 2315-2324.

23. Zimmerman, T. G. Personal area networks: near-field intrabody communication. IBM Systems Journal 35, 3-4 (1996), 609-617.

**Publication P3**

Tero Jokela and Andrés Lucero. 2014. FlexiGroups: Binding Mobile Devices for Collaborative Interactions in Medium-Sized Groups with Device Touch. In *Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '14).* ACM, New York, NY, USA, 369–378. DOI: https://doi.org/10.1145/2628363.2628376

# FlexiGroups: Binding Mobile Devices for Collaborative Interactions in Medium-Sized Groups with Device Touch

**Tero Jokela, Andrés Lucero**

Nokia Research Center

P.O. Box 1000, FI-33721 Tampere, Finland

tero.jokela@nokia.com, lucero@acm.org

## ABSTRACT

We present a touch-based method for binding mobile devices for collaborative interactions in a group of collocated users. The method is highly flexible, enabling a broad range of different group formation strategies. We report an evaluation of the method in medium-sized groups of six users. When forming a group, the participants primarily followed viral patterns where they opportunistically added other participants to the group without advance planning. The participants also suggested a number of more systematic patterns, which required the group to agree on a common strategy but then provided a clear procedure to follow. The flexibility of the method allowed the participants to adapt it to the changing needs of the situation and to recover from errors and technical problems. Overall, device binding in medium-sized groups was found to be a highly collaborative group activity and the binding methods should pay special attention to supporting groupwork and social interactions.

## Author Keywords

Collocated interaction; mobile phones; user interfaces; device ecosystem binding; group association; pairing.

## ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

## INTRODUCTION

While mobile devices have traditionally been very personal devices targeted at individual use, over the last years there has been a growing interest in systems that combine several mobile devices together to create broader ecosystems of interaction [24]. Such ecosystems allow groups of collocated users to engage in rich collaborative activities and shared experiences with their devices. Potential application scenarios include presenting and collaboratively editing documents in business meetings, sharing photographs and videos within groups of friends in a café,

or playing multi-player games with other family members in the living room.

But before a group of collocated users can engage in collaborative interactions with their mobile devices, they must first join their devices together into a multi-device ecosystem. This is a complex procedure with several steps. The necessary system and application software needs to be initiated on all devices. The devices must discover the other devices existing in the proximity and the devices intended to participate in the ecosystem must be identified. A communication channel then needs to be established between the devices participating in the ecosystem, in order to allow exchange of data and coordination of the interactions. Typically, short-range radio technologies, such as WLAN or Bluetooth, are used to transmit data between devices. This process of setting up the ecosystem is generally known as device binding or ecosystem binding [24] (also known as device association, pairing, or coupling [4]). As the intention is to enable spontaneous interactions, it should be possible to bind devices having no prior knowledge of each other in a fast and easy way. If the process of binding devices is too complicated or tedious, the users might lose interest in using multi-device interactions in the first place. As the users cannot see the wireless connections between the devices, they cannot be sure that they are really connecting to the other devices intended to. Therefore, the binding process should also provide sufficient cues and security, so that the users can ensure that the right devices are connected.

In this paper, we study establishing an ecosystem of mobile devices to support collaborative interactions within medium-sized groups of collocated users. While the problem of a single user pairing two devices has been extensively studied in prior research, more complex scenarios involving multiple users, especially more than four users, have received little attention in prior research. We present a touch-based group-binding method called FlexiGroups that builds on earlier research by Jokela and Lucero [13] and Lucero, et al. [16]. The method is highly flexible, enabling the users to apply a broad range of different group formation strategies. We also present a laboratory evaluation of FlexiGroups in a realistic photo sharing application context with four groups of six users. The evaluation results indicate that the method was generally found easy and intuitive to use. We analyze the different group creation strategies and patterns used by the

participants during the evaluation. The participants primarily followed viral patterns where they opportunistically added other participants to the group without advance planning. The participants also suggested a number of more systematic patterns, which required the group to agree on a common strategy but then provided a clear procedure to follow. The flexibility of the method allowed the participants to adapt it to the changing needs of the situation and to recover from errors and technical problems. Overall, device binding in medium-sized groups was found to be a highly collaborative group activity and the binding methods should pay special attention to supporting groupwork and social interactions.

The rest of this paper is structured as follows. First, we provide an overview of the related work. We then give a detailed description of the FlexiGroups binding method and the evaluation procedure. Finally, we present the results of the evaluation, followed by discussion and conclusions.

## RELATED WORK

The problem of ad hoc device binding has been thoroughly studied in the fields of human-computer interaction and security research. A wide range of methods for device binding has been proposed – in security research alone, over 20 different methods have been identified [20]. These methods vary in terms of device hardware requirements, amount of user involvement, and level of provided security.

The most common device-binding methods today, such as those typically used in Bluetooth and WLAN networks, are based on scanning the environment for available devices and presenting a list of the found devices to the user for selecting the other device to bind with. The connections are authenticated using short strings that the user is expected to copy or compare between devices. The authentication strings can be represented as numbers, words, graphical images, audio signals, or gestures in the user interface.

The proposed alternative methods include a variety of techniques based on synchronous user actions, for example, pressing buttons on both devices [22] or touching both devices [26] simultaneously, shaking the devices together [10], or bumping the devices together [8]. Further, device binding can be based on continuous gestures spanning from one device display to another [9]. Methods based on spatial alignment of the devices include pointing, for example, with laser light [18], touching [23], or placing the devices in close proximity of each other [15]. It is also possible to bind devices with various auxiliary devices such as tokens [1] or cameras [19]. Some of the proposed methods cover only device identification or authentication, while others combine both identification and authentication into a single user action.

Binding methods are not only means for connecting devices – they have strong social and emotional aspects. In real-life situations, people do not always pick the easiest or fastest method available, nor the one they like best. Many factors influence their choice of binding method, including the place, the social setting, the other people present, and the sensitivity of data [12, 21]. Users are willing to take security risks to comply with social norms [12].

The majority of earlier research has focused on scenarios of a single user binding two devices with each other (for example, binding a headset with a mobile device, or a mobile computer with a wireless access point). Only recently have researchers started to consider more complex scenarios involving multiple users and devices. Such multi-user scenarios differ in many respects from single-user scenarios, making the single-user device-binding methods not necessarily applicable to multi-user scenarios. In multi-user scenarios, communication between group members provides an additional source for potential errors. On the other hand, the users are typically willing to help each other and make decisions by mutual agreement, which reduces the amount of errors [14]. Methods that involve physical exchange of devices have been found to be unacceptable unless the users know each other very well, as the users are unwilling to hand in their devices to strangers [25, 5].

While numerous methods and technologies have been proposed for group association, Chong and Gellersen [2] present an interesting study on what people would spontaneously do to associate a group of devices. In their study, groups of four users were asked to suggest and rate techniques for binding together different combinations of low-fidelity acrylic prototypes of various mobile and fixed devices. Device touch based methods were found to be among the most frequently proposed methods, and were also considered popular and easy to use.

The group creation task can be divided in different ways between the members of the group [13, 14]. Leader-driven methods, which concentrate the task on a single participant, allow strong control over the group and require only one participant to be able to form a group. Peer-based methods, which distribute the work between all members, help to create a stronger sense of community and scale better to larger group sizes and distances. Further, group association can be seen as a one-step procedure of binding all devices with a single action, or as a sequence of pairwise associations [2, 24].

Finally, Chong and Gellersen [4] present a framework that sums up and categorizes the different factors that influence the usability of spontaneous device association. They identify technology, user interaction, and application context as the three most important criteria.

## STUDY

### Objectives

In this study, we were primarily interested in two research questions:

First, earlier studies have suggested that device-binding methods should be flexible, allowing people to adopt

different group creation strategies in different situations [13, 12]. Still, many of the methods tested in earlier studies have been very specific, enforcing a detailed procedure that has to be followed exactly. We wanted to test a more flexible method that would give users more freedom to adapt it to different situations. On the other hand, giving the users more options might potentially be confusing to them and provide additional possibilities for errors.

Second, we wanted to study group binding in a medium-sized group of six users. Earlier studies on device binding have focused either on individual users, pairs, or small groups of up to four participants. As the size of the group increases, a much wider variety of different approaches and strategies becomes possible. The overall process also becomes much more parallel. We wanted to better understand the different possible approaches and their strengths and weaknesses. We were also interested in group behavior, communication, and collaboration between users during the group-formation task.

### FlexiGroups

In order to study these research questions, we designed a group-binding method called FlexiGroups. The method builds on the results of an earlier comparison of three group-binding methods by Jokela and Lucero [13]. While the FlexiGroups method itself is generic and can be used in many different applications, we decided to study the group-binding method in the context of a multi-user photo sharing application to create a more realistic setting for the evaluation. The application was a simplified version of Pass-Them-Around [16]. It allowed the users to browse their own photo collections stored in their devices and also supported spatial interactions of sharing photos by throwing them from one device to another.

Figure 1 illustrates the FlexiGroups group-binding method. To create a new group, one of the persons first starts the FlexiGroups application on their device by tapping the application icon in the Application Grid (Fig. 1.a). The application starts in the Add Device view (Fig. 1.b). Visual feedback on the screen instructs the person to hold the device in portrait mode and to touch another device to add it to the group. When the person moves their device close to another device, the device detects the new device (see Fig. 2). Visual, auditory, and haptic feedback is provided to indicate that the other device has been detected and to instruct the person to hold the device still while the new device is added to the group. When the new device has been added to the group, the person can continue adding more devices by touching them following the same procedure. When the person does not want to add any more persons to the group, they should press the "Done" button to enter the Tabletop Overview view (Fig. 1.d).

Note that the person who is added to the group does not have to start the application manually from the Application Grid – their device can remain idle (Fig. 1.c). When another
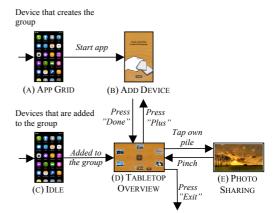


**Figure 1. FlexiGroups group-binding method.**

person touches the device and adds it to the group, the application is automatically launched and starts in the Tabletop Overview (Fig. 1.d). The Tabletop Overview (see Fig. 3) shows all the persons who have been added to the group. Each person is represented as a pile of photos with a textual name defined by the person next to the pile. If the person wants to add more persons to the group, they can press the "Plus" button at the center of the screen to enter the Add Device view (Fig. 1.b). The person can then add new devices by touching them in the same way as described in the previous paragraph. Any member of the group is allowed to add new devices and several persons can add new devices simultaneously.

FlexiGroups also supports defining the positions of the devices relative to each other in order to enable spatial interactions such as throwing photos between devices. This ordering phase is optional and can be omitted in applications that do not require the order of the devices to be defined. Alternatively, a similar mechanism can also be used to define other kinds of roles within the group, for example, to divide the group into two competing teams in a game application and to select captains for both teams. The ordering mechanism in FlexiGroups works as follows. The persons appear in the Tabletop Overview (see Fig. 3) in the order they are added to the group. If the order of the devices on the screen is different from the order of the devices in the real world, any member of the group can correct it by dragging the devices to the right positions on the screen. Only one person can change the order of the devices at a time. When one person is dragging a device to a new position, the other devices' screens are locked and grayed out to indicate that another person is reordering the devices.

When the order of the devices is correct on the screen, the people can enter the Photo Sharing view (Fig. 1.e) by tapping their piles of photos on the Tabletop Overview and start sharing photos by throwing them between devices. By pinching to zoom out in the Photo Sharing view, the person

**Figure 2. Touching another device to add it to the group.**



**Figure 3. Tabletop Overview.**

can return to the Tabletop Overview at any time to check the current members of the group, to add new members to the group, or to change the order of the devices. To leave the group, the person presses the "Exit" button in the Tabletop Overview. The group continues to run on the other devices until the last member exits the group.

## Prototype Implementation

We implemented a prototype of the FlexiGroups binding method on Nokia N9[1] mobile devices running the MeeGo operating system. The prototype was built as a native C++ application on top of the Qt 4.7 software framework. QML and Qt Quick with OpenGL ES hardware acceleration were used to implement a smooth animated user interface. Device touch interactions were detected with a radio technology, which was able to detect other devices at distances closer than 20 cm (8 inches) in approximately five seconds based on wireless signal strength. While the technology generally worked reliably, there were occasionally delays in detecting the other devices in the proximity and detections of devices further away. Detailed connectivity and initialization information was then sent to the discovered device over Bluetooth. A server, which was listening to a pre-defined Bluetooth socket on the discovered device, received the information and started the FlexiGroups application. The application then established an ad hoc WLAN network and connected to the group according to the connectivity information it had received. An ad hoc WLAN network was used for communication between the devices in order to allow the application to be used anywhere independent of the available network infrastructure. All communication was handled directly between the devices without a server backend. The prototype was fully functional with real network communication, except for the security protocols, which were only simulated in the user interface.

## Participants

We recruited a total of 24 participants for the evaluation by posting an advertisement on a local mailing list. The participants were recruited in groups of two to four people, as we wanted that each participant would know some of the other participants to make the situation more natural and comfortable to them. We assigned the participants into four evaluation sessions of six people each in the order they registered for the study, so that every participant knew at least one other participant, but no participant knew all the others in the same session. Six of the participants were female and 18 male. The ages of the participants varied between 25 and 41 years (M=32.8, SD=4.7). One participant was left-handed and 23 right-handed. The participants represented a variety of different professions, with six participants having a software engineering background, 13 having other technical background (for example, mechanical engineering), and five having a non-technical background (for example, teaching or photography). The participants were fairly advanced users of information technology – on a scale between 1-7 (1=novice, 7=expert), they evaluated their IT skills above average (M=5.5, SD=1.2). All participants were experienced smartphone users, but only two of them had used a Nokia N9 device before the study.

## Procedure

We organized a total of four evaluation sessions. The sessions were arranged in our usability laboratory. In each session, there were six participants and a moderator present. Figure 4 shows the evaluation setup. The participants were sitting around a round table with a radius of 120 cm (48 inches). Each participant was provided with a Nokia N9 mobile device with the FlexiGroups application pre-installed. While all the devices were of the same model, we used devices of different colors to make them easier to differentiate: there was one black, one white, two cyan, and two magenta devices. The average duration of the evaluation sessions was 60 minutes.

As the participants arrived in the laboratory, the moderator guided them to their seats around the table and asked them to fill in the background information and consent forms.

---

[1] http://swipe.nokia.com/

The moderator then introduced the participants to collaborative multi-device applications and demonstrated the idea with the photo sharing application. The participants could try the application with their own devices and practice throwing photos to each other. The moderator explained that before the participants could share photos like this, they first had to bind their devices together into a group, and that the purpose of the evaluation was to test a method for this task. The moderator then continued by describing the detailed evaluation procedure. The moderator also told the participants that the method to be tested was based on device touch interactions and demonstrated how to touch another device. The participants could practice touching with their own devices until they felt comfortable doing it. We considered practicing touch interactions necessary as while many participants were aware of device touch as an interaction method, few had practical experience using it.

To start the actual evaluation, the moderator played a video that demonstrated the FlexiGroups group-binding method. We used instructions recorded on video to ensure that all groups received uniform guidance on how to use the method. The instructions demonstrated the operations that an individual user could do – they did not show how a group of users should use the method together. The moderator then asked the participants to set up a group using the method that was just demonstrated to them, so that they could start sharing photos between devices. The moderator observed the situation and only intervened if the participants encountered obvious technical problems with their devices that prevented them from proceeding. When the participants had successfully created a group and could throw photos between devices, the moderator asked the participants to exit the application and create a new group so that a different person would start the group creation. Overall, the participants created three or four groups during each evaluation session.

