Designing UbiComp Experiences for Spatial Navigation and Cross-Device Interactions

Doctoral thesis for obtaining the academic degree Doctor of Natural Sciences (Dr. rer. nat.)

submitted by

Roman Rädle

at the

Universität
Konstanz

Faculty of Sciences
Department of Computer and Information Science

Date of the oral examination: January 12th, 2017

1. Reviewer: Prof. Dr. Harald Reiterer
2. Reviewer: Prof. Dr.-ing. Raimund Dachselt
3. Reviewer: Prof. Dr. Marc H. Scholl

Konstanz, April 2017

Designing UbiComp Experiences for Spatial Navigation and Cross-Device Interactions

Dissertation zur Erlangung des akademischen Grades eines Doktors
der Naturwissenschaften (Dr. rer. nat.)

vorgelegt von

Roman Rädle

an der

Universität Konstanz

Mathematisch-Naturwissenschaftliche Sektion
Fachbereich Informatik und Informationswissenschaft

Tag der mündlichen Prüfung: 12. Januar 2017

1. Referent:	Prof. Dr. Harald Reiterer
2. Referent:	Prof. Dr.-ing. Raimund Dachselt
3. Referent:	Prof. Dr. Marc H. Scholl

Konstanz, April 2017

Roman Rädle
Designing UbiComp Experiences for Spatial Navigation and Cross-Device Interactions
Doctoral thesis for obtaining the academic degree Doctor of Natural Sciences (Dr. rer. nat.)
Submission: June 2016
Publication: April 2017
Supervisor: Prof. Dr. Harald Reiterer

University of Konstanz
Human-Computer Interaction Group (HCI Group)
Faculty of Sciences
Department of Computer and Information Science
Universitätsstraße 10
78457 Konstanz

Abstract

We are witnessing a considerable growth in number and density of powerful mobile devices around us. Such devices like smartphones and tablets are our everyday companions. If not already at hand, they often wait in our bags and pockets to provide us with a ubiquitous computing (UbiComp) experience. However, most of these devices are still blind to the presence of other devices and performing tasks among them is usually tedious due to the lack of guiding principles.

This thesis closes with the gap as mentioned above by investigating in the design and evaluation of spatial and cross-device interactions. As a central theme, presented research fundamentally grounds on embodied practices by exploiting users' pre-existing practical knowledge of everyday life for human-computer interaction. These embodied practices are often applied subconsciously in our daily activities, which unfolds new — yet unexplored — potentials for fun and joyful UbiComp experiences.

Within this thesis, research is approached through both deductive and inductive reasoning. It begins with a brief history of UbiComp and its overarching vision. Contradicting opinions on this vision are discussed before leading over to recent theories and believes on embodied cognition and models on human spatial memory. This theoretical background eventually thrives arguments for yet unexplored and hidden potentials for spatial and cross-device interactions. Then, the application domain is narrowed down to academic libraries and knowledge work activities. In field studies at the Library of the University of Konstanz, the following two main knowledge work activities, and resonating issues are identified: *literature & bibliographic search* and *reading & writing across documents*. Together with the theoretical background the found issues are transformed into potentials for future knowledge work. Thereby, two fully functional research prototypes, Blended Shelf and Integrative Workplace, were implemented to explore the problem space further and to derive research questions covered in this thesis.

The research questions are tackled through controlled experiments and implementation of low-cost enabling technology. In the first experiment, the optimal size of a spatially-aware peephole display is studied. As a finding, a relatively small tablet-sized peephole display serves as "sweet-spot" between navigation performance, subjective workload, and user preference. Within the second experiment, peephole navigation is contrasted with traditional multi-touch navigation. The findings indicate that users prefer physical peephole navigation over multi-touch navigation. It also leads to better navigation trajectories, shortens task-completion-time, and hints for longer retention of object identities as well as their locations in human spatial memory. Due to the lack of appropriate technology HuddleLamp was developed in an intermediary step. HuddleLamp is a low-cost sensing technology that tracks multiple mobile devices on a table. It allows implementing spatial and cross-device applications without the need to instrument rooms, equip devices with markers or install additional software on them. This technology is used in the third experiment to understand subtleties of cross-device interactions. Findings show that, for cross-device object-movement tasks, users prefer spatially-aware interactions over spatially-agnostic interactions.

Apart from individual findings, this thesis contributes a summary and integration of all findings to general design guidelines for future spatial and cross-device applications. Eventually, these guidelines are applied by researchers and practitioners to develop UbiComp experiences that increase users' task performance, lower their individual workloads such as mental demand, effort, and frustration. At the same time, these guidelines lead to an increase of the cumulative value when working with multiple mobile devices.

Zusammenfassung

Derzeit erleben wir eine stark zunehmde Präsenz von leistungsstarken mobilen Geräten um uns herum. Geräte wie Smartphones und Tablets sind unsere täglichen Begleiter — wenn wir sie nicht bereits in der Hand halten, dann warten sie oft in unseren Hosen- oder Tragetaschen, um uns überall und jederzeit mit ihrer Rechenleistung zu unterstützen (sog. Ubiquitous Computing oder kurz UbiComp).

Allerdings erkennen sie nur die Präsenz anderer Geräte aber nicht deren genaue Lokation und sind daher sozusagen noch "blind". Somit ist auch das Ausführen von Aufgaben über Gerätengrenzen hinweg in der Regel umständlich was dem Umstand geschuldet ist, dass Richtlinien für die Gestaltung von gerätenübergreifenden Interkationen (sog. cross-device interactions) fehlen.

Diese Arbeit schließt mit der oben erwähnten Lücke und befasst sich mit der Gestaltung von räumlichen und geräteübergreifenden Interaktionstechniken. Als zentrale Themen präsentiert sie Forschung, die einerseits auf *Embodiment*-Praktiken basiert und andererseits, im Rahmen der Mensch-Computer-Interaktion bereits bestehende praktische Kenntnisse des täglichen Lebens ausnutzt. Diese *Embodiment*-Praktiken werden bei der täglichen Arbeit oft unbewusst angewandt und bieten neue — noch unerforschte — Potenziale für UbiComp Erfahrungen, die Spaß und Freude während der Bedienung bereiten sollen.

Im Rahmen dieser Arbeit wird die Forschung sowohl deduktiv als auch induktiv angegangen. Sie beginnt mit einer kurzen Einführung in UbiComp und der damit verbundenen Vision. Widersprüchliche Meinungen zu dieser Vision werden diskutiert, bevor zeitgenössische Theorien, Modelle und Rahmenwerke des *Embodiment* und des menschlichen räumlichen Gedächtnisses eingeführt werden. Dieser theoretische Hintergrund dient schließlich dazu, Argumente für unerforschte und verborgene Potenziale für die räumliche und geräteübergreifende Interaktionen auszuarbeiten. Danach wird die Anwendungsdomäne *Wissenschaftliche Bibliotheken* vorgestellt und auf Wissensarbeit eingegrenzt. In Feldstudien in der Bibliothek der Universität Konstanz wurden die folgenden und wesentlichen zwei Wissensarbeitsaktivitäten und deren damit einhergehenden Probleme identifiziert: *Literatur & bibliographische Suche* und *Lesen & Schreiben über mehrere Dokumente*. Zusammen mit dem theoretischen Hintergrund werden die gefundenen Probleme in Potentiale für zukünftige Wissensarbeit umgewandelt. Resultierend daraus wurden zwei voll funktionsfähige Forschungsprototypen, *Blended Shelf* und *Integrative Workplace* entwickelt, um den Problembereich weiter zu erforschen und, die in dieser Arbeit abgedeckten, Forschungsfragen abzuleiten.

Die Forschungsfragen werden durch kontrollierte Experimente und durch die Entwicklung von kostengünstigen Technologien adressiert. Im ersten Experiment wurde die optimale Größe für eine *Peephole*-Interaktion untersucht. Das Ergebnis zeigt, dass bereits ein relativ kleines Tablet-Gerät als "sweet-spot" für die räumliche *Peephole*-Interaktion eingesetzt werden kann, um eine gute Navigationsleistung, eine geringe Arbeitsbelastung und eine gute Benutzerpräferenz zu erreichen. Im zweiten Experiment wird die *Peephole*-Navigation der traditionellen *Multi-Touch*-Navigation gegenübergestellt. Die Ergebnisse zeigen, dass die Nutzer die egozentrische Peephole-Navigation der *Multi-Touch*-Navigation vorziehen. Des Weiteren führt die *Peephole*-Interaktion zu besseren Navigationspfaden und verkürzt die Bearbeitungszeit der gestellten Aufgabe. Zudem weist die Peephole-Interaktion auf eine verbesserte Unterstützung des räumlichen Gedächtnisses auf. Aufgrund des Fehlens geeigneter Technologien für die räumliche und geräteübergreifende Interaktion wurde HuddleLamp in einem Zwischenschritt entwickelt. HuddleLamp ist eine kostengünstige *Sensing*-Technologie, welche die lokalen Relationen mehrerer mobiler Geräte auf einem Tisch verfolgen kann. Dadurch ermöglicht sie die Entwicklung von räumlichen und geräteübergreifenden Interaktionen mit mobilen Geräten. Diese Technologie wird in einem dritten Experiment verwendet, um geräteübergreifende Interaktionen im Detail zu untersuchen. Die Ergebnisse zeigen, dass für Aufgaben, bei denen Objekte über Gerätegrenzen versendet werden, die Anwender räumliche Interaktion den räumlich-agnostischen Interaktionen vorziehen.

Abschließend werden die einzelnen Erkenntnissen zusammengefasst und in allgemeingültige Gestaltungsrichtlinien für zukünftige räumliche und geräteübergreifende Anwendungen überführt. Diese Richtlinien können von Forschern und Praktikern angewendet werden, um neue *UbiComp* Interkationsformen und Nutzererfahrungen zu entwickeln. Diese Interaktionsformen sollen letztlich die Frustration des Nutzers senken und gleichzeitig seine Leistungsfähigkeit steigern.

Acknowledgement

This dissertation will always represent a memorable episode in my life. At the beginning of my doctorate studies, I knew it was going to affect my professional life, but I was not aware of how much it would affect my personal life as well. I have learned that it is more than a job or yet another degree; it is a voyage with countless expected but also unexpected adventures. Fortunately, I had many companions and comrades that joined me during these adventures. They made them fun and joyful experiences, and the not so pleasant ones bearable – a trouble shared is a trouble halved. I would like to thank and honor my mentors, colleagues, collaborators, friends, and family who have guided, supported, and inspired me along this fulfilling journey. Retrospectively, this acknowledgments section of my dissertation was the most difficult to write. There are many people who I would like to thank and honor, including those that I may have forgotten to mention. This episode in my life would have never been as joyful and fun without their support.

First, I would like to express my endless gratitude to my mentor and advisor, *Harald Reiterer*. Our paths had crossed multiple times before I started my doctorate studies with him; in fact, they had crossed even before I started my undergraduate degree program at the University of Konstanz. In 2003, several universities accepted my application, and I was in the fortunate position to choose whether I wanted to become a student in Technical Informatics, Computer Science, or Information Engineering. At the time, my brother was enrolled as a student at the University of Konstanz in the Department of Economics with a minor in Computer Science. To inform my decision, I decided to visit the University of Konstanz and join my brother for one "trial lecture," which happened to be a course in Mensch-Computer-Interaktion (Human-Computer Interaction) lectured by Harald. His lecturing style and the lively participation of the students in his class convinced me to enroll at the University of Konstanz.

Most importantly, at the beginning of my master's studies, our paths crossed again when I became a student researcher in Harald's group. The close collaborations with him and members of his group on several research projects allowed me to gain first-hand experience in what research entails. As I approached the end of my master's program, I knew that there was still more that I could learn from him and his group. Thus, I was very honored and did not hesitate a second when he offered me an opportunity to continue working with him as a doctoral student. Harald supervises differently than is typically the norm. He gives his students great freedom to develop their own research agenda. His door is always open for his students – including bachelor's and master's students – to ask for help, guidance, or his opinion. Once in his office, he quickly externalizes thoughts on a flipchart, which is one of his many skills. Without his guidance and persistent help, this dissertation would not have been possible. As a part of his group, I was able to perfect my research and analytical skills. I hope our paths will cross again in the future and as many times as possible – not only in December at La Bodega. Thank you for your trust and support all these years.

Also, I would like to thank my committee members, *Raimund Dachselt* and *Marc H. Scholl* for their encouraging and insightful comments, meticulous corrections, and their rigorous questions that prompted me to more extensively develop my ideas.

It is challenging to fully express my appreciation and thankfulness to my colleague *Hans-Christian Jetter*. While still being a Ph.D. student in Harald's group, he slipped into the role of a postdoc guiding my research. He taught me the *"What? Why? How?"* and I learned a lot about academia from him – the undocumented facts and stories. We spent endless nights in the lab together working towards quickly approaching deadlines, with only Curry King et al. and Red Bull fueling our academic spirit (pun intended). After each deadline and many nights with very little sleep, we agreed to not do it again, just to find the both of us in the same situation only months later. He also was my personal guide at CHI '12 in Austin, TX, introducing me to this big international circus. Throughout all these years of collaboration and guidance, our collegial relationship turned into a friendship. Christian is an exceptional researcher who has an honest and remarkable personality, and I am always looking forward to our next adventure.

My dearest appreciation goes to my colleagues and lab mates at the HCI Group at the University of Konstanz, particularly *Jens 'Dude' Müller* and *Simon Butscher*. They stood at my side during tough times – my professional Armageddon – and provided mental and technical support. Jens is a wonderful friend who is always willing to contribute different perspectives and solutions to a problem. He sometimes allowed me to practice my motivational skills on him – and I like to believe with success. Simon always lent a friendly ear when I consulted him to check on statistical methods that I had applied. Thank you to *Ulrike Pfeil* for the many whiteboard discussions to brew the structure of this dissertation. She enriches the group with her positive presence and contagious happiness. *Johannes Zagermann* and I had exciting times together, especially during the TwisterSearch studies. Thank you for joining me during weekends of work in the lab. Thanks to my officemate *Daniel Klinkhammer*, who shared my same fate when we were both writing up our dissertations. Thank you to *Marianne Kirn-Reutter* for her dedicated support to the group and for keeping an eye on our contracts.

Many students worked with me on the Blended Library project and our vision of the library of the future, many of them with whom I become friends. For example, *Stephan Huber*, who always infused the group with his happy spirit. He is an organizational talent and working with him was always fun and inspirational. In my opinion, every team needs a Stephan. *Christoph Gebhardt* is forward thinking, conscientious, and gets things done. Once you give him a task, he has most likely done it already. He is very meticulous in what he is doing, and I wish him all the best in his pursuit of a Ph.D. I thank *Michael Zöllner* and *Mario Schreiner* for their valuable technical discussions and their support when we were running out of time, but writing code was desperately needed. Thank you to *Eike Kleiner* and *Tobias Baube* and their support on the Blended Library project. *Sebastian Marwecki* and his interest in "hybrid serious games" enriched my research spectrum, and it was always thrilling to conduct studies at Mutpol[1]. Thanks to the Japanese twin brothers *Stefan Feyer* and *Leonard Krämer* for entertainment during dinner times at the Menseria Gießberg. Luckily, I could distinguish the two – one had curly hair.

I thank *Werner A. König* who, together with Harald, convinced me about the value of pursuing research. Because of this and time constraints, I had to quit a well-paying industry contract at Wilken GmbH. In the long run, it turned out to be an excellent decision to agree with Werner's arguments and choose research over money. Without him, I probably would have a 9-5 job today and not even be cognizant of how many amazing things I would have missed.

While I worked for Wilken GmbH, *Jürgen Denzel* was one of the most experienced programmers, with a profound knowledge of Java and especially J2EE programming. Throughout my time at Wilken GmbH, I often consulted Jürgen and asked for advice when I got stuck. He always provided solutions or alternative approaches to tackling my current problems – Jürgen is better than Stack Overflow. Thus, I am very thankful to him for teaching me different ways of programming and his patience when others already would have gone into a rage.

In 2013, *Katherine Isbister* provided me with the great opportunity to join the Game Innovation Lab (GIL) at New York University, for which I am very thankful. I was very heartily welcomed by GIL members *Holly Robbins*, *Kaho Abe*, *Ming Jin*, *Edward Melcer*, *Bert Buchholz*, *Chris DiMauro*, and especially *Chrystanyaa Brown*. In addition to Chrystanyaa's duty to keep everything clean(est) in the lab, she was the moral

---

[1] Mutpol is a non-profit organization that works with children who are affected by autism spectrum disorders – `http://www.mutpol.de/` (last accessed: April 17th, 2017).

center of the lab. Another big thank you goes to my colleague *Hendrik Strobelt* for connecting me with Katherine and Chrystanyaa, for his mental support in tough times, and his friendship.

My sincere thanks also go to the Human Experience & Design group at Microsoft Research in Cambridge, UK, especially to *Abigail Sellen* for offering me the summer internship opportunities and to *Kenton O'Hara* and *Sean Rintel* for supervising me working on an exciting project. I would also like to thank *Debaleena Chattopadhyay*, *Stephanie Neumann*, *Tobias Grosse-Puppendahl*, *Marco Fraccaro*, and *Filip Šrajer* for teaching me how to punt, traveling with me to Bletchley Park and Dismaland, joining me on various coffee breaks, and beating me in table tennis.

I gratefully acknowledge the financial support received from the *German Research Foundation* (DFG) and the *Ministry of Research, Science, and Art in Baden-Württemberg* (for the Blended Library Project). I also appreciated being an associated Ph.D. student to the *GK-1042* (Ph.D. Graduate Program) “Explorative Analysis and Visualization of Large Information Spaces.”

Thanks to *Ricardo Langner* for sharing this LaTeX thesis template, *Nicolai Marquardt* for his support with the ProximityToolkit and his guidance on integrating OptiTrack as another input module, and *Ashley Karr* for being a fellow companion. I gratefully acknowledge her inspiring and thought-provoking perspectives, and the mental support when work was tough.

Luckily, I did not start this adventure alone, but with my former and fellow students *Andreas Weiler* and *Johannes Fuchs*. We started our Bachelor's degree program together, found ourselves together again during the master's program, and finally started our doctorate studies together. Albeit in different groups, we frequently met for coffee breaks to discuss our research. Aside from work, we also hung out together with our close friends *Hannes*, *Esther*, *Sebastian*, and *Stefanie* to enjoy life.

I would like to thank UL45 for providing me with shelter during the last weeks of writing. At an advanced age, I never thought I could go back to living in a shared apartment, but they showed me that it can be fun. Thank you to *Martin*, *Heike*, *Naza*, *Britta*, *Saskia*, and *Mirjam*.

My deepest appreciation and gratitude goes, however, to my family. They have always believed in me, supported my studies from the beginning, and provided me with the moral and emotional support needed through the entire process. This would have never been possible without them and their constant encouragement throughout all my journeys. I will be forever grateful for their unconditional love and support. In 2012, we had a tragic loss, which is why I wholeheartedly dedicate this dissertation to my father *Anton Rädle*, who continually showed passion in his works. He never gave up and taught me how much one could achieve with diligence and passion by enacting it every day of his life.

Publications

Parts of this research were previously published in the following publications. Reused material is indicated in the beginning of each chapter where applicable.

Journal Publications

Reiterer, H. **Rädle, R.** Butscher, S. Müller, J. (2016). “Blended Library – neue Zugangswege zu den Inhalten wissenschaftlicher und öffentlicher Bibliotheken”. In: *Bibliothek Forschung und Praxis* 40.1. DOI: `10.1515/bfp-2016-0010`

Gebhardt, C. **Rädle, R.** Reiterer, H. (2014a). “Employing Blended Interaction to Blend the Qualities of Digital and Physical Books”. en. In: *i-com* 13.3, pp. 35–44. DOI: `10.1515/icom.2014.0028`

Conference Publications (Peer Reviewed)

Zagermann, J. Pfeil, U. **Rädle, R.** Jetter, H.-C. Klokmose, C. Reiterer, H. (2016). “When Tablets meet Tabletops: The Effect of Tabletop Size on Around-the-Table Collaboration with Personal Tablets”. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI ’16*. New York, New York, USA: ACM Press, pp. 5470–5481. DOI: `10.1145/2858036.2858224`

Müller, J. **Rädle, R.** Reiterer, H. (2016). “Virtual Objects as Spatial Cues in Collaborative Mixed Reality Environments: How They Shape Communication Behavior and User Task Load”. In: *Proceedings of the 34th Annual ACM Conference on Human Factors in Computing Systems - CHI ’16*. New York, NY, USA: ACM Press

Chattopadhyay, D. O’Hara, K. Rintel, S. **Rädle, R.** (2016). “Office Social: Presentation Interactivity for Nearby Devices”. In: *Proceedings of the 34th Annual ACM Conference on Human Factors in Computing Systems - CHI ’16*. New York, NY, USA: ACM Press

Schubert, M. Buder, J. **Rädle, R.** Hesse, F. W. (2015). “Common ground and individual accountability in literature selection of groups: three different group learning techniques”. en. In: *Proceeding of the 11th International Conference on Computer Supported Collaborative Learning - CSCL ’15*. Gothenburg: The International Society of the Learning Sciences. DOI: `10.15496/publikation-10650`

Rädle, R. Jetter, H.-C. Schreiner, M. Lu, Z. Reiterer, H. Rogers, Y. (2015). “Spatially-aware or Spatially-agnostic?: Elicitation and Evaluation of User-Defined Cross-Device Interactions”. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI ’15*. New York, New York, USA: ACM Press, pp. 3913–3922. DOI: `10.1145/2702123.2702287`

Müller, J. **Rädle, R.** Jetter, H.-C. Reiterer, H. (2015). “An Experimental Comparison of Vertical and Horizontal Dynamic Peephole Navigation”. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI ’15*. New York, New York, USA: ACM Press, pp. 1523–1526. DOI: `10.1145/2702123.2702227`

Rädle, R. Jetter, H.-C. Marquardt, N. Reiterer, H. Rogers, Y. (2014c). “HuddleLamp: Spatially-Aware Mobile Displays for Ad-hoc Around-the-Table Collaboration”. In: *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces - ITS ’14*. New York, New York, USA: ACM Press, pp. 45–54. DOI: `10.1145/2669485.2669500`

Accompanying Demo presented at ITS ’14 received People’s Choice Best Demo Award

Gebhardt, C. **Rädle, R.** Reiterer, H. (2014b). “Employing Blended Interaction to Blend the Qualities of Digital and Physical Books”. en. In: *Mensch & Computer 2014 - Tagungsband*. Ed. by Andreas Butz, Michael Koch, and Johann Schlichter. De Gruyter Oldenbourg, pp. 35–44

Received Best Paper Award

Rädle, R. Jetter, H.-C. Müller, J. Reiterer, H. (2014a). “Bigger is not always better: display size, performance, and task load during peephole map navigation”. In: *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI ’14*. New York, New York, USA: ACM Press, pp. 4127–4136. DOI: `10.1145/2556288.2557071`

Fuchs, J. **Rädle, R.** Sacha, D. Fischer, F. Stoffel, A. (2014). “Collaborative data analysis with smart tangible devices”. In: *IS&T/SPIE Electronic Imaging*. Ed. by Pak Chung Wong, David L. Kao, Ming C. Hao, and Chaomei Chen. International Society for Optics and Photonics. DOI: `10.1117/12.2040011`

Rädle, R. Jetter, H.-C. Butscher, S. Reiterer, H. (2013a). “The effect of egocentric body movements on users’ navigation performance and spatial memory in zoomable user interfaces”. In: *Proceedings of the 2013 ACM international conference on Interactive tabletops and surfaces - ITS ’13*. New York, New York, USA: ACM Press, pp. 23–32. DOI: `10.1145/2512349.2512811`

Doctoral Consortium

Rädle, R. (2013). “Design and evaluation of proxemics-aware environments to support navigation in large information spaces”. In: *CHI ’13 Extended Abstracts on Human Factors in Computing Systems on - CHI EA ’13*. New York, NY, USA: ACM, p. 1957. DOI: `10.1145/2468356.2468710`