When the participants had tested the method several times, the moderator asked them to fill in two validated questionnaires: AttrakDiff [7], which measures the attractiveness of interactive products, and NASA-TLX [6], which measures the subjective workload experience when performing a task. To gain a broader understanding of the tested method, we extended the NASA-TLX with four additional scales: Learnability, Quickness, Security, and Overall Preference. After the participants had completed the questionnaires, the moderator interviewed them about their experiences with the FlexiGroups method. The interview was semi-structured and covered general feedback about the strengths and weaknesses of the tested method as well as specific topics such as different group formation strategies, device touching orders, and arrangement of the devices into the correct positions. The average duration of the interviews was 20 minutes. To close the evaluation session, the moderator thanked the participants and gave



**Figure 4. Evaluation setup.**

each participant a small reward to compensate them for their time.

The group creation tasks were recorded with two video cameras: the first was placed on a tripod pointing towards the table at an angle (Fig. 4) and the other was mounted in the ceiling providing a top view of the table surface (Fig. 8 and Fig. 9). The interviews were recorded with a single video camera. Two researchers independently watched the video recordings and wrote notes about their observations. They also drew diagrams that recorded the sequences of participant actions in every group creation attempt. The same two researchers then collaboratively analyzed the data and built an Affinity Diagram [11] in a series of interpretation sessions. Each researcher independently studied the notes and grouped them into clusters of related items. The clusters then evolved to broader categories that were naturally revealed and were jointly revisited, discussed, and refined. In the end, the categories were processed into more general findings that form the core of the Results section. A quantitative analysis of the AttrakDiff and NASA-TLX questionnaires was done separately.

## RESULTS

### General
The evaluation produced overall positive results. All four evaluation groups succeeded in all of their attempts to bind their devices together and form a group. While the groups encountered some problems and made some mistakes, especially in their initial attempts to use FlexiGroups, the robustness of the method allowed them to recover and continue, and to successfully complete the group creation task. Most participants (18/24) commented that FlexiGroups was generally easy and intuitive to use and it was also easy to learn.

These qualitative results are supported by the AttrakDiff questionnaire [7] results, which are illustrated in Figure 5. The main bars indicate the means for each product dimension, while the error bars indicate standard errors. Pragmatic quality (PQ) refers to the product's ability to
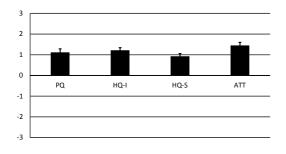
**Figure 5. AttrakDiff results (higher is better).**



**Figure 6. Extended NASA-TLX results (lower is better).**

support the achievement of behavioral goals (usability). Hedonic quality refers to the users' self: stimulation (HQ-S) is the product's ability to stimulate and enable personal growth, while identification (HQ-I) is the product's ability to address the need of expressing one's self through objects one owns. Perceived attractiveness (ATT) describes a global value of the product based on the quality perception. Both pragmatic and hedonic dimensions of AttrakDiff as well as overall attractiveness show positive and well-balanced results.
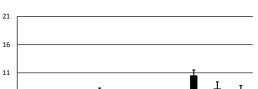
The extended NASA-TLX questionnaire [6] results are illustrated in Figure 6. The main bars indicate the means for each subscale, while the error bars indicate standard errors. The original six subscales of NASA-TLX presented on the left show an overall positive trend. The four subscales that we added for the purposes of this study are shown on the right. The low Learnability (LEA) score indicates that the method was considered very easy to learn. The method was also considered relatively secure (SEC). The Overall Preference (PRE) score further confirms the generally positive attitude towards the tested method.

However, several participants (7/24) commented that the group formation should have been faster. *"[P14] I think it is a bit slow process. It should be somehow faster. It takes a long time to set up."* The relatively high subjective Quickness (QUI) score of the extended NASA-TLX results supports these qualitative comments. We measured the fastest group creation time for each group using the video recordings. The average time to set up a six-person group was 111 seconds.

We can identify three main phases in forming a group: initiating the group creation, adding the participants to the group, and arranging the participants in the correct order for spatial interactions. We will next discuss each of these phases in more detail.

**Initiating the Group**

Every time the FlexiGroups application is started from the application grid, a new group is created. Therefore, when the participants wanted to create a new group, they had to agree who of them would start the application, while the
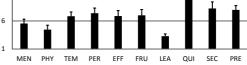
others had to wait until they were added to the group. Initially, this proved to be challenging for the participants as they were used to the common practice of each person first starting the application on their device. *"[P10] I found it confusing that by just starting the application, I started my own group."* In every evaluation group, several participants started the application in parallel during their first attempt to form a group. However, the visualization of the devices in the Tabletop Overview allowed the participants to quickly realize that there were several parallel groups. The participants solved the problem by agreeing that those people in the smaller groups should exit the application and be then added to the main group by the other participants. One of the participants saw it as an interesting opportunity that there could be several parallel subgroups within a larger group. *"[P21] If you are in a bar or restaurant and there are a lot of people in the table, you are not going to be talking to everyone. ... For me, it is quite natural that you have different groups."*

As the participants understood that only one of them should start the application, they quickly developed practices to agree verbally, or with gestures, who would start the application. These practices included announcing that one would start, asking for permission to start, and suggesting that another participant should start the application. However, half of the participants (12/24) proposed that it should be possible to start several groups in parallel and then merge the groups together. This would enable building a group from bottom up so that anybody could start and also would make the group creation faster. A few participants (4/24) commented that there should be a security mechanism to confirm that both groups really want to merge. Three participants also expressed more general concerns that somebody could add them to a group that they did not wish to join by touching their devices and suggested there should always be a confirmation before a device joins to a group.

**Touching Patterns**

FlexiGroups gave the participants a lot of freedom regarding the overall approach on how to form a group and the order in which the individual participants were added to

a) LEADER     b) CO-LEADERS     c) RING     d) TWO-WAY RING     e) VIRAL     f) VIRAL
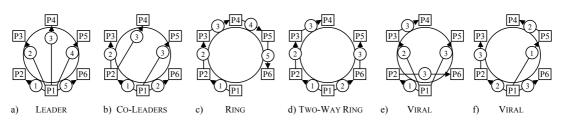
**Figure 7. Different group creation patterns. Boxes P1-P6 represent the participants. Arrows show the touch actions between the participants, while the numbered circles indicate the order of the touch actions.**

the group. A single participant could add all participants to the group, or several participants could participate in setting up the group. The participants could be added to the group one at a time, or multiple participants could be added in parallel. During the evaluation tasks and the interviews, the participants suggested a wide variety of different patterns for building a group. Figure 7 illustrates different patterns suggested by the participants.

### Systematic Patterns
In systematic patterns, the participants followed a well-defined procedure in setting up a group.

*Leader (Fig. 7.a).* The participant P1 (the leader) creates a new group and adds all the other participants one after another. The leader may proceed in clockwise or counter-clockwise order around the group.

*Co-Leaders (Fig. 7.b).* The participant P1 (the leader) creates a new group and then selects another participant P2 as a co-leader and adds the co-leader to the group. The two co-leaders P1 and P2 then add the other participants one after another proceeding in opposite directions around the group. While similar to the Leader pattern, having several co-leaders scales better to larger groups.

*Ring (Fig. 7.c).* The participant P1 creates a new group and adds another participant P2 next to them. The participant P2 then adds the next participant P3, who continues by adding P4. This way the ring proceeds around the group. The ring can be formed in clockwise or counter-clockwise order.

*Two-Way Ring (Fig. 7.d).* The participant P1 creates a new group and adds another participant P2 next to them. P1 and P2 then continue by adding the participants next to them like in Ring but in opposite directions. The Two-Way Ring pattern is similar to Ring, but more efficient as it proceeds in both directions around the group.

### Viral Pattern
While the participants suggested many systematic patterns, in practice their behavior was usually more random. *"[P14] Somebody initiated the group creation and then it started to spread around the table."* The participants opportunistically selected which device to touch next. *"[P3] It was very random. ... Going left and right. It didn't have any strict form."* This way the group membership spread like a viral infection, or fire, across the group from

one participant to another. Figures 7.e and 7.f show examples of real viral patterns employed by the participants during the evaluation.

It was common to connect to one's own neighbors first and then add other participants further away if needed. Participants who were not yet members of the group also requested the participants who were already in the group to add them to the group. If some participant had problems in detecting another device, the other participants were eager to help and add the new device with their own devices. Sometimes the participants intentionally touched and "infected" another participant on the opposite side of the table (for example, touch action #1 in Fig. 7.f) to make the group membership spread faster. It was common that several participants were touching and adding new devices simultaneously (see Fig. 8).

### Alternative Patterns
In addition to the main patterns described above, individual participants suggested a range of alternative approaches for group creation. Interesting alternatives include patterns where several devices were touched simultaneously to make the group creation more efficient. In one variation, the participant who creates the group puts their device at the center of the table. The other participants who want to be added to the group then put their devices next to it. In another variation, all devices are collected next to each other on the table and the participant who creates the group then touches all of them in one action. While these patterns



**Figure 8. Three persons touching and adding devices in parallel.**

would have been possible with the tested prototype, they were not used in practice by any group during the evaluation.

### Device Ordering

While the phase of touching and adding devices to the group was characterized by highly parallel activity with several participants adding other people simultaneously (see Fig. 8), acting either independently or in small sub-teams, the phase of device ordering required the participants to more closely co-operate and coordinate their actions within the whole group. The transition from the adding phase to the ordering phase provided an important synchronization point in the group formation task.

Device ordering was done collaboratively with different participants taking different roles (see Fig. 9). While some participants actually moved devices on the screen, others checked whether the order was correct and provided suggestions on the needed changes verbally and with hand gestures. In most cases, several participants moved devices on the screen – typically two to four participants were involved. This created problems as several participants tried to move the devices simultaneously. While the screens of the other devices were grayed out and locked when one of the participants started to move a device around, there was a small delay in locking the screens, which sometimes allowed another participant to start moving until their screen was grayed out. Many participants (13/24) complained that they felt the device ordering phase was confusing. Several participants (7/24) also commented that the ordering phased was unnecessary and should be avoided. The system should have been able to define the device positions automatically instead.

### Physical Device Handling

Half of the participants (12/24) made a natural and spontaneous gesture of pushing their devices forward when they wanted to be added to the group. Most commonly, a participant would move their device towards another participant asking them to touch the device and add it to the group. Alternatively, a participant could push their device towards the center of the table asking any of the other participants to add it to the group.

While the participants commonly touched each other's devices with their own devices to add them to the group, there was a high barrier of taking another participant's device into one's hands or manipulating it with one's fingers. This was true even if the participants did not use their own personal devices but devices we had given to them for the purpose of the evaluation. In all the sessions, there were only a few cases where a participant touched another participant's device with their hands. In those few cases, the reason was usually to help another participant who had technical problems with their device.



**Figure 9. Ordering devices.**

We encouraged the participants to keep their devices on the table as that would allow everybody to see the screens of the other participants and help create a common awareness of the group status. Keeping the devices on the table also made the proximity detection technology work faster and more reliably. Still, the participants often chose to hold their devices in their hands. In 27% of the touch actions recorded on video, both participants kept the devices in their hands, making the touch action resemble a handshake (see Fig. 8). In one of the sessions, all participants held the devices in their hands also during the reordering phase. In the other sessions, two participants liked to fiddle or toy with their devices while waiting to be added to the group.

### Collaboration

We observed a high level of communication and collaboration between the participants when they were forming a group. Participants were very eager to help each other if some participant encountered problems with the application. For example, if a participant's device could not detect and add another device to the group, another participant would use their device to help and add the participant. Helping others mostly occurred spontaneously when a participant noticed that another participant could not complete some action or was doing something incorrectly. Only rarely did the participants explicitly ask for help.

A major challenge for the evaluation groups was creating and maintaining a common understanding of the overall task status as there were six persons involved and many actions were taking place in parallel. *"[P23] I cannot keep an eye on what every other person is doing."* Several techniques were used to accomplish common awareness of the task status. Verbal communication and coordination between the participants played a major role, including the participants announcing intentions to do some actions, providing feedback on other participants' actions, instructing others to take some actions, asking and confirming facts, and stating the common group status aloud. Another important technique was to observe the other participants, including both real-world actions taken

by the others as well as the status information shown on the their device screens. When adding participants to the group, everybody could easily perceive the touch actions. The devices also provided audio feedback when another device was detected and added to the group. While the audio feedback was useful and could also be observed by all the participants, the participants had problems in identifying the source of the audio when several participants were touching devices in parallel. In the ordering phase, the participants had to rely more on information on the device screens, and peeking at and comparing information between other participants' screens was common. Figure 9 shows an example of a situation where all participants have placed their devices in a close formation at the center of the table to make the coordination easier.

**DISCUSSION**

As observed in our study and in earlier studies [13, 14], device association in large groups is a highly collaborative group activity. Considering the design of binding methods for groups, while good usability is definitely important, people as a group can help each other and are capable of overcoming and solving together most usability and technical problems they encounter. However, in larger groups, the main challenges are related to groupwork and social interactions within the group: making decisions and agreeing on a common strategy, coordinating and synchronizing actions, and keeping track of the others and the overall task status. An important consideration is also keeping everybody engaged in the process, as people easily get bored or distracted when they cannot do anything but wait for others to complete the group formation.

In the evaluation sessions, people most commonly and naturally employed Viral patterns (see Fig. 7.e and 7.f) where the group membership spread like a contagion from one person to another [24]. These required the least advance planning within the group and provided most flexibility. The Viral patterns are efficient and keep everybody involved also in large groups due to the high level of parallel activity. The possibility of forming a group bottom-up by first forming smaller groups and then merging them together into a larger group, which was suggested by some of the study participants, is interesting and should be studied further.

The participants also suggested a variety of more systematic patterns. These patterns require that the participants are aware of and agree on a common strategy to form a group. While this requires some initial planning, the systematic pattern then defines a precise sequence of actions that each participant should follow, reducing the need for coordination between the participants later. Further, the Leader pattern (see Fig. 7.a) enables the leader to have strong control over the group [13], which may be important, for example, in situations where there are unknown people present. The Co-Leaders pattern (see Fig. 7.b) provides an

extension of the Leader pattern, which scales better to larger groups.

While the participants were able to successfully arrange the devices in the right order for spatial interactions in all group creation attempts, the majority of them found the ordering phase confusing. While this can be partially attributed to implementation issues, such as delays in locking the screens of the other devices when one device was used for rearranging, the ordering technique needs to be improved. Ideally, device ordering should be automatic, removing the need for manual ordering completely. However, in real-life applications this may be difficult to achieve as it may require special tracking equipment or dedicated infrastructure that may not be widely available. As suggested in an earlier study by Jokela and Lucero [13], device ordering would probably work better when done by one person. This person could be, for example, the person who creates the group. Alternatively, one of the participants could reserve the role for a longer period of time and arrange all the people in the right places. If strictly followed, the systematic patterns might enable defining the device order based on the touch order, but this may be in many cases too restrictive [13].

When designing binding methods for groups of users, it is important to consider robustness in real-life conditions. While many methods can work well in theory or with mock-ups, in reality, multi-user multi-device applications are complex distributed systems. As multiple devices are involved, there is an increased risk of technical issues: the devices may fail to detect each other, the software may crash, and the network connections may be broken. Also, all persons may not be aware of the procedure they should follow, or they may be unable to do so, for example, because they arrive late or they are occupied with other tasks such as incoming telephone calls. Therefore, the methods should not expect an exact procedure to be followed. The methods should be flexible and robust, allowing people to adapt them to the changing needs of the situation and to recover from failures.

**FUTURE WORK**

Our experiment, like all the other experiments with group-binding methods we are aware of, was done in a usability laboratory under ideal conditions. Also, the use case was defined by the researchers and given to the participants. In real life, various contextual and situational factors influence the group creation process. Therefore, to gain a deeper understanding of group binding in realistic situations and tasks, we believe it would be important to study group-binding methods also in more realistic settings and over extended periods of time with longitudinal field trials.