Book Chapters

Rädle, R. Jetter, H.-C. Reiterer, H. (2013b). “TwisterSearch: A distributed user interface for collaborative Web search”. In: *Distributed User Interfaces: Collaboration and Usability*. Ed. by María D. Lozano, Jose A. Gallud, Ricardo Tesoriero, and Victor M.R. Penichet. London, UK: Springer London. Chap. 5, pp. 4–7. DOI: `10.1007/978-1-4471-5499-0`

Heilig, M. **Rädle, R.** Reiterer, H. (2011). “Die Blended Library: Benutzerorientierte Verschmelzung von virtuellen und realen Bibliotheksdiensten”. In: *Benutzerorientierte Bibliotheken im Web*. Ed. by Bernard Bekavac, René Schneider, and Werner Schweibenz. De Gruyter

Posters

Schreiner, M. **Rädle, R.** Jetter, H.-C. Reiterer, H. (2015a). “Connichiwa: A Framework for Cross-Device Web Applications”. In: *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA ’15*. New York, New York, USA: ACM Press, pp. 2163–2168. DOI: `10.1145/2702613.2732909`

Gebhardt, C. **Rädle, R.** Reiterer, H. (2014c). “Integrative workplace: studying the effect of digital desks on users’ working practices”. In: *Proceedings of the extended abstracts of the 32nd annual ACM conference on Human factors in computing systems - CHI EA ’14*. New York, New York, USA: ACM Press, pp. 2155–2160. DOI: `10.1145/2559206.2581186`

Marwecki, S. **Rädle, R.** Reiterer, H. (2013). “Encouraging collaboration in hybrid therapy games for autistic children”. In: *CHI ’13 Extended Abstracts on Human Factors in Computing Systems on - CHI EA ’13*. New York, New York, USA: ACM Press, p. 469. DOI: `10.1145/2468356.2468439`

Kleiner, E. **Rädle, R.** Reiterer, H. (2013). “Blended Shelf: Reality-based Presentation and Exploration of Library Collections”. In: *CHI ’13 Extended Abstracts on Human Factors in Computing Systems on - CHI EA ’13*. New York, New York, USA: ACM Press, p. 577. DOI: `10.1145/2468356.2468458`

Rädle, R. Zöllner, M. Sulger, S. (2011a). “eXLEpse: An Eclipse-based, Easy-to-Use Editor for Computational LFG Grammars”. In: *Proceedings of the LFG11 Conference*. Ed. by Miriam Butt and Tracy Hollaway King. Hong Kong, China

Demos

Rädle, R. Jetter, H.-C. Marquardt, N. Reiterer, H. Rogers, Y. (2014b). “Demonstrating HuddleLamp: Spatially- Aware Mobile Displays for Ad-hoc Around-the-Table Collaboration”. In: *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces - ITS ’14*. New York, New York, USA: ACM Press, pp. 435–438. DOI: `10.1145/2669485.2676584`

Received People’s Choice Best Demo Award

Rädle, R. Jetter, H.-C. Reiterer, H. (2012d). “TwisterSearch: Supporting Social Search with Tabletop and Mobile Displays”. In: *Mensch & Computer 2012 – Workshopband: interaktiv informiert – allgegenwärtig und allumfassend!?* Ed. by Harald Reiterer and Oliver Deussen. 2009. München: Oldenbourg Verlag, pp. 509–512

Conference & Workshop Papers, & Reports w/o Proceedings

Rintel, S. O’Hara, K. Rostami Yaganeh, B. **Rädle, R.** (2015). “Ad hoc adaptability in video-calling”. In: *1st International Workshop on Interacting with Multi-Device Ecologies in the Wild (in conjunction with ITS ’15)*

Schreiner, M. **Rädle, R.** O’Hara, K. Reiterer, H. (2015b). “Deployable Cross- Device Experiences: Proposing Additional Web Standards”. In: *1st International Workshop on Interacting with Multi-Device Ecologies in the Wild (in conjunction with ITS ’15)*

Zagermann, J. Pfeil, U. Schreiner, M. **Rädle, R.** Jetter, H.-C. Reiterer, H. (2015). "Reporting Experiences on Group Activities in Cross-Device Settings". In: *International Workshop on Cross-Surface 2015 (in conjunction with ITS '15)*. Madeira, Spain

Marwecki, S. **Rädle, R.** Reiterer, H. "Enhancing Motivation in Hybrid Therapy Games for ASD". in: *International Workshop on Supporting Children with Complex Communication Needs (in conjunction with CHI '14).*

Greenberg, S. Honbaek, K. Quigley, A. Reiterer, H. **Rädle, R.** (2014b). *Proxemics in Human-Computer Interaction (Dagstuhl Seminar 13452)*. eng. Tech. rep. 11, p. 57. DOI: `10.4230/DagRep.3.11.29`

Jetter, H.-C. **Rädle, R.** (2013). “Visual and Functional Adaptation in Ad-hoc Communities of Devices”. In: *International Workshop on Visual Adaptation of Interfaces (in conjunction with ITS ’13)*. St Andrews, Scotland, UK

Butscher, S. Müller, J. Weiler, A. **Rädle, R.** Reiterer, H. Scholl, M. H. (2013). “Multi-user Twitter Analysis for Crisis Room Environments”. In: *Collaborative Human-Computer-Interaction with Big Wall Displays (BigWallHCI 2013), 3rd JRC ECML Crisis Management Technology Workshop*. Ed. by Markus Rester, Peter Judmaier, Tom De Groeve, and Alessandro Annunziato. Luxembourg: Publications Office of the European Union, pp. 28–34

Jetter, H.-C. Schöning, J. **Rädle, R.** Reiterer, H. Rogers, Y. (2013). “Collaborative Interactions in Future Crisis Rooms”. In: *Collaborative Human-Computer-Interaction with Big Wall Displays (BigWallHCI 2013), 3rd JRC ECML Crisis Management Technology Workshop*. Ed. by Markus Rester, Peter Judmaier, Tom De Groeve, and Alessandro Annunziato. Luxembourg: Publications Office of the European Union, pp. 35–39

Rädle, R. Weiler, A. Huber, S. Jetter, H.-C. Mansmann, S. Reiterer, H. (2012a). “eBook meets tabletop: using collaborative visualization for search and serendipity in on-line book repositories”. In: *Proceedings of the fifth ACM workshop on Research advances in large digital book repositories and complementary media - BooksOnline ’12*. New York, New York, USA: ACM Press, p. 3. DOI: `10.1145/2390116.2390120`

Rädle, R. Butscher, S. Huber, S. Reiterer, H. (2012b). “Navigation Concepts for ZUIs Using Proxemic Interactions”. In: *Proxemics ’12 (a workshop in conjunction with NordiCHI ’12)*. Konstanz

Rädle, R. Jetter, H.-C. Reiterer, H. (2012c). “TwisterSearch: A distributed user interface for collaborative Web search”. In: *Proceedings of the 2nd Workshop on Distributed User Interfaces: Collaboration and Usability (in conjunction with CHI ’12)*. Ed. by Ricardo Tesoriero, José A. Gallud, María D. Lozano, Víctor M. R. Penichet, and Jean Vanderdonckt. Spain: University of Castilla-La Mancha, pp. 1–4

Rädle, R. Heilig, M. Reiterer, H. (2011b). “Interactive Reading: Serendipity in the Context of the Blended Library”. In: *Proceedings of 1st Workshop on Encouraging Serendipity in Interactive Systems, 13th IFIP TC13 Conference on Human-Computer Interaction*. Ed. by Stephann Makri, Elaine .G. Toms, Lori McCay-Peet, and Ann Blandford. Springer Verlag, pp. 24–26

Contributing Theses

The following Master's thesis that I supervised contributed to the technical implementation and the user study of the Blended Shelf (see Section 3.3.2 for details, page 79):

Kleiner, E. (2013). “Blended Shelf: Ein realitätsbasierter Ansatz zur Präsentation und Exploration von Bibliotheksbeständen”. Master Thesis. University of Konstanz

The following Master's thesis that I supervised contributed to the technical implementation and the user study of the Integrative Workplace (see Section 3.3.3 for details, page 85):

Gebhardt, C. (2013). “Integrative Workplace: Employing Reality-based Interaction to join Digital and Analog Media at a Workplace”. Master Thesis. University of Konstanz

The following Master's thesis that I supervised contributed to the technical implementation of the study prototype used to study the effect of egocentric body movements on users' spatial memory (see Chapter 5 for details, page 119):

Huber, S. (2013). “Entwicklung und Untersuchung eines räumlichen Interaktionskonzepts für die Navigation in zoombaren Benutzerschnittstellen”. Master Thesis. University of Konstanz

Contents

Introduction

1

> *Neither an explication of the principles of ubiquitous computing nor a list of the technologies involved really gives a sense of what it would be like to live in a world full of invisible widgets. Extrapolating from today's (1991 – Ed. Note) rudimentary fragments of embodied virtuality is like trying to predict the publication of Finnegans Wake shortly after having inscribed the first clay tablets.*
>
> — **Mark Weiser**
> (Father of Ubiquitous Computing)

Contents

We are witnessing a considerable growth in number and density of powerful mobile devices around us. Such devices like smartphones and tablets are our everyday companions, which we carry around with us in our bags and pockets. If not already at hand, they often wait there to be unlocked to provide us with a ubiquitous computing (UbiComp) experience. With them, we browse the world wide web, create and edit documents, take photos with integrated cameras, chat and email with colleagues and friends, and access maps and other location-based services. However, their vast majority are still blind to the presence of other devices and performing tasks among them is usually tedious due to the lack of guiding principles.

This thesis closes with this gap by investigating in the design and evaluation of spatial and cross-device interactions. As a central theme, presented research fundamentally grounds on embodied practices by exploiting users' pre-existing practical knowledge of everyday life for human-computer interaction (HCI). These embodied practices are often applied subconsciously in our daily activities, which unfolds new — yet unexplored — potentials for fun and joyful UbiComp experiences. Presumably, such hidden potentials improve knowledge work activities, specifically academic work in libraries. Thereby, they ideally increase the cumulative value when working with multiple mobile devices.

In the remainder of this chapter, I introduce the problem space and describe the driving motivation. Then, I deconstruct the thesis title and depict challenges for spatial navigation and cross-device interaction with mobile devices. Next, I point out overall research questions and describe my research approach and methods applied to answer these research questions. Finally, I give a brief outline of my thesis contributions and end with a brief overview of the dissertation structure and an overview of consecutive chapters of this thesis.

1.1 Motivation

According to Harper et al. and one recent view on UbiComp, we are nearly at the end of the mobile computing era with "several computers per user" (Harper et al., 2008) and are soon to transition to the ubiquitous computing era with "thousands of computers per user" (Harper et al., 2008) (see Figure 1.1). Surprisingly, with recent technological advancements like holographic displays (e.g., Microsoft HoloLens[1] or MagicLeap[2]), their anticipation was right, and the ubiquity era is waiting in the wings. Still, they are still in their infancy with developer units just getting delivered (March 30th, 2016) to a few pre-selected developers only.

Beyond these promising technologies, interaction with personal computers and laptops is still stuck in the 20th century. Interface paradigms have not changed since the Graphical User Interface (GUI) (e.g., Xerox Star or Apple Lisa), and the Windows, Icons, Menus, and Pointer (WIMP) (Dam, 1997) paradigms were invented. Also, interaction with mobile devices, such as smartphones and tablets, is not far from WIMP interaction where touch replaced mouses and apps replaced windows. Of course, such user interface paradigms have been proven to work dependably

---

[1] Official Website of Microsoft HoloLens – `https://www.microsoft.com/microsoft-hololens/en-us` (last accessed: July 13th, 2015)

[2] MagicLeap promises mixed-reality sensations similar to HoloLens. However, a specific device has not yet been presented. More information on their website: `http://www.magicleap.com` (last accessed: March 30th, 2016)

good. However, only if users work on a single computing device. Interaction across multiple devices, however, is a tedious task and lacks tool support due to guiding principles (Oulasvirta and Sumari, 2007; Jetter, 2013; Hamilton and Wigdor, 2014; Yang and Wigdor, 2014). In line with this argument, Klokmose and Beaudouin-Lafon "question the adequacy of the current predominant user interface paradigm, the application-based WIMP interaction model, [. . .] for building user interfaces going beyond a single desktop computer."

For this reason, recent believes in HCI call researchers to think "beyond the assumption that users only employ a single, personal computer and actively be aware of coordinate with a user's other devices" (Dearman and Pierce, 2008). Just like in Sal's day in the vision of UbiComp explicated by Weiser (Weiser, 1999). UbiComp devices support Sal throughout her day seemingly anticipating her every wish. Rest assured, Sal's day did not include mouse and keyboard and thus calls for a paradigm shift from WIMP to post-WIMP user interfaces (Wigdor and Wixon, 2011; Jetter et al., 2014). Alternatively, even a more radical paradigm shift to instrumental interaction (Beaudouin-Lafon, 2000; Klokmose, 2007; Klokmose and Beaudouin-Lafon, 2009) where interaction between "users and domain objects is mediated by interaction instruments, similar to the tools and instruments we use in the real world to interact with physical objects" (Beaudouin-Lafon, 2000).

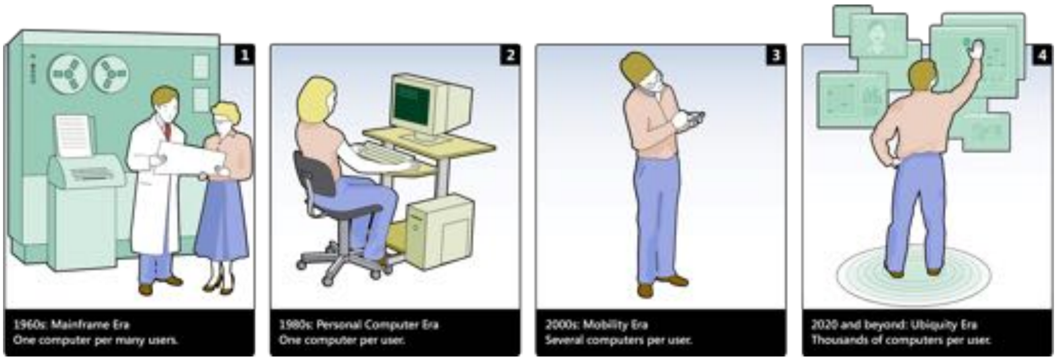


Figure 1.1.: An illustration of computing history with past decades of mainframe, personal, and mobile computing and the anticipation of the upcoming era of ubiquitous computing in the year 2020 (Harper et al., 2008).

However, a paradigm shift comes along with challenges, which might unfold new potentials but also bear risks. To identify potentials and risks, research in this thesis is operationalized by narrowing the application domain down to libraries, specifically academic libraries and therein related knowledge work activities. More precisely, it concentrates on two common but essential knowledge work activities: *literature & bibliographic search* and *reading & writing across multiple documents*. Both knowledge work activities mentioned above relate to the thesis title and will be connected in *Chapter 3 – Context & Analysis*.

1.2 Research Objectives

As a seed point, I start from a general perspective by introducing UbiComp challenges and briefly motivate them from an HCI perspective. The challenges lead to two research objectives (RO) (see Figure 1.2, page 4), which are *Spatial Navigation* (RO1) and *Cross-Device Interaction* (RO2). These two research objectives were selected because they cover a spectrum from interaction with a single mobile device up to interaction with multiple mobile devices.

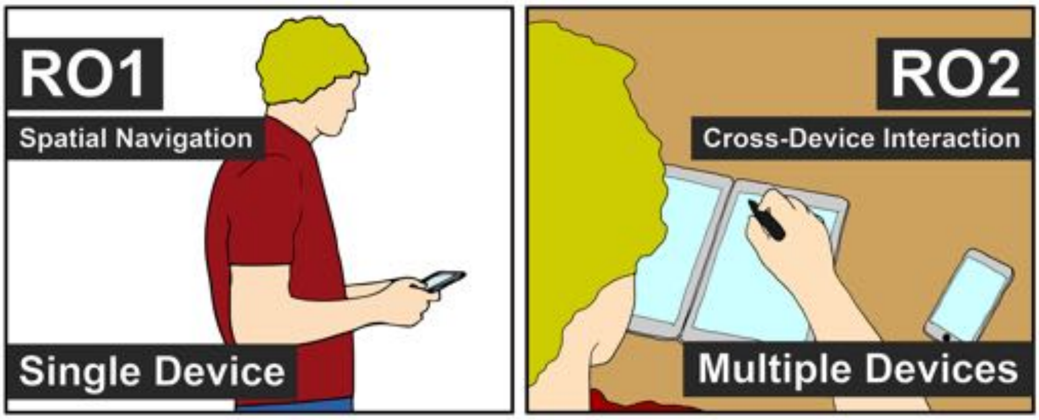


Figure 1.2.: Research objectives tackled in this thesis: *Spatial Navigation* (RO1) and *Cross-Device Interaction* (RO2). These ROs range from interaction with a single mobile device to interaction with multiple mobile devices.

A general goal of this thesis is to find — yet hidden — potentials of UbiComp. For example, by relying on emerging theories, models, and frameworks in HCI that propose to build on users' pre-existing knowledge of the world. Building on existing knowledge eventually saves users' cognitive resources. Thereby, they can focus on primary "application level tasks (e.g., reading, annotating, search)" right away rather than spending the time to and being distracted by secondary "system level tasks (e.g., view management)" (Andrews et al., 2010). Presumably, this lowers their mental effort and decreases frustration while at the same time increases experience and the cumulative value of a system.

The following sections elaborate on the research objectives by means of everyday computing tasks. For each task, I will identify co-occurring problems where users currently waste precious time on such secondary system level tasks. I also present limitations of current approaches and name potentials for improvement. Each section will end with a prospect of thesis chapters tackling particular problems. Both research objectives are then empirically validated in the context of knowledge work in academic libraries. The findings will be reported in *Chapter 3.2.4 – Empirical Findings*.

1.2.1 Spatial Navigation

With the proliferation of the graphical user interface and WIMP interface paradigms, 2D, 2.5D, and 3D navigation in virtual spaces became a subtask in everyday work. For example, scrolling down a website to reveal hidden content or zooming into a digital photograph to increase details. However, virtual spaces are still decoupled from the real physical world, which means that physical movements (e.g., of an input device) undergoes a functional transformation and transcodes to movements in virtual space. For instance, a physical movement of two fingers sensed by a horizontally laid out laptop trackpad transcodes to scrolling content displayed on a vertically positioned laptop screen. As a matter of fact, writing this thesis employs navigation as well, e.g., when frequently switching between reading and writing or to find chapters and sections when referencing them in the text. Such navigation is different from everyday navigation in the physical world.

Historically, WIMP and the graphical implementation of the desktop metaphor became the de facto replacement for command line interfaces and their conversational metaphor. Therein, drag and drop allows users to *directly* manipulate virtual objects and *engage* in a visual conversation with the computer (Hutchins et al., 1985). This conversation and engagement are based on mundane principles or basic-level concepts that users transfer from experiences in real life (e.g., containment of files within folders). Jörn Hurtienne, for instance, calls these basic-level concepts image schemas (Jörn Hurtienne, 2007) and Fauconnier and Turner, simply speaking, refers to them as blends. *Chapter 2.2.4 – Blended Interaction* will go into more detail of blends and the underlying Conceptual Blending theory by (Fauconnier and Turner, 2003). It will also describe how this theory is applied to HCI by Jetter et al. in their Blended Interaction framework (Jetter et al., 2014).

Beyond a more engaging user experience, WIMP interfaces allow their users to exploit so far unutilized capabilities such as spatial memory (Robertson et al., 1998; Tan et al., 2002; Andrews et al., 2010; Scarr et al., 2013). However, still, users are restricted to navigate in inherently abstract virtual spaces. This navigation is different from well-known physical navigation and therefore feels less direct than navigation in physical space.

Physical Navigation in Front of Wall-Sized Screens

Since the advent of interactive spaces with large screens (see Figure 1.3, page 6)[3], navigation is no longer constrained to moving in virtual space alone. The benefit of large interactive displays is that they blend the advantages of the digital world with that of familiar physical navigation (Andrews et al., 2010; Tan et al., 2006; Ball et al., 2007; Ball and North, 2008). For example, they enable users to navigate the information space from an egocentric perspective by glancing around or walking in front of the large display so that view management becomes an entirely familiar physical activity that often happens beneath the conscious awareness of the user.

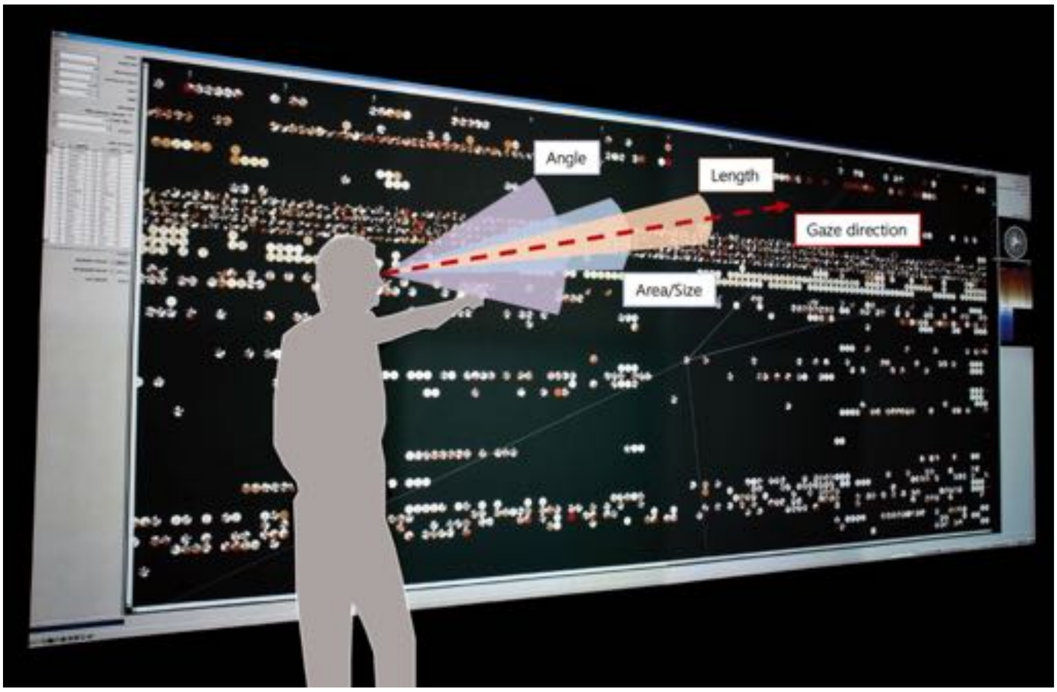


Figure 1.3.: In interactive spaces users can freely move in front of a large screen thus physically navigate in virtual information spaces. The number of perceived details, however, is dependent on their current position and viewing angle (Bezerianos and Isenberg, 2012). (Picture at Powerwall University of Konstanz)

Limitations & Potentials

However, the users' visual acuity is limited, which restricts the number of details that they can perceive on a large display depending on their current position, viewing angle (Bezerianos and Isenberg, 2012), and the display's resolution. One solution to this problem is to provide users with an additional personal mobile screen, e.g.,

---

[3] Picture at the Powerwall University of Konstanz. Changes made by thesis author. – `http://www.vis.uni-konstanz.de/en/powerwall/` (last accessed: February 2nd, 2015)

a tablet or smartphone (Zadow et al., 2014; Dachselt, 2014). Typically, a tablet is held at an optimal distance from one's eyes so that it can provide users with a view of a section of the content of a larger virtual information space and it enables natural manipulation and annotation of items with the tablet's touch or stylus input. Despite these advantages, again there is the need to reintroduce navigation techniques for view management (i.e., zooming and panning) that potentially limit the positive effect of the physical navigation in front of a large display. Eventually, interactive spaces with large wall-sized screens and mobile devices can leverage humans' acquired skills and competencies from interaction in the real physical world to also navigate in virtual spaces.

Prospect: In *Chapter 3.3.2 – Research Prototype 1: Blended Shelf*, I motivate egocentric spatial navigation in a library scenario to search and browse virtual library collections using spatially-aware mobile displays. In *Chapter 4 – Spatial Navigation – Peephole Size & Navigation Behavior*, I operationalize this scenario and determine the optimal screen size of a mobile display, so-called "sweet-spot." In a consecutive lab experiment, in *Chapter 5 – Spatial Navigation – Spatial Memory*, I determine the effect the navigation technique for view management on a mobile device has on the users' navigation performance, spatial memory, and user experience.