**CONCLUSION**

We have presented FlexiGroups, a touch-based method to bind mobile devices for collaborative interactions within a group of collocated users. The method is highly flexible,

enabling a broad range of different group formation strategies. In a laboratory evaluation with four groups of six users, the method was found to be generally intuitive and easy to use and learn. When forming a group, the participants primarily followed viral patterns where they opportunistically added other participants to the group without advance planning. The participants also suggested a number of more systematic patterns, which required the group to agree on a common strategy but then provided a clear procedure to follow. The flexibility of the method allowed the participants to adapt it to the changing needs of the situation and to recover from errors and technical problems. Overall, device binding in medium-sized groups was found to be a highly collaborative group activity and the binding methods should pay special attention to supporting groupwork and social interactions.

**REFERENCES**
1. Ayatsuka, Y. and Rekimoto, J. tranSticks: physically manipulatable virtual connections. In *Proc. CHI '05*, 251-260.

2. Chong, M. and Gellersen, H. How groups of users associate wireless devices. In *Proc. CHI '13*, 1559-1568.

3. Chong, M. and Gellersen, H. How users associate wireless devices. In *Proc. CHI '11*, 1909-1918.

4. Chong, M. and Gellersen, H. Usability classification for spontaneous device association. *Personal and Ubiquitous Computing* 16, 1 (2012), 77-89.

5. Hang, A., von Zezschwitz, E., De Luca, A. and Hussman, H. Too much information!: user attitudes towards smartphone sharing. In *Proc. NordiCHI '12,* 284-287.

6. Hart, S. G. and Staveland, L. E. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In Hancock and Meshkati (eds.) *Human Mental Workload*, North Holland Press, 1988.

7. Hassenzahl, M. The Interplay of Beauty, Goodness, and Usability in Interactive Products. *Human-Computer Interaction* 19, 4 (2004), 319-349.

8. Hinckley, K. Synchronous gestures for multiple persons and computers. In *Proc. UIST '03*, 149-158.

9. Hinckley, K., Ramos, G., Guimbretiere, F., Baudisch, P. and Smith, M. Stitching: pen gestures that span multiple displays. In *Proc. AVI '04*, 23-31.

10. Holmquist, L.E., Mattern, F., Schiele, B., Alahuhta, P., Beigl, M. and Gellersen, H. Smart-Its friends: a technique for users to easily establish connections between smart artefacts. In *Proc. UbiComp '01*, 116-122.

11. Holtzblatt, K., Wendell, J. B. and Wood, S. *Rapid Contextual Design*. Morgan Kaufmann, 2004.

12. Ion, I., Langheinrich, M., Kumaraguru, P. and Čapkun, S. Influence of user perception, security needs, and social factors on device pairing method choices. In *Proc. SOUPS '10*.

13. Jokela, T. and Lucero, A. A comparative evaluation of touch-based methods to bind mobile devices for collaborative interactions. In *Proc. CHI '13*, 3355-3364.

14. Kainda, R., Flechais, I. and Roscoe, A. Two heads are better than one: security and usability of device associations in group scenarios. In *Proc. SOUPS '10*.

15. Kray, C., Rohs, M., Hook, J. and Kratz, S. Group coordination and negotiation through spatial proximity regions around mobile devices on augmented tabletops. In *Proc. TABLETOP 2008*, 1-8.

16. Lucero, A., Holopainen, J. and Jokela, T. Pass-Them-Around: collaborative use of mobile phones for photo sharing. In *Proc. CHI '11*, 1787-1796.

17. Lucero, A., Jokela, T., Palin, A., Aaltonen, V. and Nikara, J. EasyGroups: binding mobile devices for collaborative interactions. In *CHI EA '12*, 2189-2194.

18. Mayrhofer, R. and Welch, M. A human-verifiable authentication protocol using visible laser light. In *Proc. ARES '07*, 1143-1148.

19. McCune, J.M., Perrig, A. and Reiter, M.K. Seeing-is-believing: using camera phones for human-verifiable authentication. In *Proc. SOUPS '05*, 110-124.

20. Nithyanand, R., Saxena, N., Tsudik, G. and Uzun, E. Groupthink: usability of secure group association for wireless devices. In *Proc. Ubicomp '10*, 331-340.

21. Rashid, U. and Quigley, A. Interaction techniques for binding smartphones: a desirability evaluation. In *Proc. HCD '09*, 120-128.

22. Rekimoto, J. SyncTap: synchronous user operation for spontaneous network connection. *Personal and Ubiquitous Computing* 8, 2 (2004), 126-134.

23. Rekimoto, J., Ayatsuka, Y., Kohno, M. and Oba, H. Proximal interactions: a direct manipulation technique for wireless networking. In *Proc. INTERACT '03*, 511-518.

24. Terrenghi, L., Quigley, A. and Dix, A. A taxonomy for and analysis of multi-person-display ecosystems. *Personal and Ubiquitous Computing* 13, 8 (2009), 583-598.

25. Uzun, E., Saxena, N. and Kumar, A. Pairing devices for social interactions: a comparative usability evaluation. In *Proc. CHI '11*, 2315-2324.

26. Zimmerman, T. G. Personal area networks: near-field intrabody communication. *IBM Systems Journal* 35, 3-4 (1996), 609-617.

**Publication P4**

Tero Jokela, Jarno Ojala, Guido Grassel, Petri Piippo, and Thomas Olsson. 2015. A Comparison of Methods to Move Visual Objects Between Personal Mobile Devices in Different Contexts of Use. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services (Mobile-HCI '15)*. ACM, New York, NY, USA, 172–181. DOI: https://doi.org/10.1145/2785830.2785841

# A Comparison of Methods to Move Visual Objects Between Personal Mobile Devices in Different Contexts of Use

**Tero Jokela[1], Jarno Ojala[3], Guido Grassel[2], Petri Piippo[1], Thomas Olsson[3]**
[1] Nokia Technologies, P.O. Box 1000, FI-33721 Tampere, Finland
[2] Nokia Technologies, P.O. Box 226, FI-00045 Espoo, Finland
[3] Tampere University of Technology, P.O. Box 553, FI-33101 Tampere, Finland
{tero.jokela, guido.grassel, petri.piippo}@nokia.com, {jarno.ojala, thomas.olsson}@tut.fi

## ABSTRACT

As people increasingly own multiple mobile and portable devices (such as smartphones, tablets, and laptops), situations where several devices are used together have become more common. A frequent problem in such situations is how to move virtual visual objects (such as content items or application windows) between device displays. We present a comparative evaluation of three methods for moving objects between personal mobile devices: Tray, Transfer Mode, and Device Touch. The participants' preferences of the methods in different real-life scenarios were found to strongly depend on the task and the context of use, making the design of a single optimal cross-display object movement method a challenging task. We identify several clusters of contextual factors that influenced the users' preferences. We also report more detailed differences in efficiency, novelty, learnability, physical device handling, and task completion strategies between the three methods included in the evaluation.

## Author Keywords

Multi-device interaction; cross-display object movement; mobile devices; smartphones; tablets; laptops.

## ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

## INTRODUCTION

As the number of mobile and portable devices that people have in their possession has steadily increased, situations and tasks where several devices are combined and used together have become common. A frequent problem in such situations is how to move virtual visual objects from the display of one device to the display of another device. For example, a user might take a photo with their camera phone

and want to immediately attach it to a document that they are writing on their laptop. As another example, a user might be watching a video on a family tablet, but when another family member would like to play games with the tablet, the user might want to move the video player to their phone and continue watching the video with it.

In this paper, we present a comparative evaluation of three different methods for moving visual objects between device displays: (1) Tray, which enables the user to move objects through a virtual tray shared between devices; (2) Transfer Mode, which requires the user to first connect the devices but then allows moving objects by simply tapping them; and (3) Device Touch, which supports moving objects by using one device to physically touch the other device. While the earlier studies comparing different cross-display object movement methods have primarily addressed moving visual objects between handheld devices and large tabletop or wall displays, we focus on situations where a person wants to move visual objects between the displays of their personal mobile and portable devices (such as smartphones, tablets, and laptops) located within hand's reach. Further, while the earlier comparisons have involved a single (often abstract) task in a laboratory environment, we explore the participants' preferences of different methods more broadly by also presenting them with four real-life scenarios involving cross-display object movement and asking them to evaluate the suitability of the different methods for each scenario. Importantly, the participants were found to prefer different methods in different scenarios, indicating that the preferences of the methods were strongly influenced by several factors related to the task and the context of use. This makes the design of a single optimal cross-display object movement method a difficult task. We also report more detailed observations on differences in efficiency, novelty, learnability, physical device handling, and task completion strategies between the three methods. The results inform the design of future multi-device systems that involve moving virtual visual objects between device displays.

## RELATED WORK

Studies on computing with multiple devices have found that accessing and managing content across devices is one of the key concerns in multi-device environments [10, 18, 16]. Conventional solutions for moving content items between

devices include connecting them directly using physical cables or wireless short-range radio technologies, such as Bluetooth or Wi-Fi. Alternatively, objects can be moved through portable storage media such as memory cards, USB memory sticks, and external hard drives. It is also a common practice to move objects between devices by sending them as e-mails or other messages to oneself. Further, objects can be saved in network folders or storage services and then re-opened with other devices. Many cloud services support transferring specific content types (for example, browser tabs, photographs, video, or music) between devices. Despite a broad range of commercially available solutions, the users still commonly encounter problems in accessing content across devices [10, 18].

In the field of HCI research, the general problem of cross-display object movement has been studied in the context of multi-monitor computer setups, large composite displays, collaborative interactive spaces, tabletop displays, and ad-hoc ecosystems of mobile devices. A wide variety of different techniques for moving visual objects between displays has been proposed in prior literature. For example, in Pick-and-Drop [17], the user can pick up an object by touching it with a digital pen and then drop it by repeating the touch action on another display. Conduit [4] uses the metaphor of transferring the object through the user's body with the user first designating the target display by touching it with the non-dominant hand and then selecting the object to be transferred with the dominant hand on the source display. Objects can also be moved from one display to another by making a throw touch screen gesture towards the target [11]. Many variations of the basic throw gesture, for example, Drag-and-Throw and Push-and-Throw [8], have been suggested, aiming to provide better control over and to improve the accuracy of the throw.

Drag-and-Pop [2] extends traditional drag-and-drop to work across multiple displays by bringing the possible targets from other displays near the object when the user starts to drag it, allowing the user to complete the movement action on the same display. In ConnecTables [22], the displays of two devices positioned side-by-side form a single continuous display area and objects can be directly dragged from one display to another. Objects can also be transferred between device displays by stacking the devices on top of each other [5]. In ARIS [3], the user can manage objects across screens on a radar-like map view showing all displays in proximity. Further, objects can be manipulated across displays by using pointing gestures, for example, with laser pointers [15]. Synchronized clipboards [12] are shared between multiple devices, so that an object copied to the clipboard on one device can be pasted on another device. In Conductor [6], the user may broadcast objects to all other devices as cues which can then be acted upon or ignored. In MediaBlocks [23], objects can be bound to physical tokens that can be moved between displays.

Nacenta, et al. [14] present a taxonomy and classification of cross-display object movement methods based on three conceptual levels. On referential domain level, the methods can reference displays either spatially or non-spatially (for example, using textual names). On display configuration level, the mapping between spatial arrangement of the displays and the method's input model can be planar, perspective, or literal. On control paradigm level, methods can be open-loop, closed-loop, or intermittent.

Despite the large variety of proposed methods, only a few studies have compared different methods against each other. These studies have mostly focused on moving objects between handheld devices and large tabletop or wall displays. Nacenta, et al. [13] report a study that evaluated the performance of six different methods for moving objects with a stylus from a tablet computer to a tabletop display at different distances within and beyond hand's reach. Substantial differences between the methods were identified. Radar View and Pick-and-Drop were found to be the fastest methods and were also preferred by the participants.

Bachl, et al. [1] present a study where pairs of users solved collaborative tasks which involved moving objects from a shared tabletop display to personal tablet computers. Three different one-way object movement methods were included in the study: pressing a button attached to the object, dragging the object on top of an icon representing the target device, and using the tablet as a lens to capture content beneath it on the shared display. The participants considered the slowest button method to be the easiest to use, while the fastest lenses method was considered too complicated.

Scott, et al. [21] report another study where groups of users played a game that involved moving virtual cards between tablet computers and a large tabletop display. Three different cross-display object movement methods were included in the study: two variants of the Pick-and-Drop method and a method called Bridges where objects were moved through a virtual portal area shared between devices. All methods were found to effectively support the movement tasks. The participants' preferences were equally divided between the methods, with each method providing unique advantages that suited different participants' playing styles.

## OUR STUDY

### Objectives

Our study on cross-display object movement methods was motivated by two primary objectives:

First, while the earlier studies comparing different cross-display object movement methods have primarily focused on moving visual objects between handheld devices and large tabletop or wall displays, we were interested in situations where a single user wants to move visual objects

**Figure 1. Different cross-display object movement methods: a) Tray, b) Transfer Mode, and c) Device Touch.**

between two or more mobile or portable devices (such as smartphones, tablets, and laptops). We assumed that all devices would have touch screens and that they would be located within hand's reach. As people possess increasing numbers of mobile devices, such computing situations are common in today's mobile multi-device environments.

Second, while the earlier studies comparing different cross-display object movement methods have concentrated on a single (often abstract) task in a laboratory environment, we wanted to explore the users' preferences of cross-display object movement methods over a range of different situations, tasks, and environments. We were also interested in the different factors that influence the users' preferences and selections of the methods in different situations. Overall, we wanted to consider not only the pragmatic aspects (such as efficiency and number of errors) but also the hedonic aspects (such as emotional and social factors) of the different methods.

**Cross-Display Object Movement Methods**
One of the key challenges in designing a comparative study of cross-display object movement methods is deciding which methods to include in the study. While a wide variety of different methods have been proposed, the study participants may easily become overwhelmed if the number of methods to learn and evaluate is too large. Therefore, following an approach similar to those adopted by Bachl, et al. [1] and Scott, et al. [21], we decided to restrict the number of methods to three to limit the mental load of the participants. After experimenting with different methods, we finally selected three methods, (1) Tray, (2) Transfer Mode, and (3) Device Touch, that featured significantly different approaches to the cross-display object movement problem and appeared to perform well in our initial testing. In the Tray method, the user could move objects through a virtual tray shared between all devices. It represented the conventional model of using an intermediate storage, such as a shared clipboard or a network folder, for moving objects between displays. The Transfer Mode method was based on the concept of virtual connections between devices. The user first created a connection by making a screen gesture and could then share objects between the

connected devices by simply tapping objects on the displays. In the Device Touch method, the user could move objects by physically touching the screen of another device with the corner of their phone. Instead of virtual interactions on the display, it relied on tangible interactions between physical devices in the real world. We will describe the methods in detail in the following subsections.

While all three methods were generic and could be used to move a wide range of different kinds of objects in different applications, we decided to prototype them in the context of a simple multi-device photo browsing and management application. The application allowed the user to browse the local photo collection stored in the device in a grid view. It was also possible to rearrange photos by dragging them to a new position in the grid. By tapping a photo, it could be opened in a full screen view for more detailed viewing. In the full screen view, the user could browse the photos by swiping left or right. It was also possible to delete photos. Finally, the photo browser allowed the user to move photos from one device to another using the three cross-display object movement methods, for example, in order to sort and organize them between devices.

*Tray*
In the Tray method (see Figure 1.a), the photos were transferred through a virtual tray that was shared between all devices. Conceptually, the Tray was similar to a shared visual clipboard or a temporary network folder, or the Bridges virtual portal method used by Scott, et al. [21]. The user could open the tray by making a swipe gesture. The user could then move photos to the tray by dragging. When the user opened the tray on another device, they could see the photos that were moved there on the first device. The user could then drag any of the photos from the tray to a desired position in the grid on the screen of the second device.

*Transfer Mode*
In the Transfer Mode method (see Figure 1.b), the user first indicated which devices were involved in photo transfer by making a swipe gesture on each device. Each device activated the transfer mode which was indicated by blue

Figure 2. Device Touch implementation.



Figure 3. Evaluation setup.

frames around the display. The user could then select photos on any device by tapping them on the screen. The selected photos appeared on the other devices as scaled-down preview thumbnails in the grid, similarly to the broadcasted cues in Conductor [6]. The user could confirm the transfer by tapping the thumbnails on those devices where they wanted to move the photos. The user could then deactivate the transfer mode by making the swipe gesture again. When the transfer mode was deactivated, the thumbnails of unconfirmed photos disappeared on the other devices.

*Device Touch*

In the Device Touch method (see Figure 1.c), the user could move photos by touching the display of a tablet (or a laptop) with the corner of a phone. The Device Touch method was inspired by the earlier work of Schmidt, et al. [19, 20]. To move a photo from a phone to a tablet, the user first selected a photo to be transferred by making a long press on the screen of the phone. The photo then slid to the top of the screen and the user could transfer the photo to the tablet by touching the screen of the tablet with the corner of the phone. A preview of the photo appeared on the screen of the tablet and the user could drag the photo to the desired position in the grid by moving the phone like a stylus. Lifting the phone up from the screen of the tablet completed the transfer. To move a photo from a tablet to a phone, the user picked the photo from the screen of the tablet by touching it with the corner of the phone. The photo then appeared at the top of the screen of the phone and the user could drag it down to the desired position on the grid. Compared to conventional file transfer with NFC, the Device Touch method did not require the user to touch a specific point on the tablet surface but the user could touch anywhere on the tablet screen. The tablet was able to detect the exact point that the user was touching, allowing them to put a photo on a specific position or to pick a specific photo from the tablet screen.

**Prototype Implementation**

We built a prototype of the three cross-display object movement methods on Windows 8 and Windows Phone 8

platforms. The prototype was implemented as a hybrid web and native application. The user interface was implemented in HTML, CSS, and JavaScript, while the lower level functions were implemented in native C#. The prototype implemented all the three object movement methods but only one of them was active at a time. It was possible to select the active method through a hidden menu. The devices were connected over a Wi-Fi network. A simple star topology was used for communication between devices with one of the devices acting as a central hub.

The implementation of Device Touch followed the general description given by Schmidt, et al. [19, 20]. However, instead of camera-based detection used by Schmidt, et al., we used a novel implementation approach of applying capacitive sensing technology to detect the device touch events. This implementation had the important benefit that the method could be used with standard commercially available phones, tablets, and laptops with capacitive touch screens. We attached stylus heads to the corners of the phones (see Figure 2). Conductive tape was used on the device sides to connect the user's body to the stylus. Device internal sensors were used for separating finger touch events from device touch events and for identifying the touching device.

**Participants**

We recruited a total of 18 participants for the evaluation by posting an advertisement on a local mailing list. Nine of the participants were female and nine male. The ages of the participants varied between 21 and 52 years (M=30.4, SD=8.1). Six of the participants were left-handed and twelve right-handed. The participants represented a variety of different backgrounds, with ten participants having a background related to information technology and eight not related to information technology (for example, a teacher, a civil engineer, and a nurse). The participants were fairly advanced users of information technology: on a scale between 1 and 7 (1=novice, 7=expert), the participants rated their familiarity with information technology above average (M=5.1, SD=1.5). All participants were experienced smartphone users and fourteen of them were also active tablet users.

**Procedure**

The evaluation sessions were organized in a usability laboratory. In each session, there was a participant and a researcher who served as a moderator present. In some sessions, another researcher participated as an observer. Figure 3 shows the evaluation setup.

The prototype system used in the evaluation consisted of three devices: a Sony Vaio Windows 8 Pro tablet and two Nokia Windows Phone smartphones, a Lumia 625 and a Lumia 620. The devices were carefully selected to represent different device categories used today: a tablet with an 11.6-inch screen, a smartphone with a large 4.7-inch screen, and a smartphone with a small 3.8-inch screen. Two different sets of phones were used: one set of phones with extra styluses in the corners for the Device Touch method and another set of unmodified phones for the Tray and Transfer Mode methods. The photo browser prototype software, which implemented the three object movement methods, was installed on all devices.

As the participant arrived in the laboratory, the moderator first asked them to fill in the background information and consent forms. The moderator then introduced the participant to the idea of cross-display object movement and gave an overview of the study. The moderator also demonstrated the basic operations of the photo browser application and allowed the participant to practice them. While photos were used as example content in the prototype, the moderator emphasized that the methods could be used for moving any kind of objects.

To begin the actual evaluation, the moderator explained and showed how to use the first object movement method. The participant was then asked to practice the method by transferring single photos between each phone and the tablet in both directions. When the participant felt comfortable with the method, the moderator gave the participant two different tasks in which the participant was asked to move groups of photos between the phones and the tablet, again in both directions. The tasks also required the participant to arrange the photos in a specified order in the grid. A number was printed on each photo to define a clear order for them. The tasks were as follows:

*Task T1.* Initially, the tablet was empty, while the phones contained different selections of photos of different animals. The participant was asked to copy all photos of chicken to the tablet, so that they were in the numbered order. There was a total of eight photos of chicken, four on each phone.

*Task T2.* Initially, the phones were empty, while the tablet contained a selection of photos of different animals. The participant was asked to copy all photos of chicken to the small phone and all photos of sheep to the large phone, so that they were in the numbered order. There were four photos of chicken and four photos of sheep.

As the participant had completed both tasks, the moderator asked the participant to fill in an Extended NASA-TLX questionnaire [7], which measures the subjective workload experience when performing a task. To gain a broader picture of the overall user experience, the moderator also asked the participant to fill in an AttrakDiff questionnaire [9], which measures the attractiveness of an interactive product. The moderator then briefly interviewed the participant about their immediate experiences with the method and ideas about how the method could be improved. The same procedure was then repeated for the second and the third methods. The order of the methods was systematically varied between evaluation sessions to counter-balance any learning effects.

After the participant had experimented with all the three methods, the evaluation session was concluded with a summative interview contrasting the methods with each other. In the interview, the participant was presented four different Scenarios S1-S4 involving cross-display object movement and was asked to rate the applicability of each method to each scenario using a scale similar to NASA-TLX. Each scenario was accompanied with a photograph illustrating the context of use. The scenarios were based on a recent user study on practices of multi-device use [10] and represented real-life situations where people today use multiple devices. The scenarios were carefully crafted to balance different content types, device combinations, directions of object movement, and environments.

*Scenario S1: Photos.* When traveling in a train, I wrote an e-mail message to my friend with a laptop. I selected five holiday pictures from my phone and attached them to the message.

*Scenario S2: Phone Number.* At the office, I searched the client's phone number with my laptop. I copied the number to my phone and called the client with it.

*Scenario S3: Game.* I was sitting on the couch in the living room at home and was playing a game with my phone. When the phone ran out of battery, I moved the game to a tablet and continued playing with it.

*Scenario S4: Browser Tabs.* In a café, I browsed the web with a tablet and found three interesting pages. When I had to leave, I moved the browser tabs to my phone, so that I could continue reading the pages on the bus on my way home.

The evaluation tasks and the interviews were recorded with a video camera. The video material was transcribed and thematically analyzed by two researchers. A quantitative analysis of the Extended NASA-TLX and AttrakDiff questionnaires was done separately.

**RESULTS**

We first give an overview of the quantitative results. We then present the qualitative results in detail and contrast them with the quantitative results when relevant.
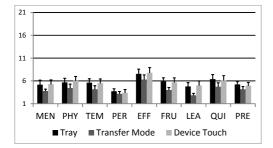
**Figure 4. Extended NASA-TLX results (lower value is better).**



**Figure 5. AttrakDiff results (higher value is better).**

### General

The overall results were positive for all the cross-display object movement methods included in the evaluation. The participants commented that in general, the methods were easy, simple, and intuitive to use. These positive comments were supported by the questionnaire results.

Figure 4 illustrates the results of the Extended NASA-TLX questionnaire [7]. The main bars indicate the means for each subscale, while the error bars indicate standard errors. The original six subscales of NASA-TLX are presented on the left. To gain a broader picture of the evaluated methods, we extended the questionnaire with three additional scales, Learnability (LEA), Quickness (QUI), and Overall Preference (PRE), which are shown on the right. All methods received good scores on all scales, indicating that the participants were able to complete the evaluation tasks with any method reasonably well. Transfer Mode consistently received the best score on each scale, suggesting that it performed the best for the photo sorting tasks used in the evaluation. However, the differences between the methods were small. One-way repeated measures ANOVA tests were conducted to compare the results for statistically significant differences. The results indicate that the object movement method had significant effect on Learnability ($F(2, 16) = 6.34$, $p = .009$). Pairwise comparisons show that the level of Learnability was significantly lower for Transfer Mode compared to Device Touch ($p = .04$).

Figure 5 shows the results of the AttrakDiff questionnaire [9] for the three object movement methods. Pragmatic Quality (PQ) refers to the product's ability to support the achievement of behavioral goals (usability) and the results are consistent with the extended NASA-TLX with Transfer Mode receiving the best score. Hedonic Quality refers to the users' self: Identification (HQ-I) is the product's ability to address the need of expressing one's self through objects one owns, while Stimulation (HQ-S) is the product's ability to stimulate and enable personal growth. There are only minor differences in the Identification scores, indicating that the participants did not see major distinctions in social aspects related to the use of the three methods. Device Touch received a higher Stimulation score than the other
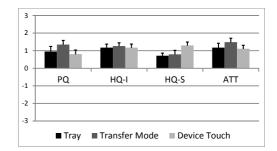
methods, meaning that it was considered the most novel and exciting of the three methods. Perceived Attractiveness (ATT) describes a global value of the product based on the quality perception and shows results similar to Overall Preference in the Extended NASA-TLX, with Transfer Mode receiving a slightly better score than the other two methods. One-way repeated measures ANOVA tests found no statistically significant differences between the methods.

### Performance

We measured the completion times for both tasks with all the three methods from the video recordings. Figure 6 illustrates the mean completion times, with the error bars showing the standard errors. On the average, Transfer Mode was the fastest method for the Task T1. Otherwise, the differences between the methods were small. One-way repeated measures ANOVA tests indicated no statistically significant differences between the methods.

Regarding the perceived performance, most of the participants (13/18) commented that Transfer Mode was fast and efficient. *"[P3] When you have activated the Transfer Mode, all you need to do is to click the photos."* The Extended NASA-TLX also suggests similar results, with the Transfer Mode receiving slightly better perceived MEN, PHY, TEM, and QUI scores than the other methods.

Many participants (7/18) considered Tray inefficient because they had to first drag the object to the tray on the first device and then drag the object from the tray to the final position on the second device, effectively duplicating the number of necessary movement actions. *"[P8] It is not the tray that is the destination. I want to move the objects to the other device."* Similarly, several participants (8/18) felt that Device Touch was slow and laborious as they had to physically touch the other device for every object they wanted to transfer. Two participants suggested that the user could select several objects and then transfer them all with a single touch action. On the other hand, many participants (8/18) appreciated the immediateness of Device Touch, as it worked instantly without any preparations or setup. *"[P11] You just touch the photo and it is directly in your phone."*
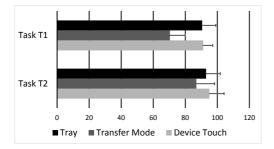
**Figure 6. Task completion times (in seconds).**



**Figure 7. Different grips for Device Touch: a) from above (top), and b) from below (bottom).**

### Novelty and Learnability

Most participants (11/18) commented that Device Touch was the most novel method. They also found the method entertaining and playful. *"[P11] It was new. Exciting. I haven't thought you could move [objects] like that."* These comments were supported by the high HQ-S score of Device Touch in AttrakDiff. On the other hand, novelty made the method difficult to learn, as also indicated by the LEA score in Extended NASA-TLX. Learning the method was further complicated by the asymmetry of the method: the participants had to actually learn and use two different methods, one for moving objects from phone to tablet and another for moving objects from tablet to phone. This initially confused most users (13/18). *"[P11] It required a lot of learning. It was difficult to remember which device I should start with. What to touch and do."* After learning the method, half of the participants (9/18) commented that it felt very natural and intuitive.

The participants considered Tray and Transfer Mode to be more traditional and anticipated solutions for moving objects between devices. The similarity of Tray and Transfer Mode with familiar touch interfaces made them easier to learn.

### Physical Device Handling

While some participants (5/18) were familiar with NFC, Device Touch was a new interaction method for all of them. When touching the screen of another device, most participants (11/18) held the phone from above (see Figure 7.a), so that the device was hanging below the palm. Only five participants held the phone from below (see Figure 7.b), so that the phone was resting on the palm. Two participants did not have a preferred grip but continued to use different grips throughout the evaluation. Many participants (10/18) accidentally pressed the side buttons of the phone (for example, the camera or volume buttons) when using Device Touch. Three participants considered Device Touch to be easier to use with the smaller phone – the others did not see differences between the phones. When working with two phones, most participants (12/18) always held the phone in their dominant hands when they touched the display of the tablet. Four participants used both hands, holding one phone in their right hands and the

other in their left hands. One participant always used their non-dominant hand, while one participant did not have a clear preference.

Some participants (6/18) were concerned that the display of the tablet might be damaged when touched with the corner of the phone. *"[P6] Sharp glass corners, hard surfaces… It feels a bit scary to touch. You could scratch the tablet screen."* Two participants suggested that the phone could have a special touch corner made of softer material that would not damage the other device. A few participants (4/18) accidentally tried to use the phone as a stylus for normal finger touch interactions, for example, to drag and rearrange objects on the display. Three participants who held the phone in their left hands complained that it was difficult to see the preview of the photo. The device should detect which hand the participant is using to hold the phone and position the preview picture so that it is not behind the phone.

When using Tray and Transfer Mode methods, most participants held the phones on the table. The phones easily slid on the table when the participants made swipe gestures or dragged objects on the screen, requiring them to hold the device with the other hand. Three participants rearranged the devices on the table or moved their chair sideways depending on the focus of the task.

### Strategies

The methods allowed for many possible approaches and strategies on how to complete the tasks. The objects could be moved as a single operation involving all three devices, or as a series of moves between pairs of devices. The objects could be moved individually one at a time, in several groups, or all as a single group. The objects could be arranged in the correct order when selecting them on the source device, when moving them to the target device, or as a separate step after the move had been completed. The method influenced the strategy, with the participants using different strategies with different methods.

In Tray, most of the participants (11/18) opened the tray on each device only when they needed to access it. Others

(7/18) opened the tray first on all devices. The participants (12/18) preferred to move objects in groups – typically all objects to be moved between each pair of devices were transferred as one group. The tray worked as an intermediate storage where to collect objects. It provided a good visualization of the currently selected set of objects and supported maintaining the task status and planning the next steps. Most participants (13/18) collected the objects to the tray in a random order and arranged them when they were moved from the tray to the grid on the target device.

In Transfer Mode, the participants were divided between two different strategies. About half of the participants (8/18) quickly collected all objects to the target device, selecting and confirming objects in small groups, and then arranged them separately on the target device in the end. Other participants (7/18) preferred to select the objects in the correct order. Many of these participants (4/18) moved objects one at a time, confirming each transfer on the target device before selecting the next object.
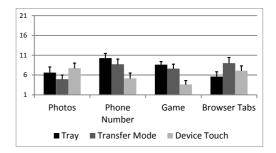
In Device Touch, most participants (12/18) preferred to select and move the objects in the correct order, adding new objects to the end of the grid on the target device. The possibility to move the object to the correct position on the tablet display when touching or confirming the move was little used. In Task T1, where the objects were distributed across two source phones in a random order, the participants switched phones, checking which phone the next object was on and then selecting that phone. In Task T2, they first collected all objects to the first phone and then put the phone on the table. Then they took the other phone and picked all objects going to that phone.