1.2.2 Cross-Device Interaction

Nowadays, people often own several personal and shared computing devices such as smartphones, tablets, and laptops. They use them both sequentially or in parallel[4]. Multi-device activities and tasks range from emailing, internet browsing, social networking, playing games, searching, work documents, and watching a video. Similar multi-device use is also reported by Jokela et al. (Jokela et al., 2015a). Figure 1.4, for instance, illustrates parallel use of a laptop and a smartphone to prepare a presentation. Such a multi-device configuration, and often related cross-device interaction, is very common nowadays and, not solely but often, found when people work outside of their offices (e.g., in shared office spaces, in café, meeting rooms, or libraries). Scharf et al., for example, defines cross-device interaction as "the type of interaction, where human users interact with multiple separate input and output devices, where input devices will be used to manipulate content on output devices within a perceived interaction space with immediate and explicit feedback" (Scharf et al., 2013).

---

[4] The New Multi-screen World: Understanding Cross-platform Consumer Behavior – `https://think.withgoogle.com/databoard/media/pdfs/the-new-multi-screen-world-study_research-studies.pdf` (last accessed: February 5th, 2016)

Cross-Device Interaction Example

Figure 1.5 exemplifies a multi-device activity when information is received on one device but needed on another device. It is a real world scenario occurred to me while preparing for a research seminar talk. I created slides in Microsoft PowerPoint on the laptop, my primary device, and chatted with a colleague in WhatsApp[5] on the smartphone asking him for suggestions on relevant literature (see Figure 1.5a). My colleague sent me a link to a research paper, which might be of importance for the seminar topic. Instead of opening and reading the document on the small smartphone screen, I very much preferred reading it on the larger laptop screen. However, to have the document on the large screen it required seven steps (see Figure 1.5). These steps are necessary each time when sending information across device boundaries. Of course, there exist other options but out of subconscious action, and possibly due to past experiences, I opted for sending the document via email to my personal account (see Figure 1.5b-e). Once received in my inbox, I could open the email inbox on my laptop and download and comfortably read the document on its larger screen (see Figure 1.5f,g).

Apple tackles this issue with *Handoff*, which makes steps c-f obsolete. Handoff is part of the Continuity feature that "[...] lets you seamlessly move between your **iOS** devices and your **Mac**, or use them together"[6] (emphasizes done by the thesis author). However, as highlighted in the quote before, it only works within the Apple ecosystem and only with particular apps specifically designed to support Handoff. Since I also frequently work with Microsoft Windows and interchangeable use Windows, iOS, and Mac, I was used to the previous general working approach. When I realized it might be easier to use Handoff, in this case, it was already too late. Not to mention that Handoff only works for one's personal ecosystem but is inapplicable when a user needs to send information to another user.

Limitations & Potentials

Still, the vast majority of devices are blind to the presence of other devices and performing tasks among them is usually tedious (Greenberg et al., 2011) due to the lack of guiding principles (Oulasvirta, 2008). UbiComp technologies preordain, formerly monolithic applications, to be distributed across multiple interconnected devices (Tan et al., 2004; Wigdor et al., 2009; Yang and Wigdor, 2014). However,

---

[5] WhatsApp is an instant text messenger for mobile devices. Recently it offers access to messages through a web browser. – `https://www.whatsapp.com` (last accessed: April 28th, 2016)

[6] Apple Continuity – `https://support.apple.com/en-us/HT204681` (last accessed: March 7th, 2016)

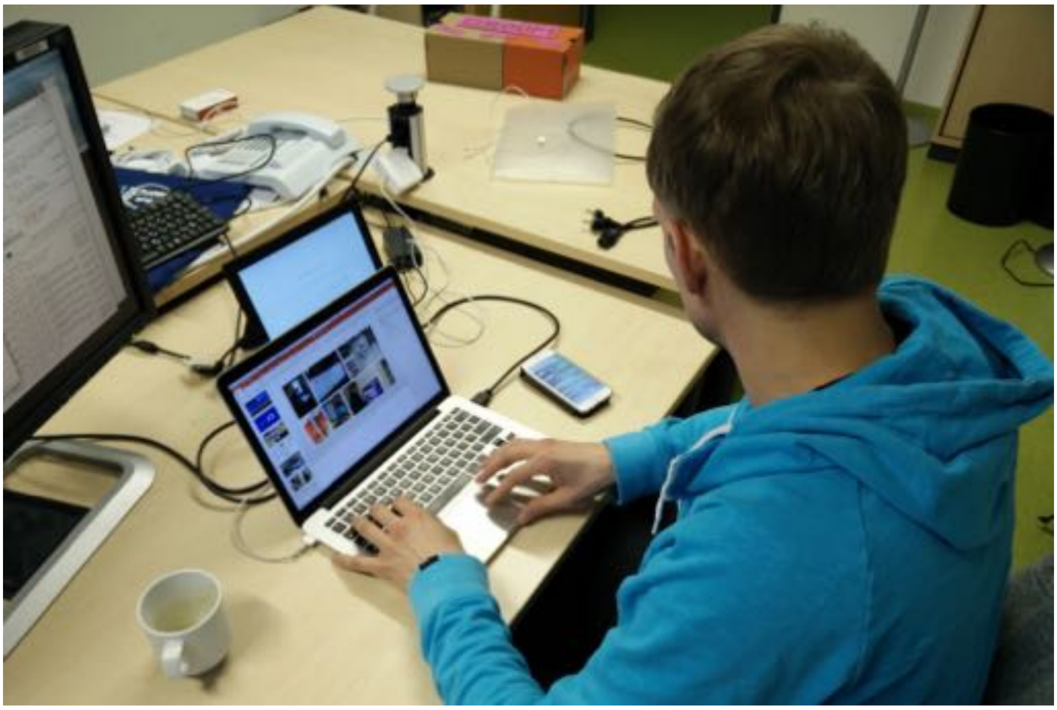
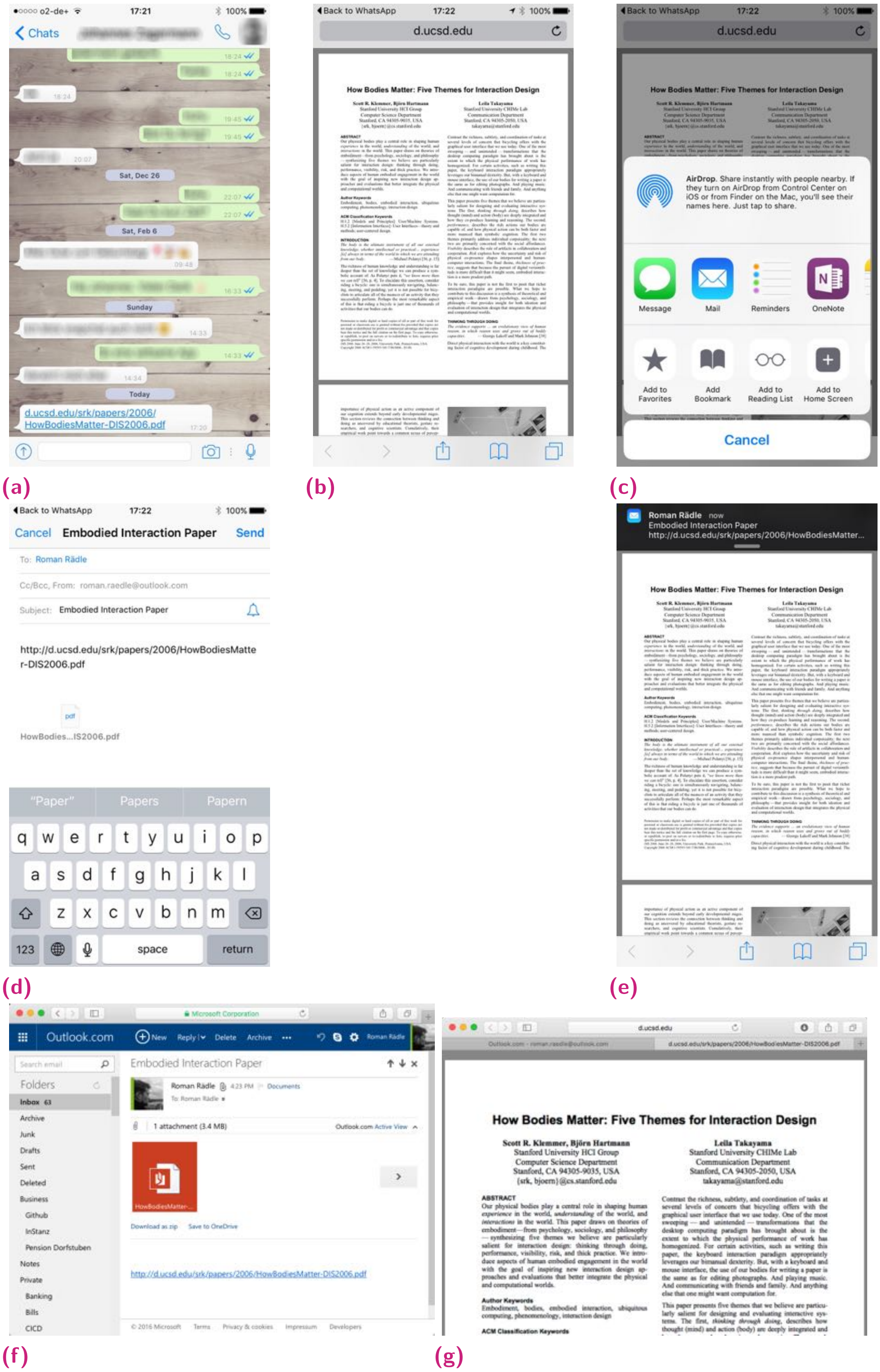


Figure 1.4.: Parallel use of laptop and smartphone to prepare a presentation. Such a multi-device configuration is very common nowadays and often found when people work in shared office spaces and café.

this “poses the challenge of developing UIs that span over this potentially wide range of computing platforms.” (Gallud et al., 2011, Forward by Jean Vanderdonckt and Vanâtorii Mari). Thus, distributing activities and tasks across device boundaries is a documented challenge in HCI and “there are very few means by which a user may take advantage of [. . .] large number of screens” (Hamilton and Wigdor, 2014).

Current cross-device interactions can be tedious, and owners of devices most often misuse services that were developed for other purposes. For example and as illustrated before, sending information in an email as an attachment to themselves to be able to open it on another device. This behavior has been discovered as a challenge by many HCI researchers (Dearman and Pierce, 2008; Jokela et al., 2015a) and even most recently by Cecchinato et al. (Cecchinato et al., 2016).

However in research, cross-device interaction between multiple mobile devices is an increasingly popular field of research in HCI (Chen et al., 2013; Hamilton and Wigdor, 2014; Li and Kobbelt, 2012; Lucero et al., 2011; Lucero et al., 2010; Nacenta et al., 2013). It can be regarded as the latest incarnation of the UbiComp vision (Weiser, 1991) in which user experiences truly begin to cross devices (Hamilton and Wigdor, 2014) and the co-located devices can be easily joined to create ad-hoc device communities (Jetter and Reiterer, 2013; Hamilton and Wigdor, 2014). Ideally, users experience such a community as a single seamless and natural UI (or even a



AirDrop. Share instantly with people nearby. If they turn on AirDrop from Control Center on iOS or from Finder on the Mac, you'll see their names here. Just tap to share.

To: Roman Rädle

Cc/Bcc, From: roman.raedle@outlook.com

Subject: Embodied Interaction Paper

http://d.ucsd.edu/srk/papers/2006/HowBodiesMatter-DIS2006.pdf

HowBodies...IS2006.pdf

Roman Rädle now
Embodied Interaction Paper
http://d.ucsd.edu/srk/papers/2006/HowBodiesMatter...

Embodied Interaction Paper

Roman Rädle 4:23 PM Documents
To: Roman Rädle

http://d.ucsd.edu/srk/papers/2006/HowBodiesMatter-DIS2006.pdf

Figure 1.5.: A real world and multi-device activity when information is received on one device but needed on another device. It illustrates a workflow when information is sent from a smartphone to a laptop. Figure(a-e) shows the display of a mobile device and Figure(f,g) shows a laptop screen.