**Scenarios**

Figure 8 shows participants' preferences of methods for the four object movement scenarios that were presented to them. One-way repeated measures ANOVA tests indicated that the object movement method had significant effect on Scenario S3 Game ($F(2, 16) = 6.28$, $P = .01$). Pairwise comparisons showed that the level of preference was significantly lower for Device Touch compared to both Tray ($p = .006$) and Transfer Mode ($p = .043$). There were no statistically significant differences between methods in other scenarios.

The Scenario S1 Photos was closest to the photo sorting tasks that the participants were asked to do during the evaluation. Similarly to the Extended NASA-TLX and AttrakDiff results, Transfer Mode was preferred for the Photos scenario. The primary criteria for preferring Transfer Mode was the efficiency of transferring multiple photos. Some participants (5/18) were also concerned that using Device Touch might be inconvenient in the limited working space available in a crowded train and might draw unwanted attention from the other persons present.

The Scenarios S2 Phone Number and S3 Game involved transfers of a single content item and an application



**Figure 8. Method preferences for different scenarios (lower value is better).**

window. For these tasks, the participants preferred the Device Touch method. The primary criteria was that it was immediately available and did not require any time-consuming setup operations. This was particularly important in Scenario S3 Game as one of the devices was running out of battery. The initialization steps of Tray and Transfer Mode were considered too heavy for moving a single item. *"[P11] If I would have to use Tray or Transfer Mode, I would rather type the phone number directly on my phone."*

In Scenario S4 Browser Tabs, the participants were again asked to move multiple objects, which made Device Touch unattractive. The participants preferred Tray as it allowed them to divide the task into two parts. They could move the browser tabs to the tray on the tablet in the café and then put the tablet in the bag. The participants did not have to immediately continue with the phone but could postpone it until they were sitting in the bus and it was convenient for them to continue reading the pages.

**DISCUSSION**

**Factors Influencing Method Preference**

The study results indicate that the best method for moving objects between devices strongly depends on the task and context of use. In different scenarios, the participants preferred different methods. This makes designing a single optimal method for moving objects between devices a challenging task. We identified several clusters of contextual factors that influenced the participants' perceptions on the preferences of the methods.

*Number of objects.* Moving a single object is a common special case and object movement methods should offer efficient performance in such a case. However, methods that are optimized for moving a single object may not be efficient for moving multiple objects. When moving a large number of objects, even very complicated initialization operations can be justified if they improve the performance of the actual object movement actions as the initialization cost can be allocated over a large number of objects.

*Number of devices.* With three or more devices, the user needs to explicitly indicate which of the other devices the

object should be moved to. However, if only two devices are involved in the movement operation, it is obvious which device the user wants to move the object to and indicating the target device is not necessary.

*Device characteristics.* Devices have different physical characteristics and technical capabilities, which influence the suitability of object movement methods. As a practical example, touch interactions are easier on devices with large screens, while small devices are easier to manipulate and move around with hands. If the devices have very different characteristics, different methods may be needed for moving objects in different directions. However, such asymmetric behavior may easily confuse the user.

*Access to devices.* Transfer Mode and Device Touch required the user to have simultaneous access to all devices. With Tray, the user could move the objects to the tray with the first device, even if they had no access to the other devices. The user could then put the first device away and, possibly much later, access the second device and complete the transfer.

*Social situation.* The presence of other people in proximity raised issues related to privacy and security. In Tray and Transfer Mode, where the movement actions were based on virtual concepts, the participants were concerned that they would accidentally move the object to some other person's device, or another person could somehow catch the object while it was being transferred. There were also differences between the methods in how easily the others could perceive what the person was doing. Many participants preferred discreet methods in public places with unknown people present. On the other hand, some participants thought that using a novel method, such as Device Touch, could impress their friends. While beyond the focus of our study, moving objects between devices of different persons is naturally also very relevant in social situations.

*Physical environment.* Finally, the scenarios raised some potential issues related to the physical environment of the object movement task. Some participants commented that if the available space was limited, such as in a café table full of coffee mugs and other tableware, the use Device Touch, which was the most physical of the methods, would be difficult. Regarding the other methods, the participants were also concerned about making precise touch gestures in a shaking environment, such as in a train.

### Recommendations on the Use of the Studied Methods

Considering the methods included in the study, all of them performed reasonably well in the evaluation and the differences between them were small. Which method works best, depends on the application and the context of use.

Of the three methods, Tray was the most conventional and closest to traditional touch user interfaces. It was therefore easy to use, but it also had the least novelty value. The method provided very good feedback about the task status,

but the tray visualization consumed a lot of screen space, making the method better suited for devices with large displays. Tray was considered inefficient and demanding numerous steps in the object transfer. Tray method might be a good choice if object movement actions are only rarely done and if the participants are not very experienced with information technology. The Tray method also gave the possibility to divide the task into several independent steps, since the tray did not have to be open on all devices at the same time. This feature could be very useful if the participant cannot easily access all the devices simultaneously.

Transfer Mode was the most efficient of the methods. It was well suited for moving large amounts of objects, as after the method had been set up, making individual moves was very fast. However, the necessary initialization steps made it too heavy for moving single objects. Transfer mode was also considered simple, making it easy to learn. The method was preferred for photo sorting tasks that the participants were asked to do during the evaluation. On the other hand, the method was probably also the most specific, being optimized for photo sorting tasks, which may make it more difficult to apply in other applications.

Device Touch was considered to be the most exciting method and also to have the most novelty value. As such, it might be a good choice for entertainment applications and games. On the other hand, the method was considered difficult to learn, partly because it worked differently depending on which direction the user was moving the object. This makes the method better suited for experienced users, or situations where objects need to be moved in one direction only. The method worked well for quick transfers of individual objects as it could be immediately used without complex initialization steps. On the other hand, it was considered laborious for larger numbers of objects, as the user had to physically touch the other device for every object that was moved. This might be solved by allowing the user to select multiple objects and then move them all with a single touch action. As a physical action, the use of Device Touch is more easily perceived by other persons in proximity, which can be a benefit or drawback depending on the context. Special care should be taken to address the participant concerns about scratching the device displays with carefully selected design and materials of the touch corner.

### CONCLUSION

We have presented a comparative evaluation of three different methods for moving visual objects between displays of personal mobile devices: Tray, which enables the user to move objects through a virtual tray shared between devices; Transfer Mode, which requires the user to first connect the devices but then allows moving objects by simply tapping them; and Device Touch, which supports moving objects by using one device to physically touch the other device. The study results indicate that all methods

performed reasonably well in the evaluation and the differences between them were small, but still some interesting distinctions were observed. Tray was perceived as familiar, provided excellent visual feedback about the task status, and allowed the movement task to be divided into logical parts, but was considered somewhat inefficient and boring. Transfer Mode was considered the most efficient method for transferring large numbers of objects, but the required initialization steps were too heavy for moving single objects. Device Touch was considered the most novel method and it was preferred for quick transfers of single objects, but it was difficult to learn (partly due to the asymmetry of the method) and laborious for moving multiple objects. Importantly, it was found that the participants preferred different methods in different scenarios of use, with the preferences strongly depending on contextual factors: the numbers of objects and devices involved, the characteristics of the devices, the user's access to the devices, and the physical and social environment. This makes designing a single optimal method for moving objects between devices a challenging task.

**ACKNOWLEDGMENTS**

**REFERENCES**

1. Bachl, S., Tomitsch, M., Kappel, K. and Grechenig, T. The effects of personal displays and transfer techniques on collaboration strategies in multi-touch based multi-display environments. In *Proc. INTERACT '11*, 373-390.

2. Baudisch, P., Cutrell, E., Robbins, D., Czerwinski, M., Tandler, P., Bederson, B. and Zierlinger, A. Drag-and-Pop and Drag-and-Pick: techniques for accessing remote screen content on touch- and pen-operated systems. In *Proc. INTERACT '03*, 57-64.

3. Biehl, J. and Bailey, B. ARIS: an interface for application relocation in an interactive space. In *Proc. GI '04*, 107-116.

4. Chen, N., Guimbretiere, F. and Sellen, A. Designing a multi-slate reading environment to support active reading activities. *ACM TOCHI* 19, 3 (2012), article 18.

5. Girouard, A., Tarun, A. and Vertegaal, R. DisplayStacks: interaction techniques for stacks of flexible thin-film displays. In *Proc. CHI '12*, 2431-2440.

6. Hamilton, P. and Wigdor, D. Conductor: enabling and understanding cross-device interaction. In *Proc. CHI '14*, 2773-2782.

7. Hart, S. G. and Staveland, L. E. Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In Hancock and Meshkati (eds.) *Human Mental Workload*, North Holland Press, 1988.

8. Hascoët, M. Throwing models for large displays. In *Proc. HCI '03*, 73-77.

9. Hassenzahl, M. The interplay of beauty, goodness, and usability in interactive products. *Human-Computer Interaction* 19, 4 (2004), 319-349.

10. Jokela, T., Ojala, J. and Olsson, T. A diary study on combining multiple information devices in everyday activities and tasks. In *Proc. CHI '15*, 3903-3912.

11. Lucero, A., Holopainen, J. and Jokela, T. Pass-them-around: collaborative use of mobile phones for photo sharing. In *Proc. CHI '11*, 1787-1796.

12. Miller, R. and Myers, B. Synchronizing clipboards of multiple computers. In *Proc. UIST '99*, 65-66.

13. Nacenta, M., Aliakseyeu, D., Subramanian, S. and Gutwin, C. A comparison of techniques for multi-display reaching. In *Proc. CHI '05*, 371-380.

14. Nacenta, M., Gutwin, C., Aliakseyue, D. and Subramanian, S. There and back again: cross-display object movement in multi-display environments. *Human-Computer Interaction* 24, 1-2 (2009), 170-229.

15. Olsen, D. and Nielsen, T. Laser pointer interaction. In *Proc. CHI '01*, 17-22.

16. Oulasvirta, A. and Sumari, L. Mobile kits and laptop trays: managing multiple devices in mobile information work. In *Proc. CHI '07*, 1127-1136.

17. Rekimoto, J. Pick-and-drop: a direct manipulation technique for multiple computer environments. In *Proc. UIST '97*, 31-39.

18. Santosa, S., and Wigdor, D. A field study of multi-device workflows in distributed workspaces. In *Proc. UbiComp '13*, 63-72.

19. Schmidt, D., Chehimi, D., Rukzio, E. and Gellersen, H. PhoneTouch: a technique for direct phone interaction on surfaces. In *Proc. UIST '10*, ACM, New York, NY, USA, 13-16.

20. Schmidt, D., Seifert, J., Rukzio, E. and Gellersen, H. A cross-display interaction style for mobiles and surfaces. In *Proc. DIS '12*, 318-327.

21. Scott, S., Besacier, G. and McClelland, P. Cross-device transfer in a collaborative multi-surface environment without user identification. In *Proc. CTS '14*, 219-226.

22. Tandler, P., Prante, T., Müller-Tomfelde, C., Streitz, N., and Steinmetz, R. Connectables: dynamic coupling of displays for the flexible creation of shared workspaces. In *Proc. UIST '01*, 11-20.

23. Ullmer, B., Ishii, H. and Glas, D. mediaBlocks: physical containers, transports, and controls for online media. In *Proc. SIGGRAPH '98*, 379-386.

**Publication P5**

Tero Jokela, Parisa Pour Rezaei, and Kaisa Väänänen. 2016. Natural Group Binding and Cross-Display Object Movement Methods for Wearable Devices. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '16)*. ACM, New York, NY, USA, 206–216. DOI: https://doi.org/10.1145/2935334.2935346

# Natural Group Binding and Cross-Display Object Movement Methods for Wearable Devices

**Tero Jokela**
Nokia Technologies
Tampere, Finland
tero.jokela@nokia.com

**Parisa Pour Rezaei**
Tampere Univ. of Technology
Tampere, Finland
parisa.pourrezaei@tut.fi

**Kaisa Väänänen**
Tampere Univ. of Technology
Tampere, Finland
kaisa.vaananen@tut.fi

## ABSTRACT

As wearable devices become more popular, situations where there are multiple persons present with such devices will become commonplace. In these situations, wearable devices could support collaborative tasks and experiences between co-located persons through multi-user applications. We present an elicitation study that gathers from end users interaction methods for wearable devices for two common tasks in co-located interaction: group binding and cross-display object movement. We report a total of 154 methods collected from 30 participants. We categorize the methods based on the metaphor and modality of interaction, and discuss the strengths and weaknesses of each category based on qualitative and quantitative feedback given by the participants.

## Author Keywords

Co-located interaction; multi-device user interfaces; wearable devices; smartwatches; smartglasses; device ecosystem binding; group association; pairing; cross-display object movement; guessability study; elicitation study.

## ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous

## INTRODUCTION

We are seeing an increasing diversity of digital devices with the emergence of new wearable form factors, such as smartwatches and smartglasses. Currently, wearable devices are still relatively rare and the early applications focus on individual use. However, as more and more people acquire and start to use such devices, situations where there are multiple persons present with wearable devices will become commonplace. In such situations, wearable devices provide interesting opportunities to support collaborative tasks and experiences between co-located persons through

multi-user applications [12]. For example, two families on a holiday trip together could collect all photos and videos that they capture with their smartglasses into a shared album; a group of friends biking together could use their smartwatches to track the locations and performances of each other; or a team of rescue workers arriving at an accident scene could use their wearable devices to communicate and to share situational information with each other.

But before a group of co-located persons can engage in spontaneous collaborative interactions with their devices, they must first connect their devices together into a group to make the devices aware of each other and to enable data communications between them. This problem is called ad-hoc group binding or group association [6] (also known as pairing or coupling). Another common challenge in co-located interaction is moving virtual objects (such as content items or application windows) between device displays. This problem is generally known as cross-display object movement [21]. Both tasks involve complex technical procedures, but still from a user's perspective, it should be possible to perform these tasks in a fast and easy way – if the process is too time-consuming or tedious, the users might lose their interest in multi-user applications in the first place. Both device binding and cross-display object movement have been extensively studied in prior research and a wide range of potential solutions have been proposed to both problems. However, most of the existing solutions have been designed for individual users rather than for groups and have been driven by technology and security considerations rather than by user experience. Most of the existing methods have also been designed for conventional devices and may not be suitable for or take full advantage of the features of wearable devices [12].

In this paper, we present an elicitation study which aims to collect methods for group binding and cross-display object movement tasks on wearable devices from groups of ordinary end-users. The question we address is what would a group of co-located users naturally do to connect their wearable devices or to move objects between them. We cover both smartglasses and smartwatches, currently the two most common types of wearable devices. We report a total of 154 methods collected from 30 participants, categorize the methods based on the metaphor and modality of interaction, and discuss the strengths and weaknesses of each category based on qualitative and quantitative

feedback given by the participants. The results inform the design of future wearable device user interfaces and applications that involve group binding and cross-display object movement tasks.

## RELATED WORK

### Ad-Hoc Device Binding Methods

The problem of ad-hoc device binding has been extensively studied in human-computer interaction, ubicomp, and security research, and a wide range of different methods for device binding has been proposed [6, 29]. The most common device-binding method today, which is widely used in Bluetooth and WLAN networks, is called Scan & Select. It is based on scanning the environment for available devices and presenting a list of the found devices to the user who can then select the other devices they wish to bind with. The connections are authenticated with short strings that the user should copy or compare between devices.