"symphony of devices" (Hamilton and Wigdor, 2014)) that is flexible regarding use and is not restricted to a few possible configurations or predefined sequences of use (Jetter and Reiterer, 2013).

Existing research examples of cross-device interaction are diverse, ranging from collaborative photo sharing or brainstorming with smartphones (Lucero et al., 2010; Lucero et al., 2011) to multi-tablet active reading (Chen et al., 2012; Chen et al., 2013) and sensemaking (Hamilton and Wigdor, 2014). User studies have shown that multi-tablet systems can be successfully used in the wild (Chen et al., 2013) and that users can effectively manage cross-device interactions with 5 to 10 devices (Hamilton and Wigdor, 2014). However, many questions remain unanswered: How should cross-device interaction between mobile devices be designed so that they are easy to learn and easy to use? What role should increasingly popular technologies for sensing spatial configurations and detecting mid-air gestures play in their design? Should interactions follow a traditional, yet robust, non-spatial model, e.g., menu-based selection of devices (Hamilton and Wigdor, 2014)? Alternatively, should systems sense locations and use gestures to make cross-device interactions more like familiar non-digital interactions (Greenberg et al., 2011)?

These questions are important since our devices are still rather limited concerning sensing their mutual spatial relations (Hamilton and Wigdor, 2014) without using expensive or custom-built sensing hardware such as instrumented rooms with motion tracking systems (Marquardt et al., 2011b). Such spatial information could, however, enable interfaces to grow easily across nearby devices and annex them in natural ways (Hinckley et al., 2004; Pierce et al., 2003), ideally as a byproduct of natural use in space, e.g., by putting tablets on a table, moving them around, placing them side-by-side, performing pick-and-drop gestures between them.

We envision a future where mobile devices can contribute their interaction resources (e.g., their multi-touch displays) to a community of devices (Jetter and Rädle, 2013) in their proximity that then serves as one seamless multi-device user interface (UI). At any time, users can dynamically compose and fluidly reconfigure this UI according to their current needs and the task at hand. Ideally, cross-device interaction with mobile devices becomes a familiar experience almost similar to working with paper documents.

Prospect: In *Chapter 3.3.3 – Research Prototype 2: Integrative Workplace*, I motivate fluid device configurations and interaction across them. As a first incarnation and in *Chapter 6 – Cross-Device Interaction – Enabling Technology*, I present HuddleLamp, a desk lamp with an integrated low-cost RGB-D camera that detects and identifies mobile displays (e.g., smartphones or tablets) on tables allowing for fluid configurations of multiple mobile devices. In *Chapter 7 – Cross-Device Interaction – Understanding*

Spatial Cues, I explore the design space of mobile cross-device interactions and study the importance of spatial cues for cross-device object-movement tasks.

1.3 Research Questions

This thesis deals with limitations as mentioned earlier and discovers potentials for improvements. It particularly addresses two challenges within them:

- Enable users to exploit pre-existing knowledge to navigate and interact in virtual information spaces.
- Seek for opportunities to utilize commodity and off-the-shelf hardware to enable users to work across multiple mobile devices.

All challenges are concerned with the use of space and interaction in space. It seeks for an understanding of the importance of space as a cognitive resource by observing various spatial and cross-device interactions and their impact on users' performance (e.g., navigation and object recall) and subjective workloads (e.g., physical and mental demand). Research begins with egocentric spatial navigation with a single device and continues with interactions across multiple devices. Thereby, it addresses three research questions fundamental to future UbiComp and HCI:

RQ1 Does an egocentric spatial navigation improve users' navigation performance and increase their ability to recall information from memory?

RQ2 How can technology support egocentric spatial and cross-device interaction, so it seamlessly integrates into people's everyday practices?

RQ3 What are the benefits of spatially-aware cross-device interactions; and are they superior to non-spatial or spatially-agnostic cross-device interactions?

Answers to all three research questions will (i) contribute and foster understanding of spatial and cross-device interactions in the imminent mobility era and (ii) guide the design of future UbiComp experiences. Ideally, it will spawn new UbiComp experiences that lower users' subjective workload during navigation in virtual spaces and cross-device interaction. For example, by decreasing mental and physical demand and at the same time reducing user frustration. Overall, they seek to improve user experience and increase the cumulative value of UbiComp applications.

In the following, I introduce research approaches and methods applied for this thesis research. Then I give a brief overview of my research that attributes above research questions. In *Chapter 2 – Theoretical Background*, I concisely put my research in a broader context of HCI and relate it to UbiComp.

1.4 Research Approaches and Methods

Research presented in this thesis is four-layered and core research consists of two phases. Figure 1.6 (page 13) provides an overview of the overall research strategy illustrated as four layers (colored rectangles with rounded corners). It particularly denotes the two phases *Context & Analysis* (Phase 1) and *Empiricism & Technology* (Phase 2). Each of the four layers groups together common steps. The visual language of this research overview is explained in the next section.

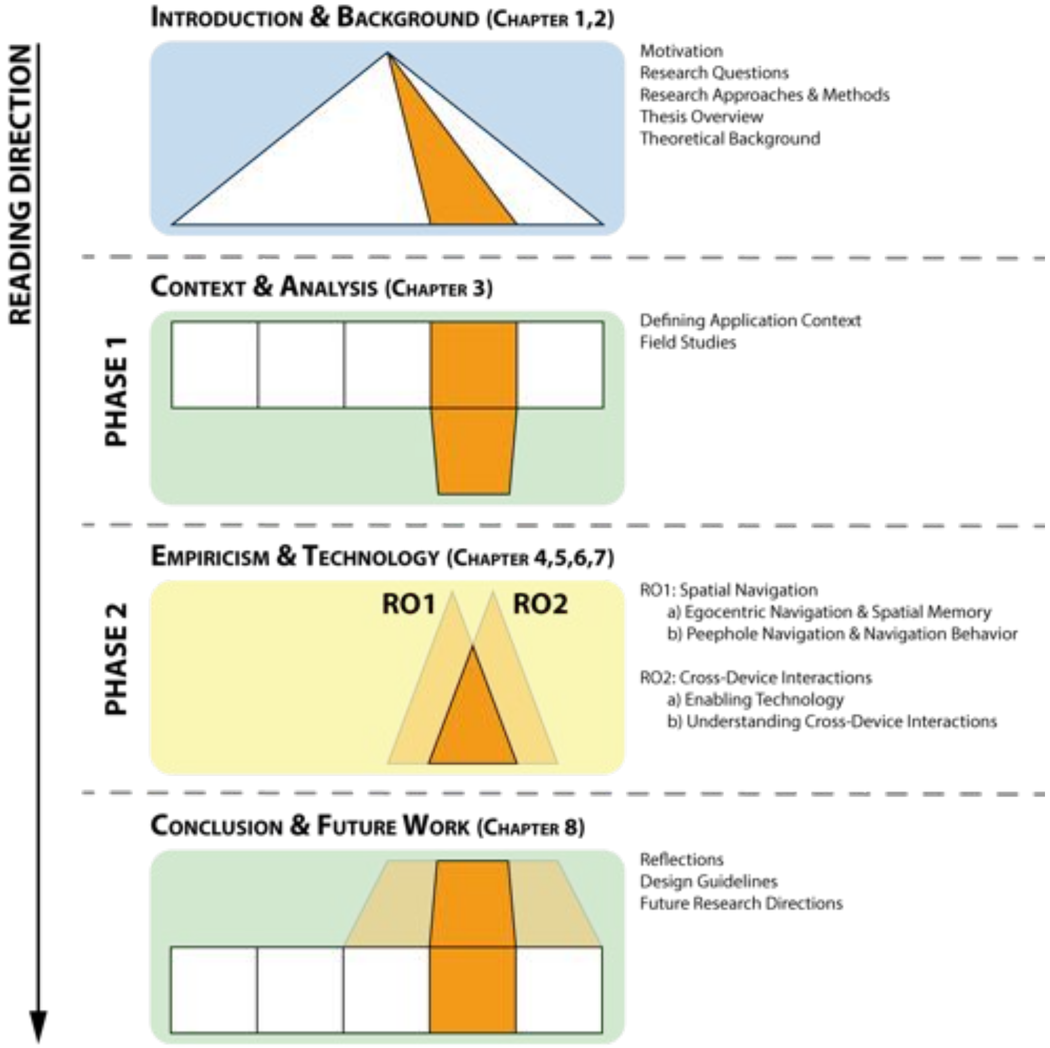


Figure 1.6.: The illustration gives a brief thesis overview with related chapters and a list of covered topics. It further illustrates research approaches including the two phases *Context & Analysis* and *Empiricism & Technology*.

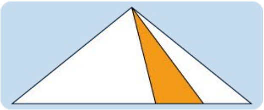
Research is approached through both deductive and inductive reasoning. In a pre-phase, it begins with a brief history of UbiComp and its overarching vision. Contradicting opinions on this vision are discussed before leading over to recent theories and believes on embodied cognition and models on human spatial memory. This theoretical background eventually thrives arguments for yet unexplored and hidden potentials for spatial and cross-device interactions.

In **Phase 1**, the application domain is narrowed down to academic libraries and knowledge work activities. In field studies at the Library of the University of Konstanz, the following two main knowledge work activities and resonating issues are identified: *literature & bibliographic search* and *reading & writing across documents*. Together with the theoretical background the found issues are transformed into potentials for future knowledge work. Thereby, two fully functional research prototypes, Blended Shelf and Integrative Workplace, were implemented to explore the problem space further and to derive research questions covered in this thesis.

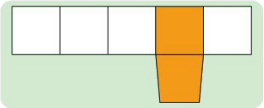
In **Phase 2**, research questions are tackled through controlled experiments and implementation of low-cost enabling technology. Several research prototypes were built to study potentials for embodied interaction and phenomena surrounding them. These small-scale research prototypes helped to answer specific research questions with high internal validity. Phase 2 also includes novel software technology that enables research of spatial and cross-device interactions outside of expensive and instrumented research facilities.

Visual Language of Research Overview

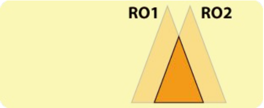
Figure 1.6 (page 13) follows a visual language, which is used throughout this thesis to offer anchor points as additional visual orientation for the reader. The four glyphs in this language are explained sequentially in the following paragraphs.



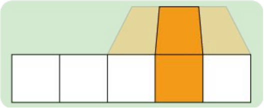
The **first layer** and introductory part motivates research from an HCI perspective. It is metaphorically represented as a white triangle. From the top seed point and in an act of exploration, it spans over the vision of UbiComp. Theoretical background thriving this thesis' research is highlighted in orange.



The **second layer** narrows research down to academic libraries. In this layer, each square glyph illustrates an individual application domain. For example, related domains like public libraries or special libraries but also unrelated domains like museums, schools, sports & fitness, or office work. The thesis focuses on a single domain, academic libraries (orange colored square). The orange colored trapezoid below insinuates exploring the problem space for this specific domain. Thereby it refines problems to operable units.



Similar to layer 1, each triangle in the **third layer** metaphorically starts from a single point. Each seed point reflects one of the two research objectives: *Spatial Navigation* (RO1) and *Cross-Device Interaction* (RO2). Even though both are subject to individual experimental research, their findings overlap and will be integrated in layer 4.



In the **fourth layer**, all findings from layer 3 are summarized and integrated to overall conclusions. Eventually, they are reflected and generalized. In this act of generalization, findings are transformed into design guidelines. Ideally, other researchers and practitioners can reuse and apply these design guidelines to other application domains (two semi-transparent trapezoids).

1.5 Thesis Contributions

With this research approach and applied research methods, this thesis provides a solid foundation for understanding of spatial navigation and cross-device interactions and the appropriate design of them. To answer my research questions, I further divided research into smaller research projects and prototypes instead of developing an overarching research application that spans across all research questions. First, this was important to avoid confounding factors and thus achieve valid results. Second and in a final step, it allows for generalization of research findings.

In detail, this thesis disseminates four types of contributions that are relevant to the HCI community: (i) empirical, (ii) technological, (iii) methodological, and lastly (iv) helps to spark academic discourse.