Examples of alternative device binding methods proposed in prior literature include a range of methods based on synchronous user actions, for example, pressing buttons on both devices simultaneously [25] or shaking the devices together [10]. Methods based on spatial alignment of devices include pointing [19], touching [26], or placing the devices in close proximity of each other [17]. Devices can also be bound by using auxiliary devices, for example, by attaching tokens to the devices [1].

Binding methods are not only means for connecting devices – they have strong social and emotional aspects [12]. Many factors influence persons' preferences of binding methods, including the place, the social setting, the other people present, and the sensitivity of the data [11]. Device association in large groups is a highly collaborative group activity [13, 14, 16]. Therefore, the group binding methods should pay special attention to supporting groupwork and social interactions. The group creation task can be divided in different ways between the members of the group, including leader-driven and peer-based approaches [13, 16]. Group association can be seen as a one-step procedure of binding all devices with a single action, or as a sequence of pairwise associations [5, 29]. The group binding methods should also be flexible and robust, allowing people to adapt them to the changing needs of the situation and to recover from failures [14].

### Cross-Display Object Movement Methods

Conventional solutions for moving content items between devices include connecting the devices directly using physical cables or wireless short-range radio technologies, such as Bluetooth or WLAN. Objects can also be moved through portable storage media, such as memory cards, USB memory sticks, and external hard drives. It is also a common practice to move objects between devices by sending them as e-mails or other kinds of messages. Further, objects can be saved to network folders or storage services and re-opened with other devices. Many cloud services support transferring specific content types (for example, browser tabs, photographs, video, or music) between devices.

In the field of human-computer interaction research, the problem of cross-display object movement has been studied in the context of multi-monitor computer setups, large composite displays, collaborative interaction spaces, tabletop displays, and ad-hoc ecosystems of mobile devices [21]. A wide range of methods for moving visual objects between device displays has been proposed in the prior literature. For example, in Pick-and-Drop [24] the user can pick up an object by touching it with a digital pen and then drop it by repeating the touch action on another display. Objects can also be moved from one display to another by making a throw touch screen gesture towards the target. Many variations of the basic throw gesture, for example, Drag-and-Throw and Push-and-Throw [9] have been suggested, aiming to provide better control over and to improve the accuracy of the throw. In ARIS [3], the user can manage objects across screens on a radar-like map view showing all displays in proximity. Further, objects can be manipulated across displays by using pointing gestures [23]. In MediaBlocks [30], objects can be bound to physical tokens that can be moved between displays.

Nacenta et al. [21] present a taxonomy and classification of cross-display object movement methods based on three conceptual levels: on referential domain level, the methods can reference displays either spatially or non-spatially (for example, using textual names); on display configuration level, the mapping between spatial arrangement of the displays and the method's input model can be planar, perspective, or literal; and on control paradigm level, methods can be open-loop, closed loop, or intermittent. A few studies have compared different cross-display object movement methods against each other. In an early study comparing six different methods, Nacenta et al. [20] identified significant differences between methods, with Radar View and Pick-and-Drop as the fastest and most preferred methods. More recent studies by Jokela et al. [15] and Scott et al. [28] have found that many factors can influence the users' preferences of methods, including the numbers of objects and devices involved, device characteristics, physical and social environment, and personal working styles.

### Elicitation Studies

To generate a set of user-defined input actions, Nielsen et al. [22] and Wobbrock et al. [32] suggest similar elicitation approaches: the participants are first presented with the end effect of an operation and are then asked to perform the action that caused it. Elicitation studies have been most commonly used to produce sets of touch screen gestures [for example, 31] but they have been adapted to other kinds of input actions as well, including hand gestures [27], kick gestures [2], and bend gestures [18]. Closest to our study, Chong and Gellersen report two elicitation studies that

collected device binding methods from individual users [4] and groups of users [5]. While their studies covered a broad range of different device combinations, they did not address group binding methods for wearable devices.

## OUR STUDY

### Objectives

In our study, we were interested in collaborative multi-user applications for groups of co-located users with wearable devices. In particular, we were interested in interaction methods for two common tasks in co-located interaction: group binding and cross-display object movement. Regarding devices, we decided to focus on the two most common wearable device categories today: smartglasses and smartwatches.

Both device binding and cross-display object movement have been extensively studied in prior research and a wide range of potential solutions have been proposed to both problems. However, most of these solutions have been designed for a single user interacting with two devices. Scenarios involving multiple users and devices differ in many respects from single-user scenarios, making the single-user solutions not necessarily applicable to multi-user scenarios. Most of the prior research has also been driven by technology and security considerations. The existing solutions have often been invented by researchers and system designers, focusing on a particular technology solution, while the intuitiveness of the solution to non-technical users has been given little consideration. Finally, most of the existing methods have been designed for conventional devices such as computers, phones, and tablets, and are not necessarily applicable to wearable devices, which are far more personal and intimate. They also may not take advantage of the unique features of wearable devices that could enable more natural and innovative solutions [12].

To address these concerns, we wanted to approach the problem from a perspective of what pairs and small groups of people would naturally and intuitively do to connect their wearable devices into a group or to move virtual objects between their wearable devices. Therefore, we adopted an elicitation study approach and asked groups of participants to come up with their own techniques for these tasks. The participants took turns to suggest different methods and then immediately tested and evaluated the methods with mock-up devices. The participants could suggest any technique with only one restriction: as earlier studies suggest that people are unwilling to hand in their personal devices to strangers [8], the participants were not allowed give their devices to the other participants. The primary objective of our study was to collect and preliminarily evaluate a broad set of suggestions for natural group-binding and cross-display object movement methods for wearable devices, in order to inform the design of future multi-user applications for groups of co-located users with wearables.



**Figure 1. Device surrogates used in the study: smartglasses (left) and smartwatch (right).**

Our study was inspired by similar studies on device binding methods by Chong and Gellersen [4, 5]. Our study extends their work in three major directions: we address (1) groups of users binding their watches together (Chong and Gellersen only considered a single user binding a digital watch with a phone, a tablet, or a display); (2) binding of smartglasses (not considered by Chong and Gellersen); and (3) cross-display object movement (not considered by Chong and Gellersen).

### Devices

We decided not to use any commercially available wearable devices in the study, as the devices' native user interface styles and technical features and limitations might have guided the participants' proposals and limited their creativity. Instead, we provided the participants with simple mock-up devices that acted as surrogates of real devices. Figure 1 illustrates the device surrogates used in the study. We simulated smartglasses with ordinary 3M safety spectacles. For simulating smartwatches, we built custom mock-ups by attaching a Casio watch band to a small block of polystyrene foam. The dimensions of the block were 48x38x12 mm (1.9x1.5x0.5 inches) which is comparable to the currently available commercial smartwatches.

### Participants

We recruited eight groups of four participants for the study by posting an advertisement on local mailing lists and social media. However, two persons canceled their participation at the last moment, which resulted in two of the groups having only three participants and reduced the total number of participants to 30. We encouraged the participants to recruit also their friends, so that some participants would know each other to make the study situation more natural and comfortable. 15 of the participants knew one of the other participants in the same group and eight participants knew several of the other participants, while seven knew none of the other participants. 14 of the participants were female and 16 male. The ages of the participants varied between 15 and 56 years (M=32.7, SD=12.6). Two of the participants were left-handed and 28 right-handed. 14 participants had educational backgrounds related to information technology,

while the other 16 participants represented a wide variety of different professions (for example, a school teacher, a nurse, and a bus driver). On a scale between 1 and 7 (1=novice, 7=expert), the participants rated their familiarity with information technology slightly above average (M=4.7, SD=1.9). Seven participants had tried smartglasses (for example, Google Glass) and 11 participants had tried a smartwatch or other smart wrist device. 11 participants were wearing eyeglasses continuously and six as needed, while 13 participants did not have eyeglasses. Eight participants were wearing a wristwatch every day and seven sometimes, while 15 participants did not have a wristwatch.

## Procedure

We organized a total of eight study sessions. The sessions were arranged in our usability laboratory. In each session, there were three or four participants and a moderator present. Figure 2 shows the evaluation setup. The durations of the sessions varied between 75 and 120 minutes.

As the participants arrived in the laboratory, the moderator asked them to fill in the background information and consent forms. The moderator then introduced the participants to wearable devices and the idea of several co-located persons using their wearables together for collaborative applications. The moderator explained that the objective of the study was to invent and evaluate interaction methods for two common tasks in multi-user applications: group binding and cross-display object movement. The moderator then continued by describing the detailed study procedure.

To begin the actual study, the moderator gave each participant the first mock-up device. If the participant was wearing eyeglasses, they could wear the protective glasses used to simulate smartglasses over the eyeglasses if needed. However, if the participant was wearing a wristwatch, they were asked to remove the wristwatch when wearing the smartwatch mock-up. The moderator then described the device in more detail and gave several examples of different visual, audio, and haptic interaction methods that the device could support. However, the moderator emphasized that these were just examples and that the devices could have any capabilities the participants wanted them to have. After that, the moderator gave the participants the first task and explained it in detail. The moderator asked one of the participants to suggest a method how the task could be achieved with the devices. The moderator encouraged the participants to suggest the first intuitive ideas that spontaneously occurred in their minds and reminded that they could do anything but to give their devices to another participant.

When the first participant had described a method, the moderator asked all participants to stand up and try it out with the device mockups, first in pairs and then as a single group of three or four persons. The moderator portrayed an external person not to be included in the group or to receive the object. Trying the method with the mockup devices



Figure 2. Participants suggesting and evaluating methods during the study.

clarified the details of how it would work and gave the participants a better understanding of its strengths and weaknesses in practice. After the participants had tried the method, the moderator asked them to fill in a paper form to evaluate the method in terms of practicality (that is, how easy, effortless, efficient, and error-free it was to use) and pleasantness (that is, how human, connective, inspiring, and inventive it was). After filling in the forms, the moderator asked the participants to provide brief immediate free-form verbal comments about the method. The moderator then asked the next person to suggest another method, which was similarly tested and evaluated. When every participant had suggested a method, the moderator offered an opportunity to any participant to suggest further ideas. Overall, the groups tested three to seven methods for each device-task combination.

The same procedure was then repeated for the other task with the same device, and after that for both tasks with the other device. The order of the devices and methods was systematically varied between study sessions to counter-balance any learning effects. At the end of the study session, the moderator briefly interviewed the participants for general comments about the wearable devices and tasks used in the study. To close the session, the moderator thanked the participants and gave each participant a small reward to compensate for their time.

The study sessions were recorded with a video camera. All the proposed methods were documented and categorized by three researchers. All participant comments about the methods as well as interview responses were transcribed and thematically analyzed. A quantitative analysis of the evaluation responses was done separately.

## RESULTS

### General

The participants proposed a total of 154 different methods during the study. Of these methods, 73 were intended to be used with smartglasses, including 38 group binding methods and 35 cross-display object movement methods. The remaining 81 methods were intended to be used with

smartwatches and consisted of 40 group binding methods and 41 cross-display object movement methods.

Both group binding and cross-display object movement tasks could be divided into three main phases: preparation, target identification, and confirmation. The *preparation phase* consisted of various activities that were needed to initiate the task. It could include discussing with the other participants and agreeing on a common target and strategy to achieve it. It could also include initiating the necessary technical functions on the devices, such as activating a certain device mode, or creating a new group for others to join, or browsing and selecting the object to be shared with the others. In the *target identification phase*, links between the devices were created, that is, it was indicated which devices were intended to participate in the group or to receive the object. The *confirmation phase* consisted of making sure that the intended action was successfully completed. The person who created the group or shared the object wanted to make sure that all the intended persons had been included in the group or received the object, and that there were no external persons included. On the other hand, the persons who joined in the group or received the object wanted to make sure that they were only in those groups that they wanted to be part of, or received only those objects that they wanted to have. It is important to note that all phases consisted of both technical actions involving interactions with the devices and social actions involving interactions with the other participants. Many actions had both technical and social dimensions – for example, if one person was interacting with their device by tapping the screen with their finger, this could be observed by the other persons and was intertwined with the social interaction within the group.

After an initial analysis of the collected methods, we decided to categorize the methods based on the techniques used in the target identification phase, as that appeared to be the most characteristic phase of the interaction where the actual connections between the participants were formed. The same categorization was used for both group binding and cross-display object movement methods. The categorization was primarily based on the *metaphor of interaction* and secondarily on the *modality of interaction*. Three researchers first independently analyzed and categorized each method. The individual categorizations were then compared and the methods where the researchers disagreed were commonly discussed and resolved. Table 1 presents the resulting categorization. We will next explain and discuss each main category in detail.

### Spatial

The methods in the Spatial category were based on the relative or absolute positions of the participants. Spatial methods were the most popular category for smartglasses, but they were also commonly suggested for smartwatches. Spatial methods were equally proposed for both group binding and object movement tasks. The methods in this

category could be divided into three subcategories: pointing-based, proximity-based, and area-based methods.

The participants suggested several different approaches for selecting the target person by pointing. For smartglasses, by far the most popular approach was to point at the target with gaze. *"[P25] Eye contact. I look directly at my target person and then the [object] will be transferred to her."* In some variants, it was enough just to look in the direction of the target person, while in other variants it was required to establish an eye contact between the persons, providing the target with a possibility to reject the interaction. While some participants considered looking at another person a natural element of human communication, others felt that it was disturbing to have to stare at the eyes of another, possibly unfamiliar, person. Gaze pointing, like all pointing-based methods, was also thought to be slow for larger groups of people as the person had to point at every target person in sequence, one at a time. As a solution, some participants suggested that all the persons in the field of view could be selected with a single operation. The person could then select their viewpoint so that only the target persons they wished to select were visible, or alternatively they could hide the unwanted persons, for example, with their hands. For smartwatches, the most popular approach was to point at the target with the device, aiming and shooting like with a camera or a laser ray. These methods were considered as fun and playful. *"[P21] There is the Star Trek fun factor. Pew, pew, pew… You could add a lazer sound."* A third possible approach, which was suggested for both smartglasses and smartwatches, was to point at the target with a finger or with a hand. While generally considered easy and natural, some participants were concerned that pointing with a finger could appear as rude or ridiculous in some social situations.

In proximity-based methods, the devices were capable of determining the closest person or group of persons. The participants could then make selections by moving close to each other. While simple, the methods were considered error-prone in situations where there were a lot of people present. Some participants also considered moving very close to unfamiliar persons annoying. Finally, in area-based methods the participants defined an area to select all persons within that area. The area could be defined either by setting the radius of a circle around the user, or by explicitly drawing the area boundaries. Defining an area was seen as an efficient though imprecise way to select a large group of people. Selecting a small group people in a crowded area could force the participants to define a very small area and to move very close to each other to fit in that area.

### Touching

The most commonly suggested method for smartwatches was device touch, that is, bringing the watches into physical contact or very close proximity of each other for a brief period of time. *"[P32] You select [the object to share] and*

| | Smartglasses | | | | | | | | | | | | Smartwatch | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | All | | | | Group binding | | | | Object movement | | | | All | | | | Group binding | | | | Object movement | | | |
| | S | N | PRA | PLE | S | N | PRA | PLE | S | N | PRA | PLE | S | N | PRA | PLE | S | N | PRA | PLE | S | N | PRA | PLE |
| ALL | 8 | 73 | 5,03 | 4,99 | 8 | 38 | 5,03 | 4,99 | 8 | 35 | 5,03 | 4,99 | 8 | 81 | 5,34 | 5,32 | 8 | 40 | 5,38 | 5,32 | 8 | 41 | 5,31 | 5,32 |
| SPATIAL | 7 | 30 | 4,92 | 4,99 | 7 | 14 | 5,02 | 5,07 | 7 | 16 | 4,84 | 4,91 | 7 | 13 | 4,89 | 5,02 | 6 | 7 | 5,17 | 5,13 | 4 | 6 | 4,53 | 4,88 |
| Pointing | 7 | 25 | 5,04 | 4,99 | 5 | 11 | 5,19 | 5,05 | 7 | 14 | 4,92 | 4,94 | 5 | 8 | 4,91 | 4,98 | 2 | 3 | 5,62 | 5,5 | 3 | 5 | 4,41 | 4,62 |
| *With gaze* | 7 | 21 | 5 | 4,97 | 5 | 9 | 5,21 | 5,09 | 7 | 12 | 4,84 | 4,88 | 1 | 1 | 4,12 | 4,25 | 0 | 0 | N/A | N/A | 1 | 1 | 4,12 | 4,25 |
| *With hand* | 4 | 4 | 5,23 | 5,1 | 2 | 2 | 5,06 | 4,88 | 2 | 2 | 5,43 | 5,36 | 1 | 2 | 4 | 3,67 | 0 | 0 | N/A | N/A | 1 | 2 | 4 | 3,67 |
| *With device* | 0 | 0 | N/A | N/A | 0 | 0 | N/A | N/A | 0 | 0 | N/A | N/A | 4 | 5 | 5,37 | 5,55 | 2 | 3 | 5,62 | 5,5 | 2 | 2 | 4,93 | 5,64 |
| Proximity | 3 | 3 | 3,64 | 4,27 | 2 | 2 | 3,57 | 4,29 | 1 | 1 | 3,75 | 4,25 | 1 | 2 | 5 | 5,67 | 1 | 1 | 4,83 | 5 | 1 | 1 | 5,17 | 6,33 |
| Define area | 2 | 2 | 5,31 | 5,94 | 1 | 1 | 5,75 | 6,75 | 1 | 1 | 4,88 | 5,12 | 3 | 3 | 4,79 | 4,79 | 3 | 3 | 4,79 | 4,79 | 0 | 0 | N/A | N/A |
| TOUCHING | 5 | 9 | 5,01 | 4,82 | 5 | 7 | 5,17 | 4,85 | 2 | 2 | 4,5 | 4,75 | 8 | 21 | 5,53 | 5,63 | 8 | 11 | 5,4 | 5,58 | 6 | 10 | 5,67 | 5,68 |
| Device-device | 4 | 6 | 5,27 | 4,95 | 4 | 5 | 5,31 | 4,94 | 1 | 1 | 5,12 | 5 | 8 | 19 | 5,64 | 5,72 | 8 | 10 | 5,54 | 5,71 | 6 | 9 | 5,75 | 5,74 |
| Person-person | 3 | 3 | 4,54 | 4,58 | 2 | 2 | 4,88 | 4,62 | 1 | 1 | 3,88 | 4,5 | 1 | 1 | 5 | 5,25 | 0 | 0 | N/A | N/A | 1 | 1 | 5 | 5,25 |
| Person-device | 0 | 0 | N/A | N/A | 0 | 0 | N/A | N/A | 0 | 0 | N/A | N/A | 1 | 1 | 4,12 | 4,38 | 1 | 1 | 4,12 | 4,38 | 0 | 0 | N/A | N/A |
| COMMAND | 6 | 15 | 4,88 | 4,62 | 4 | 6 | 5 | 4,89 | 5 | 9 | 4,79 | 4,46 | 6 | 16 | 5,06 | 4,75 | 4 | 5 | 4,95 | 4,61 | 6 | 11 | 5,11 | 4,82 |
| GUI | 3 | 7 | 4,25 | 3,77 | 3 | 3 | 4,27 | 3,95 | 3 | 4 | 4,23 | 3,63 | 4 | 7 | 5,26 | 4,91 | 3 | 3 | 5,14 | 4,77 | 3 | 4 | 5,34 | 5 |
| Voice | 3 | 4 | 4,96 | 5,25 | 2 | 2 | 5,5 | 5,64 | 2 | 2 | 4,43 | 4,86 | 6 | 7 | 4,96 | 4,87 | 1 | 1 | 4,38 | 4,75 | 6 | 6 | 5,07 | 4,89 |
| Thought | 3 | 4 | 5,81 | 5,47 | 1 | 1 | 6,12 | 6,12 | 3 | 3 | 5,71 | 5,25 | 1 | 2 | 4,69 | 3,88 | 1 | 1 | 5 | 4 | 1 | 1 | 4,38 | 3,75 |
| SCAN & SELECT | 3 | 3 | 4,41 | 3,86 | 3 | 3 | 4,41 | 3,86 | 0 | 0 | N/A | N/A | 8 | 13 | 5,62 | 5,48 | 7 | 7 | 5,44 | 5,19 | 4 | 6 | 5,81 | 5,79 |
| List | 3 | 3 | 4,41 | 3,86 | 3 | 3 | 4,41 | 3,86 | 0 | 0 | N/A | N/A | 6 | 10 | 5,97 | 5,68 | 5 | 5 | 5,81 | 5,42 | 4 | 5 | 6,12 | 5,92 |
| Spatial map | 0 | 0 | N/A | N/A | 0 | 0 | N/A | N/A | 0 | 0 | N/A | N/A | 3 | 3 | 4,83 | 4,83 | 2 | 2 | 4,62 | 4,69 | 1 | 1 | 4,25 | 5,12 |
| SHARED ACTION | 3 | 5 | 5,47 | 5,81 | 3 | 5 | 5,47 | 5,81 | 0 | 0 | N/A | N/A | 5 | 10 | 5,39 | 5,39 | 5 | 9 | 5,61 | 5,6 | 1 | 1 | 3,62 | 3,75 |
| Hand gesture | 1 | 1 | 4,5 | 5,38 | 1 | 1 | 4,5 | 5,38 | 0 | 0 | N/A | N/A | 4 | 6 | 5,23 | 5,55 | 3 | 5 | 5,58 | 5,94 | 1 | 1 | 3,62 | 3,75 |
| Touch gesture | 0 | 0 | N/A | N/A | 0 | 0 | N/A | N/A | 0 | 0 | N/A | N/A | 1 | 1 | 6,33 | 6,5 | 1 | 1 | 6,33 | 6,5 | 0 | 0 | N/A | N/A |
| Eye gesture | 1 | 1 | 6 | 6,67 | 1 | 1 | 6 | 6,67 | 0 | 0 | N/A | N/A | 0 | 0 | N/A | N/A | 0 | 0 | N/A | N/A | 0 | 0 | N/A | N/A |
| Button press | 2 | 2 | 5,57 | 6,14 | 2 | 2 | 5,57 | 6,14 | 0 | 0 | N/A | N/A | 1 | 1 | 5,17 | 3,83 | 1 | 1 | 5,17 | 3,83 | 0 | 0 | N/A | N/A |
| Passkey | 1 | 1 | 5,88 | 5 | 1 | 1 | 5,88 | 5 | 0 | 0 | N/A | N/A | 2 | 2 | 5,57 | 5,07 | 2 | 2 | 5,57 | 5,07 | 0 | 0 | N/A | N/A |
| VIRTUAL OBJECT | 4 | 7 | 5,67 | 5,65 | 1 | 1 | 4,88 | 5 | 4 | 6 | 5,82 | 5,77 | 3 | 6 | 5,73 | 6,02 | 0 | 0 | N/A | N/A | 3 | 6 | 5,73 | 6,02 |
| Throwing | 4 | 5 | 5,34 | 5,34 | 1 | 1 | 4,88 | 5 | 4 | 4 | 5,47 | 5,43 | 3 | 5 | 5,69 | 5,94 | 0 | 0 | N/A | N/A | 3 | 5 | 5,69 | 5,94 |
| *Touch screen* | 1 | 1 | 6,5 | 7 | 0 | 0 | N/A | N/A | 1 | 1 | 6,5 | 7 | 3 | 3 | 5,73 | 5,91 | 0 | 0 | N/A | N/A | 3 | 3 | 5,73 | 5,91 |
| *AR* | 3 | 4 | 5,12 | 5,03 | 1 | 1 | 4,88 | 5 | 3 | 3 | 5,21 | 5,04 | 0 | 0 | N/A | N/A | 0 | 0 | N/A | N/A | 0 | 0 | N/A | N/A |
| *Imaginary* | 0 | 0 | N/A | N/A | 0 | 0 | N/A | N/A | 0 | 0 | N/A | N/A | 2 | 2 | 5,64 | 6 | 0 | 0 | N/A | N/A | 2 | 2 | 5,64 | 6 |
| Giving | 2 | 2 | 6,57 | 6,5 | 0 | 0 | N/A | N/A | 2 | 2 | 6,57 | 6,5 | 1 | 1 | 5,88 | 6,38 | 0 | 0 | N/A | N/A | 1 | 1 | 5,88 | 6,38 |
| *Imaginary* | 1 | 1 | 6,83 | 6,83 | 0 | 0 | N/A | N/A | 1 | 1 | 6,83 | 6,83 | 1 | 1 | 5,88 | 6,38 | 0 | 0 | N/A | N/A | 1 | 1 | 5,88 | 6,38 |
| *AR* | 1 | 1 | 6,38 | 6,25 | 0 | 0 | N/A | N/A | 1 | 1 | 6,38 | 6,25 | 0 | 0 | N/A | N/A | 0 | 0 | N/A | N/A | 0 | 0 | N/A | N/A |
| NATURAL BEHAVIOR | 3 | 3 | 5,79 | 5,67 | 1 | 1 | 5,5 | 5,12 | 2 | 2 | 5,94 | 5,94 | 2 | 2 | 5,38 | 5,12 | 1 | 1 | 6,25 | 5,88 | 1 | 1 | 4,5 | 4,38 |
| Natural discussion | 3 | 3 | 5,79 | 5,67 | 1 | 1 | 5,5 | 5,12 | 2 | 2 | 5,94 | 5,94 | 1 | 1 | 4,5 | 4,38 | 0 | 0 | N/A | N/A | 1 | 1 | 4,5 | 4,38 |
| Similar behavior | 0 | 0 | N/A | N/A | 0 | 0 | N/A | N/A | 0 | 0 | N/A | N/A | 1 | 1 | 6,25 | 5,88 | 1 | 1 | 6,25 | 5,88 | 0 | 0 | N/A | N/A |
| REAL OBJECT | 1 | 1 | 4 | 4,62 | 1 | 1 | 4 | 4,62 | 0 | 0 | N/A | N/A | 0 | 0 | N/A | N/A | 0 | 0 | N/A | N/A | 0 | 0 | N/A | N/A |
| Clothes | 1 | 1 | 4 | 4,62 | 1 | 1 | 4 | 4,62 | 0 | 0 | N/A | N/A | 0 | 0 | N/A | N/A | 0 | 0 | N/A | N/A | 0 | 0 | N/A | N/A |

**Table 1. Categorization, numbers, and evaluation scores of the suggested methods (S = number of sessions where the method was proposed, N = number of times the method was proposed, PRA = practicality, PLE = pleasantness).**

*then you just bump your watch with the other person's watch."* Device touch was equally suggested for group binding and object movement tasks. The participants considered device touch easy, fun, straightforward, and accurate to use. Some participants commented that bumping devices strengthened the team spirit and compared it to clinking glasses when making a toast. As an interaction method, device touch was familiar to many participants as they had experience of using NFC. However, a few participants were concerned that touching might physically damage their devices. In some variants of device touch, the target persons had to be selected one at a time, making the method very inefficient with large amounts of people. In most of the suggested variants, however, it was possible that several devices could touch each other simultaneously which improved the efficiency in large groups. Still, the method required the participants to be within touching distance of each other. In a few of the suggested methods, the target persons who had already been selected could also select further persons, making the group membership or the shared object to spread like a chain reaction or a virus

within the group. Several participants also suggested device touch for smartglasses, especially for group binding. These methods required the participant to take off their smartglasses to touch the other person's glasses, which was considered impractical, especially if the person had poor eyesight and the smartglasses were assumed to have prescription lenses.

Another form of touching that was suggested by the participants was two persons touching each other. This typically took the form of a handshake. *"[P9] When you shake hands, both persons are added to the group."* While shaking hands was considered a familiar and natural gesture, it was primarily used in formal situations and felt comical in more relaxed situations such as when meeting with friends. People usually shook hands once when they met – shaking hands many times, first to greet and then to form groups or move objects felt inconvenient. A few participants were also worried that they could accidentally share files when shaking hands, for example, in a job interview. With smartglasses, the persons could shake

hands naturally with their right hands. However, most participants wore the smartwatch on their left hands, requiring them to shake hands with the "wrong hand". Sometimes the participants wore their watches in different hands making the handshake very cumbersome. Finally, one participant also suggested a method where the person creating a group could add other devices to the group by touching the devices with their hand, that is, by pressing the device screen with their thumb. The touching person was identified by the fingerprint which raised concerns about privacy and security.

**Command**

Yet another category of methods that was common for both smartglasses and smartwatches was methods that were based on giving direct commands to the device. Such methods were particularly popular for object movement, but were suggested also for group binding. In Command methods, the target persons were referred to using pre-defined identifiers such as names or contact cards. The most common subcategory of Command methods was methods based on traditional GUI interactions. *"[P7] You have a list of contacts on your watch. You can tap the persons you want to add to the group."* Using pre-defined contacts enabled the participants to prepare groups in advance and include persons who were not currently present, but on the other hand it was limited to persons that the participants knew beforehand. On a smartwatch, the GUI was shown on the watch screen and the user could interact with standard touch screen techniques. This was seen as a familiar, reliable, and discreet way to interact but the screen was considered very small for handling large amounts of contacts. On smartglasses, the participant could see the GUI on a virtual screen floating in the air and interact with gaze, hand gestures, or some kind of a pointing device, for example, a touchpad, integrated into the frames of the smartglasses. The participants were worried that using a virtual display might be dangerous, for example, when walking, as it partly blocked their vision, and had doubts about how well the novel interaction methods would work in practice.

Giving commands by voice was also commonly suggested. *"[P23] I say: 'Create group: [first person's name], [second person's name], [third person's name].' And then it is done. I use the names I have stored on my watch."* Voice commands were more often suggested for smartwatches, and some participants commented that using voice was more natural with a smartwatch, as there was a physical device to talk to that they and the other people in proximity could see. Many participants had experience of using voice commands with conventional devices, and they raised several common problems of voice commands, for example, that they could be annoying to use in some social situations or sensitive to background noise. The voice commands also required the participant to remember the names of the other persons, and listing many names could

be slow, which would be problematic especially with larger groups of unfamiliar people.

Finally, some participants envisioned methods where they could give commands to their device by thought. Using thought commands was slightly more often suggested for smartglasses than smartwatches. While considered futuristic, the participants usually were initially excited about the possibility and perceived thought commands as a very easy, effortless, and discreet way to interact. However, when the participants tested the method in practice, the experience was described as strange and unsocial. *"[P2] I had a feeling that I was missing [the] speech. I think sharing is interaction between people. First, it felt like WOW, this was easy and nice, but when we tried it, it suddenly felt that we are all here silent and everybody is doing something alone in his own bubble with his own brain UI. You don't even have to look [at the others]."* As there were no perceivable indications, the other participants found it very difficult to know whether a person had completed a task, or even started it, and had to query the status by asking verbal questions.

**Scan & Select**

Scan & Select is currently the most commonly used device binding method and it is widely implemented in commercially available phones, tablets, computers, and other devices. As such, it was familiar to most participants and many participants suggested to use it for the study tasks, particularly with smartwatches, which were perceived to be closer to conventional devices. *"[P23] My watch shows all the devices in proximity, like in Bluetooth. Then I select and mark the devices I want to share with from the list and press 'Send'."* Scan & Select methods were equally suggested for both group binding and object movement tasks. In most variants of the method, a participant scanned the environment for available devices and then selected the devices to join in the group or to receive the object. In other variants, a participant first created a group or selected an object and made it accessible to other devices in the proximity. The other participants could then scan the environment with their devices for available groups or objects and select the one they wanted to join or receive.