Empirical

This thesis provides insights into benefits of human spatial memory and muscle memory and how these insights can be leveraged for human-computer interaction. It supports manifestation of recent theories like embodied interaction (Dourish, 1999; Dourish, 2001) and the Reality-based Interaction framework (Jacob et al., 2008). Findings were published in a conference paper (Rädle et al., 2013a) at the ACM international conference on Interactive tabletops and surfaces – ITS ’13.

It adds further insights into human navigation behavior in physical space and beyond that shows empirical evidence of the existence of a learning phase and navigation phase during dynamic peephole navigation. Dynamic peephole navigation will be introduced in *Chapter 4 – Spatial Navigation – Peephole Size & Navigation Behavior*. The existence of a navigation phase indirectly hints for the existence of spatial memory, which will be discussed in later chapters.

It further provides insights into navigation performance and users' subjective workload for different peephole sizes. Empirical findings of this conducted research inform the design of navigation techniques for large information spaces that exceed the size of a single display. Findings of this work were published in a conference paper (Rädle et al., 2014a) at the 32nd annual ACM conference on Human factors in computing systems – CHI '14.

For cross-device interaction, this thesis provides findings that inform the design of object movement interactions across multiple mobile devices. Findings of this work were published in a conference paper (Rädle et al., 2015) at the 33rd annual ACM conference on Human factors in computing systems – CHI '15.

Results of conducted studies, further uncover users' subjective workload and experience during spatial navigation in virtual spaces and object movement tasks across multiple mobile devices. For instance, users' cognitive load, mental demand, physical demand, and user frustration.

Technological

According to views from ecological psychologists, gathering data in a laboratory (lab) and thus, in an artificial setting to the user can distort or even tamper with study results. During such lab experiments, participants are situated in an environment that is different to their workplaces and therefore can introduce confounding factors (e.g., Hawthorne-Effekt (Preim and Dachselt, 2010, p. 76)). Most often labs are clean (or even sterile), and participants rather feel like being in a clinical setting. Moreover, participants can be overwhelmed by technology. For instance, technology needed to track 3D location of devices. Such technologies require space or even entire rooms to be equipped with expensive tracking hardware. However, we also learned from ecological psychology that human behavior depends on the environment in which a study is conducted.

This thesis contributes technology to track multiple mobile devices outside of research facilities and allows for low-cost implementation of spatial navigation and cross-device interaction. In the long-run, it will allow research to be conducted in "the wild" and thus embracing methods from ecological psychology. Within the technological contribution, this thesis provides (i) new ways of tracking mobile devices that aim for less environmental instrumentation during user studies and (ii) contributes a novel software framework to develop cross-device interactions.

Methodological

For most of the research conducted and presented in this thesis, existing research methods had to be altered, and new research methods and approaches had to be developed.

First, to gather data with high internal validity during peephole interaction confounding factors like weight, resolution, and brightness had to be isolated. A new research method was developed to simulate different peephole sizes but keep all other variables same.

Second, the original method for user elicitation studies can introduce legacy bias (Morris et al., 2014). Independent of Morris et al., various aspects have been introduced to the existing method of user elicitation study to compensate for this phenomena.

Third, a new method had to be developed to compare tracking quality of multi-device and cross-device technology.

Methodological contributions are published as part of several publications (Rädle et al., 2013a; Rädle et al., 2014a; Rädle et al., 2014c; Rädle et al., 2015).

Academic Discourse

Research presented in this thesis is inspired by academic workshops and seminars, which the author of this thesis co-organized. For example, the workshops "Visual Adaptation of Interfaces" (Dostal et al., 2013) at the ACM international conference on Interactive tabletops and surfaces – ITS '13 and "Proxemics 2012" [7] at the 7th Nordic Conference on Human-Computer Interaction – NordiCHI '12. As a result of the latter, the author co-organized the Dagstuhl Seminar 13452 "Proxemics in Human-Computer Interaction" (Greenberg et al., 2014b) together with Saul Greenberg, Kasper Hornbæk, Aaron Quigley, and Harald Reiterer. The seminar was held

---

[7] Proxemics 2012 Workshop Website – `http://hci.uni-konstanz.de/proxemics/` (last accessed February: 8th, 2016)

from November 3rd to November 8th, 2013[8] and triggered several joint research publications. For example, HuddleLamp presented in *Chapter 6 – Cross-Device Interaction – Enabling Technology* was sparked from the Dagstuhl Seminar. Publications of other researchers participated in the seminar resulted in (Greenberg et al., 2014a; Mueller et al., 2014).

1.6 Chapter Preview

The following paragraphs provide an overview of thesis chapters and brief excerpts of their contents. This thesis is generally structured in eight chapters including the current introduction chapter. Consecutive chapters provide a theoretical background (Chapter 2), application context and problem space analysis (Chapter 3), studies on spatial navigation and cross-device interactions (Chapter 4, 5, 7), a framework for multi-device tracking (Chapter 6), and a summary of findings and research conclusion (Chapter 8). Contributions to each research questions **RQ1**, **RQ2**, and **RQ3** are highlighted in each excerpt.

Chapter 2 – Theoretical Background provides a deeper understanding of theoretical foundations that guided this research. The theoretical foundations are based on recent believes on embodied cognition. It further establishes an understanding of human spatial memory and its relevance to HCI.

Chapter 3 – Context & Analysis narrows down the research context to academic libraries. Therein, it focusses on knowledge work activities. Analysis of data gathered through interviews, questionnaires, and observations reveal prevalent issues of knowledge work. Findings of the analysis are discussed with respect to individual and group work practices in libraries. They spark particular research efforts conducted in consecutive chapters.

Chapter 4 – Spatial Navigation – Peephole Size & Navigation Behavior presents findings on the effect of display size during spatial peephole navigation. They provide insights into an optimal display size ("sweet-spot") for spatial egocentric navigation. It also reports on differences in users' navigation behavior with distinct learning phase and navigation phase (**RQ1**,**RQ2**).

Chapter 5 – Spatial Navigation – Spatial Memory provides insights into benefits of spatial egocentric navigation over traditional multi-touch interaction. To explore the design space, different designs of egocentric spatial navigation techniques are presented. Further empirically gathered data contrasts a traditional multi-touch

---

[8] Dagstuhl Seminar Website – `http://www.dagstuhl.de/de/programm/kalender/semhp/?semnr=13452` (last accessed: February 8th, 2016)

navigation technique to one selected egocentric spatial navigation technique. A comparative study design was conducted to study impact on users' navigation performance. Both techniques are compared according to users' capabilities to exploit their spatial memory and long-term spatial memory (**RQ1**). Further data points were gathered to contrast both techniques according to users' subjective workload. For example, mental and physical demand, effort, and user frustration.

Chapter 6 – Cross-Device Interaction – Enabling Technology researches technology to support cross-device interactions. It focusses on minimizing *a priori* setup of technology to allow for ad-hoc use of multiple mobile devices (**RQ2**). This technology eventually will also allow for fluid device configurations. It provides a high-level API, which is based on latest HTML5, CSS3, and JavaScript standards. This API eases implementation of cross-device applications. Finally, the resulting cross-device interaction design space is explored by various examples of spatial navigation, multi-device configurations, and cross-device interactions.

Chapter 7 – Cross-Device Interaction – Understanding Spatial Cues presents findings on the appropriate design of cross-device interaction (**RQ3**). It provides insights into cross-device object-movement tasks and differences between competing approaches such as spatially-aware versus spatially-agnostic techniques.

Chapter 8 – Conclusion summarizes thesis contents. It integrates findings of experimental research and reflects on them. Eventually, it generalizes results and derives design guidelines, which can be reused and applied by practitioners and other researchers. Finally, it concludes with an outlook on future work.

1.6.1 Additional Digital Media

All throughout this thesis will be QR codes located at page margins. The QR codes link to accompanying digital media content if available. Some link to video figures to illustrate interaction techniques or to present study footage. Others link to websites. For technical contributions, the QR codes link to applications, study prototypes, or open source software repositories. For devices that do not have a QR code reader, please visit `https://www.romanraedle.com/phd-thesis/introduction` or use the *MediaBrowser* application on the USB stick. To show the media content in the MediaBrowser, open the MediaBrowser application and enter the corresponding *Code #* (below the QR code) into the text box.

Code #<NUMBER_TO_ENTER>

Parts of the next Chapter 2 appear in the following publications:

Rädle, R. (2013). “Design and evaluation of proxemics-aware environments to support navigation in large information spaces”. In: *CHI ’13 Extended Abstracts on Human Factors in Computing Systems on - CHI EA ’13*. New York, NY, USA: ACM, p. 1957. DOI: `10.1145/2468356.2468710`

Rädle, R. Jetter, H.-C. Butscher, S. Reiterer, H. (2013a). “The effect of egocentric body movements on users’ navigation performance and spatial memory in zoomable user interfaces”. In: *Proceedings of the 2013 ACM international conference on Interactive tabletops and surfaces - ITS ’13*. New York, New York, USA: ACM Press, pp. 23–32. DOI: `10.1145/2512349.2512811`[9]

---

[9] The responsibilities for this joint publication were divided as follows: I formulated the research question, designed and conducted the study, analyzed the study data, and spearheaded the writing. Hans-Christian Jetter helped in formulating the research question and writing the paper. Simon Butscher helped in analyzing the study data. Harald Reiterer supervised the work.

Theoretical Background

2

> *All theories are legitimate, no matter. What matters is what you do with them.*
>
> — **Jorge Luis Borges**
> (Librarian, public lecturer, writer, and poet)

Contents

This chapter starts with an overview of the origins of UbiComp and the long history related to it. It highlights the increasing trend and interest of HCI researchers to turn the UbiComp vision into reality, and ideally into an everyday experience. Importantly, this chapter reflects on various — somewhat contradicting — opinions of scientists on the success or failure of this overarching vision. As a result of this discussion, it exposes the need and space for further research.

Despite these contradictions, there is not doubt that the UbiComp vision opened up new potentials for human-computer interaction. For example, it eventually sparked research on tangible and social computing, which both redound to embodied interaction (Dourish, 1999). It emphasizes on new views of embodied interaction and highlights on related HCI theories, models, and frameworks.

Also, this chapter briefly introduces human spatial memory as significant cognitive resource leveraged by embodied interaction. It further highlights the use of spatial memory in HCI.

Overall, the theoretical background spans like an umbrella over work presented in consecutive chapters. It drives the design of research prototypes, resonating experimental studies and feeds the discussion of research findings.

2.1 The UbiComp Trend

In 1991, Mark Weiser published "*The computer for the 21st century*" in the Scientific America Journal (Weiser, 1991). In this article, he describes the vision of "*Ubiquitous Computing*" (UbiComp). Therein, he also coins the term “Embodied Virtuality”. Embodied Virtuality is an alternative and even stronger notion of UbiComp and as opposition to the concept of "virtual reality." It highlights on the embodiment of digital powers. Embodied Virtuality envisions a world where computers are drawn out of their electronic shells and digital functions become an appearance in the real world, so they tightly interweave themselves with the physical world (Weiser, 1991). Ideally digital functions become indistinct — invisible as Weiser calls it — from natural, real world objects. Weiser proposes pads, tabs, and boards as first incarnations of such computing objects. Nowadays they are known as smartphones, tablets, and wall-sized interactive screens, henceforth just called devices.

Through this embodiment, computing becomes graspable by humans. It opens up possibilities for them to sense computing input and output by other modalities beyond just sight and hearing. Such multimodal interaction “expand the size of [. . .] available working memory” (Oviatt, 2006) and improve error handling and reliability. It “minimizes users’ cognitive load, which effectively frees up mental resources for performing better while also remaining more attuned to the world around them” (Dumas et al., 2009).

A new belief by Jacob et al. summarizes these advantages. They “believe that all of these new interaction styles draw strength by building on users’ pre-existing knowledge of the everyday, non-digital world to a much greater extent than before.” (Jacob et al., 2008). Ever since the vision of UbiComp has inspired many scientists and researchers all around the globe with different research backgrounds. Ranging from social sciences (Bell and Dourish, 2006), psychology (Ohlsson, 1995; Rogers, 2006), design (Jetter, 2013; Greenberg et al., 2014a; Aylett and Quigley, 2015), media studies (Durrant et al., 2011), engineering (Hinckley, 2003; Hinckley et al., 2004; Marquardt et al., 2011b; Marquardt et al., 2012b), computer science (Greenberg et al., 2011; Klokmose et al., 2015), politics (Luger, 2012; Luger and Rodden, 2013), and several other fields of study. Evidently, the paper "*The computer for the 21st century*" is one of the most cited papers in the scientific literature with over twelve thousand citations[1] up to today.