In general, Scan & Select was considered a practical but not a very exciting solution. Some participants were concerned that in a crowded location, such as in a sports stadium, with a lot of persons and devices, scanning might take a long time and result in a large number of possible targets, which could be difficult to manage on a small smartwatch display. In the vast majority of suggested methods, the discovered targets were presented as a simple list of textual names, posing an additional problem of mapping the names to real-world persons and devices [21]. A few participants suggested spatial representations, for example, showing the targets on a map, which could enable easier mappings between targets on the screen and in the real world.

## Shared Action

For the group binding task, the participants also suggested a range of methods where the persons indicated that they wanted to join in the group by doing simultaneously a common action. *"[P17] I make some movement, for example, shake my hand like this. If you do the same, it detects a similar movement in proximity and connects us into the same group."* While a similar approach could also have been applied to object movement, only one such method was suggested, indicating that a shared action was primarily considered appropriate only for the group binding task. The shared action could be a gesture, for example, a hand gesture, a touch screen gesture (that is, drawing a picture on the screen), or an eye gesture. It could also be a button press or entering a textual passkey. The gestures could be exactly similar for every user, or they could form symmetric pairs, for example, with the first user pushing their hand forward and the second user pulling their hand backward. Typically, all participants could do the action in parallel, making the methods efficient in large groups.

In general, the methods based on shared actions were considered as practical and simple to use and remember as there was only a single action that was done by every person. The participants appeared to enjoy doing a physical action together and sometimes spontaneously accompanied the action by shouting a common phrase together, like in team gestures in sports. The actions could also be clearly observed, making it easy to follow the status within the group. On the other hand, some participants were concerned that the actions might appear as ridiculous to external observers. Several participants were worried that if the actions were common everyday actions, such as nodding or rotating your head, external persons could accidentally join in the group, especially in places with large numbers of people. A malicious external person could also observe the action and do it on their device, in order to join the group without permission. Pressing a button or entering a passkey was also considered somewhat old-fashioned and boring. Text entry in general was considered difficult with wearable devices.

## Virtual Object

In the Virtual Object methods, the object to be shared was represented as a virtual object that could be manipulated like a real object. *"[P13] I can see the file [virtually floating in the air]. I can grab the file and throw it towards you."* With the exception of a single group binding method, the Virtual Object methods were suggested only for the cross-display object movement task. The virtual object could be a purely imaginary object, or it could be visualized as a graphical object on the smartwatch display or as an augmented reality object on the smartglasses. The participants suggested two different interactions for delivering the object: throwing and giving. In the throwing methods, the sender grabbed the object and threw it towards the recipient. In the giving methods, the sender took the object in their hand and offered it on their palm. The recipient could then grab the object and collect it from the sender's palm.

In general, the Virtual Object methods were well liked by the participants and they received high ratings in the questionnaire. The methods were considered practical, natural, fun, and intuitive – even magical. However, like with the Shared Action methods, some participants were worried that they could accidentally share their files by making unintentional gestures. Several participants were also concerned that they might miss the target when throwing and send the file to a wrong destination. The giving methods were considered more secure to use. The throwing methods were also considered inefficient with a large number of people as they required the sender to throw the object to each recipient separately. The giving methods were more efficient as several recipients could take the object from the sender's palm simultaneously.

## Natural Behavior

In addition to methods where the participants did explicit actions to form groups or move objects, a few participants also suggested methods where the devices monitored the participants' natural behavior and pro-actively executed commands on their behalf. These methods could be further divided into two subcategories. In the first subcategory, the devices were monitoring the natural discussion within the group, identified the persons based on their voices, and automatically executed commands based on the discussion. *"[P29] If I wanted to share a photo with a lot of guys, I would just say: 'I have this cool photo. Do you want to see it?' Everyone who would like to have it just said: 'Yes.' Then the photo would be magically shared."* In the second subcategory, the devices were monitoring the participants' behavior in general and clustered in the same group everybody who was behaving in a similar way (for example, moving together or doing similar actions at the same time).

Overall, the methods based on monitoring natural behavior were considered very easy to use and intuitive. *"[P3] This would probably be the most pleasant method. It feels natural."* The methods also did not require any advance setup effort but could be used at any time. They received high scores in both practicality and pleasantness in the questionnaire. However, the participants were skeptical about whether such methods could be implemented in practice and were concerned that they might trigger actions unintentionally, for example, by saying 'yes' accidentally at the wrong moment. A few participants also raised privacy issues related to continuous monitoring of users.

## Real Object

One participant suggested a method where the group membership was indicated by wearing a specific piece of clothing, for example, a certain kind of a t-shirt or a baseball cap, that was recognized by the smartglasses. This method can be seen as an example of a more generic category of methods [1, 30] where the group membership or

the object to be moved is represented by a real physical object (a token) that is attached to the target user or device to indicate that the user should join the group or receive the object.

The participants considered that the method was fast and effective but required that somebody prepared the tokens in advance. The method was seen suitable for organized events where every participant was handed out a piece a clothing, for example, an event t-shirt. Different subgroups, for example, competing teams, could have different kinds of t-shirts.

## DISCUSSION

Compared to conventional devices such as smartphones, tablets, and computers, the wearable devices were seen to provide many new possibilities for interaction. As the wearables were seen as extensions of the user's body, the interaction methods were expected to be more instant, physical, and simple. Of the two wearable device types included in the study, smartwatches were considered to be closer to conventional devices that the participants had plenty of experience with. *"[P7] Smartwatch is like a small smartphone on your wrist."* The smartwatch was also considered to be more discreet and unnoticeable than the smartglasses – this might change in the longer term future, however, as the form and appearance of smartglasses approaches ordinary eyeglasses. The similarity of smartwatches to conventional devices made it easier to invent methods for them but on the other hand it resulted in a large share of traditional methods such as Scan & Select. Smartglasses, on the other hand, were considered futuristic devices with a lot of novel opportunities. *"[P21] Watches and glasses offer different magnitudes of potential. With glasses, you can do unbelievable things if we just invent them."* As the smartglasses had features like near-eye displays and augmented reality that most participants had little experience of using, it was more difficult to generate ideas for them but on the other hand the generated ideas were more versatile and novel.

As reported earlier by Chong and Gellersen [5], our study further confirms that the perceived affordances of the devices significantly influence the suggested methods. The perceived affordances were defined primarily by two factors: first, the natural physiological capabilities and social functions of the body part that the device was attached to (that is, the participant's head vs. the wrist), and second, the technical features that were perceived possible for a device attached to that body part (for example, what kind of body signals could be measured from that location). As a practical example, gaze pointing and thought commands were primarily suggested for smartglasses, as both the eyes and the brain are located in the user's head near the position of the glasses. On the other hand, the methods suggested for smartwatches emphasized actions made with hands, for example, touching and shared hand gestures, while it would have been equally possible to

detect such actions with cameras mounted on the smartglasses.

The participants commented that the group binding and cross-display object movement tasks were similar and felt that in principle, similar solutions could suit to both tasks. However, while some method categories were equally suggested for both tasks, in many categories there was a bias towards one of the tasks. For example, Shared Action and Scan & Select methods were more commonly proposed for the group binding task, while Virtual Object and Command methods were more popular for the object movement task.

The level of how well the participants knew the other participants did not seem to have a major influence on the perceived goodness of the methods. Rather, as there were unknown persons present in the same study group, the participants suggested only methods that were appropriate both with friends and strangers (for example, while handshakes were suggested, hugging was not).

### Future Work

While our study has explored a wide range of methods for group binding and cross-display object movement tasks for both smartglasses and smartwatches, a single study can only cover a limited set of potential scenarios of use. There exist many other possible wearable device configurations that should be addressed in future studies. In addition to smartglasses and smartwatches, there are other wearable device types such as smart shoes, belts, and clothes. Users might wear different types of devices, for example, some wearing smartglasses and others smartwatches, requiring methods which can work across device types. Some of the users might have conventional devices, such as smartphones and tablets. A single user might be wearing multiple devices, such as smartglasses and a smartwatch, opening up a design space of methods that combine several devices [7].

Another factor that may have a strong impact on the group binding and cross-display object movement methods is the group size [29]. In our study, we have addressed pairs of users and small groups of three or four users. As the size of the group increases, a much wider variety of different approaches and strategies becomes possible, and the overall process may become more parallel [14]. Future studies could explore group binding and cross-display object movement with wearable devices in larger groups. Regarding group binding, our study covered only the initial group setup – it did not consider managing the changing group membership over time (that is, adding or removing participants later).

In addition to the methods themselves, we aimed to collect as much feedback as possible about the methods from the participants. However, the participants only evaluated the methods that were suggested in their own groups, implying that different groups evaluated different variants of the

same method and some of the less common methods were not suggested at all in every group. Therefore, it is not possible to reliably prioritize the method categories based on our data – rather our study focused on collecting a broad range of ideas and initial feedback about those. To reliably evaluate the methods, follow-up studies that systematically present the same methods to every participant would be needed. Preferably these studies should use real prototypes as while many methods can work with mockups, in real life many technical aspects need to be considered [14]. The studies should also address different contexts of use [15, 13]. Ideally, the methods should be tested over extended periods of time in real-life environments [6].

## CONCLUSION

We have presented an elicitation study collecting interaction methods for group binding and cross-display object movement tasks on wearable devices from ordinary end users. We have reported a total of 154 methods collected from 30 participants. We have categorized the methods based on the metaphor and modality of interaction, and discussed the strengths and weakness of each category based on qualitative and quantitative feedback given by the participants. The results of our study inform the design of future multi-user applications for wearable devices.

## ACKNOWLEDGMENTS

## REFERENCES

1. Yuji Ayatsuka and Jun Rekimoto. 2005. tranSticks: physically manipulatable virtual connections. In *Proc. CHI '05*, 251-260.

2. Jason Alexander, Teng Han, William Judd, Pourang Irani, and Sriram Subramanian. 2012. Putting your best foot forward: investigating real-world mappings for foot-based gestures. In *Proc. CHI '12*, 1229-1238.

3. Jacob T. Biehl and Brian P. Bailey. 2004. ARIS: an interface for application relocation in an interactive space. In *Proc. GI '04*, 107-116.

4. Ming Ki Chong and Hans Gellersen. 2011. How users associate wireless devices. In *Proc. CHI '11*, 1909-1918.

5. Ming Ki Chong and Hans Gellersen. 2013. How groups of users associate wireless devices. In *Proc. CHI '13*, 1559-1568.

6. Ming Ki Chong, Rene Mayrhofer, and Hans Gellersen. 2014. A survey of user interaction for spontaneous device association. *ACM Computing Surveys*. 47, 1, article 8, 40 pages.

7. Jens Grubert, Matthias Heinisch, Aaron Quigley, and Dieter Schmalstieg. 2015. MultiFi: multi fidelity interaction with displays on and around the body. In *Proc. CHI '15*, 3933-3942.

8. Alina Hang, Emanuel von Zezschwitz, Alexander De Luca, and Heinrich Hussman. 2012. Too much information!: user attitudes towards smartphone sharing. In *Proc. NordiCHI '12,* 284-287.

9. Mountaz Hascoët. 2003. Throwing models for large displays. In *Proc. HCI '03*, 73-77.

10. Lars Erik Holmquist, Friedemann Mattern, Bernt Schiele, Petteri Alahuhta, Michael Beigl, and Hans Gellersen. 2001. Smart-Its Friends: A technique for users to easily establish connections between smart artefacts. In *Proc. Ubicomp '01*, 116-122.

11. Iulia Ion, Marc Langheinrich, Ponnurangam Kumaraguru, and Srdjan Čapkun. 2010. Influence of user perception, security needs, and social factors on device pairing method choices. In *Proc. SOUPS '10*, article 6.

12. Tero Jokela, Ming Ki Chong, Andrés Lucero, and Hans Gellersen. 2015. Connecting devices for collaborative interactions. *interactions* 22, 4, 39-43.

13. Tero Jokela and Andrés Lucero. 2013. A comparative evaluation of touch-based methods to bind mobile devices for collaborative interactions. In *Proc. CHI '13*, 3355-3364.

14. Tero Jokela and Andrés Lucero. 2014. FlexiGroups: binding mobile devices for collaborative interactions in medium-sized groups with device touch. In *Proc. MobileHCI '14*, 369-378.

15. Tero Jokela, Jarno Ojala, Guido Grassel, Petri Piippo, and Thomas Olsson. 2015. A comparison of methods to move visual objects between personal mobile devices in different contexts of use. In *Proc. MobileHCI '15*, 172-181.

16. Ronald Kainda, Ivan Flechais, and A. W. Roscoe. 2010. Two heads are better than one: security and usability of device associations in group scenarios. In *Proc. SOUPS '10*, article 5.

17. Christian Kray, Michael Rohs, Jonathan Hook, and Sven Kratz. 2008. Group coordination and negotiation through spatial proximity regions around mobile devices on augmented tabletops. In Proc. *TABLETOP '08*, 1-8.

18. Byron Lahey, Audrey Girouard, Winslow Burleson, and Roel Vertegaal. 2011. PaperPhone: understanding the use of bend gestures in mobile devices with flexible electronic paper displays. In *Proc. CHI '11*, 1303-1312.

19. Rene Mayrhofer and Martyn Welch. 2007. A human-verifiable authentication protocol using visible laser light. In *Proc. ARES '07*, 1143-1148.

20. Miguel A. Nacenta, Dzmitry Aliakseyeu, Sriram Subramanian, and Carl Gutwin. 2005. A comparison of techniques for multi-display reaching. In *Proc. CHI '05*, 371-380.

21. Miguel Nacenta, Carl Gutwin, Dzmitry Aliakseyeu, and Sriram Subramanian. 2009. There and back again: cross-display object movement in multi-display environments. *Human-Computer Interaction* 24, 1-2, 170-229.

22. Michael Nielsen, Moritz Störring, Thomas B. Moeslund, and Erik Granum. 2004. A procedure for developing intuitive and ergonomic gesture interfaces for HCI. In *Gesture-Based Communication in Human-Computer Interaction*, 409-220.

23. Dan Olsen, Jr. and Travis Nielsen. 2001. Laser pointer interaction. In *Proc. CHI '01*, 17-22.

24. Jun Rekimoto. 1997. Pick-and-drop: a direct manipulation technique for multiple computer environments. In *Proc. UIST '97*, 31-39.

25. Jun Rekimoto. 2004. SyncTap: synchronous user operation for spontaneous network connection. *Personal and Ubiquitous Computing* 8, 2 (May 2004), 126-134.

26. Jun Rekimoto, Yuji Ayatsuka, Michimune Kohno, and Hauro Oba. 2003. Proximal interactions: a direct manipulation technique for wireless networking. In *Proc. INTERACT '03*, 511-518.

27. Jaime Ruiz, Yang Li, and Edward Lank. 2011. User-defined motion gestures for mobile interaction. In *Proc. CHI '11*, 197-206.

28. Scott, Besacier, and McClelland. 2014. Cross-device transfer in a collaborative multi-surface environment without user identification. In *Proc. CTS '14*, 219-226.

29. Lucia Terrenghi, Aaron Quigley, and Alan Dix. 2009. A taxonomy for and analysis of multi-person-display ecosystems. *Personal and Ubiquitous Computing* 13, 8, 583-598.

30. Brygg Ullmer, Hiroshi Ishii, and Dylan Glas. 1998. mediaBlocks: physical containers, transports, and controls for online media. In *Proc. SIGGRAPH '98*, 379-386.

31. Jacob O. Wobbrock, Meredith Ringel Morris, and Andrew D. Wilson. 2009. User-defined gestures for surface computing. In *Proc. CHI '09*, 1083-1092.

32. Jacob O. Wobbrock, Htet Htet Aung, Brandon Rothrock, and Brad A. Myers. 2005. Maximizing the guessability of symbolic input. In *CHI EA '05*, 1869-1872.