In particular, HCI researchers pursued Mark Weiser's vision of disappearing, invisible computing technologies that "weave themselves into the fabric of everyday life until they (computing technologies – Ed. Note) are indistinguishable from it" (Weiser, 1991). HCI is an interdisciplinary field of research and located at the intersection of studies mentioned above. It is also the hosting discipline of UbiComp. HCI is "concerned with the design, evaluation and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them."[2]

As one phenomenon, Mark Weiser's pioneering work and vision of UbiComp triggered an avalanche of scientific publications. They are concerned with pervasive[3] and ubiquitous computing; often urging the need for post-WIMP interactions. Figure 2.1 (page 28) reveals this, ever since 1991, increasing trend and interest in UbiComp and its related device form factors: tabs, pads, and boards.

The plot was created by analyzing the number of returned Google Scholar search results[4] when searching for "*Ubiquitous Computing*" and UbiComp device form factors as additional keywords. All returned numbers might not reflect the exact number of publications for each particular year, but they certainly reflect a close approximation. The analysis starts in 1991 when Mark Weiser first published work on ubiquitous computing and ends most recently in 2014[5].

The light blue line indicates the overall[6] number of publications containing the term "*Ubiquitous Computing*", grouped by year with its highest peak in 2013 and nearly fifteen thousand papers. Additional searches drill down into the three form factors tabs, pads, and boards. For instance, searching for "*Ubiquitous Computing*" *AND ("tablets" OR "slates")* returns the approximate number of publications for the

---

[1] A search with Google Scholar shows more than 12 thousand citations of the original article from 1991. (last accessed: April 10th, 2015)

[2] ACM SIGCHI Curricula for Human-Computer Interaction — `http://old.sigchi.org/cdg/cdg2.html` (last accessed: July 13th, 2015)

[3] Pervasive computing is synonymously used for ubiquitous computing.

[4] Number of publications were retrieved using `http://www.csullender.com/scholar/`

[5] Valid data for 2015 and 2016 was not available at the time of writing.

[6] A sentiment analysis has purposefully not been conducted since both positive as well as negative statements count as equally thought provoking and spark discussion about UbiComp.

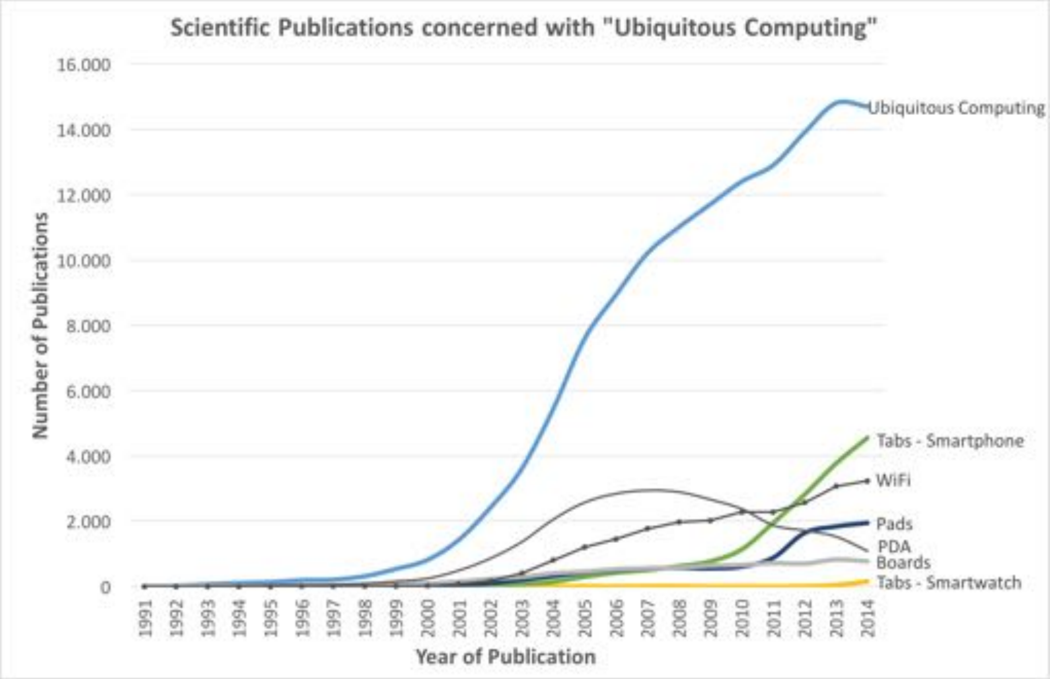


Figure 2.1.: The increasing trend and interest in Ubiquitous Computing and its related device form factors: tabs, pads, and boards.

device form factor pads (dark blue line). The query for boards (gray line) contained tabletop and interactive whiteboard but also included particular devices and brands such as Samsung SUR40, Smartboard, PixelSense, or Perceptive Pixel. Smartphones (green line) and smartwatches (yellow line) are split into two independent categories because a form factor of a size of a smartwatch has not been considered in the original vision of ubiquitous computing. However, it is becoming increasingly popular in HCI (Houben and Marquardt, 2015; Nebeling et al., 2014).

Interestingly, this plot resonates with release dates of UbiComp-related consumer devices. For example, the IBM Simon delivered on August 16th, 1994. It is considered the first so-called "smart phone"[7]. Another example is Nokia's Communicator 9000-series released in 1996, which became the world's best-selling personal digital assistant (PDA). As visualized in Figure 2.1 (page 28), since 1996 PDAs were mentioned in nearly 50% of all scientific publications. It flattens out in 2007 when Apple released the first generation of the Apple iPhone (release date on June 29th, 2007). Another trend that is visible in the plot is the increase of tablet sales attributable to Apple's release of the first iPad generation on April 3rd, 2010.

Apart from devices, Mark Weiser also proposed wireless communication between these devices. Devices that “will be interconnected in a ubiquitous network” (Weiser, 1991). With the invention and standardization of radio-based communication

---

[7] Buxton Collection: `http://www.chi2011.org/program/buxtoncollection.pdf` (last accessed: March 30th, 2016)

Figure 2.2.: The Nokia Communicator 9000 and first successful and best-selling "smart phone" (left). The first Apple iPhone released in 2007 (right).

technology like Wi-Fi and the 802.11b standard[8], technology for UbiComp was not fiction anymore but broadly available for consumers as off-the-shelf devices. If connected to the World Wide Web, those devices could access resources from servers located in different countries of even other continents. Most recently, Bluetooth 4.0 LE (low energy) standard was developed for ad-hoc short-range communication. These standards allow their users to create private local networks (e.g., through Wi-Fi hotspot) to share resources among a small community of devices without tedious setup of Wi-Fi routers, Wi-Fi access points, and alike.

2.1.1 The Dream of UbiComp

Now, 26 years (1991-2017) later and on the bottom line, the vision of UbiComp is becoming true (Bell and Dourish, 2006). Weiser anticipated that interaction between humans and computers for everyday tasks could become just as ubiquitous and handy as ink & paper became a ubiquitous "literacy technology" (Weiser, 1991). Beyond “light switches, thermostats, stereos and ovens” (Weiser, 1991), computing technology found its way further into our everyday life and vanished into everyday objects such as lamps, blinds, coffee machines, electric kettles, and nowadays even toothbrushes (see Figure 2.3, page 30). They are interconnected through above-mentioned ubiquitous networks and exchange data to communicate with each other. We can control our homes while being on vacation, see how effective we brush our teeth, let us tell that we need to start workout when we get lazy, or when and what we need to eat to live a healthier life. Just like Sal in the UbiComp vision (Weiser, 1991). So, did we achieve UbiComp? Can we tick the UbiComp vision of our list and continue to start working on other visions[9] for future human-computer interaction?

---

[8] The 802.11b standard permits 11 Mbit/s link speeds and was already faster than traditional coaxial cabled-based BNC network interface adapters.

[9] A list of other visions of computing can be found in the Wiki at visionsofcomputing.com

Figure 2.3.: Examples of off-the-shelf UbiComp devices finding their way into our homes. It ranges from toothbrush[10], over thermostat[11], to electric kettles[12], and coffee machines[13].

---

[10] Kolibree Ara – `https://www.kolibree.com/` (last accessed: April 12th, 2017)

[11] Nest Thermostat – `https://nest.com/thermostat/` (last accessed: April 12th, 2017)

[12] Smarter iKettle – `http://smarter.am/ikettle/` (last accessed: April 12th, 2017)

[13] Philips Saeco GrandBaristo – `http://www.philips.com/c-m-ho/saeco-espresso/granbaristo-avanti` (last accessed: April 12th, 2017)

It would be naïve to answer these questions with "yes" or "no." Even well-established HCI researchers come to inconsistent conclusions. In contrast to Bell and Dourish who argue that “ubiquitous computing has [. . .] arrived” (Bell and Dourish, 2006), Aylett and Quigley debate about the broken dream of ~~pervasive sentient ambient calm invisible~~ ubiquitous computing (Aylett and Quigley, 2015)(strikethrough adopted from original paper title – Ed. Note). And yet others, propose an “alternative agenda,” which for instance “focuses on designing UbiComp technologies for engaging user experiences” rather than technology dictating the user what they have to do (Rogers, 2006).

Figure 2.4.: Microsoft’s product family and device ecosystem ranges from large public and shared displays (Microsoft Surface Hub) and game consoles (Microsoft Xbox) over personal devices such computers, laptops, tablets, smartphones, to IoT (Internet of Things) devices.

No doubt, people’s everyday life is close to the vision of UbiComp, where computing technology is virtually embodied in devices. Nowadays, multifunctional and personal mobile devices (e.g., smartphones, pads, tablets) and collaborative and interactive shared spaces (e.g., large interactive walls and tabletops) exist and are either available as off-the-shelf consumer products (see Figure 2.4, page 31) or pre-installed in offices, meeting rooms, public spaces, libraries, or museums. Such mobile devices are our personal companions and if not already in hand, they idle away in our pockets and bags always operable and waiting to be activated.

Regardless of devices’ availability, their cost is still beyond being considered disposable goods, for which reason we do not own a magnitude of them. Even aforementioned interactive spaces do not yet “contain hundreds of [. . .] tiny computers” (Weiser, 1991). On the one hand, most of these UbiComp devices are yet too expensive. On the other hand, their users hesitate to share them just like pen & paper.

Figure 2.5.: Mobile devices are far beyond computing devices. They are customized devices with which their users express their current lifestyle.

Such devices are often highly customized, and they arguably established themselves as lifestyle products (see Figure 2.5, page 32). For instance, the look & feel of the operating system is changed to owners' favorite color or even more shoved by them in goofy smartphone cases to express their current lifestyle. Besides, industry has no incentive to encourage people to share mobile devices for maximization of companies profit, as Aylett and Quigley surmise (Aylett and Quigley, 2015).

Of course, new emerging technologies such as print screens (Olberding et al., 2014) or mixed-reality glasses (e.g., Microsoft HoloLens[14]) could be game changers. Print screens could be a low-cost alternative to produce a magnitude of cheap and disposable displays when needed. Trading for haptic feedback, holographic or mixed-reality displays would even further circumvent production of physical displays. Both technologies eventually change the way people think about UbiComp technology (see Figure 2.6, page 33) — just as it happened to pen & paper centuries ago.

However and despite these technological advancements, Mark Weiser's vision was more than literally invisible computing technology. As outlined at the beginning of this chapter, embodied virtuality takes the human world and physical space into account, instead of solely focusing on (computing) technology. UbiComp technology should only enhance "the world that already exists" (Weiser, 1991) rather than replacing it or dictate users living in that world. It is supposed to stay calm at the

---

[14] Official Website of Microsoft HoloLens – `https://www.microsoft.com/microsoft-hololens/en-us` (last accessed: July 13th, 2015)

Figure 2.6.: Examples of future UbiComp systems, which promise to create cheap and disposable UbiComp devices or simulate entire UbiComp environments. Left is PrintScreen allowing fabrication of thin touch-displays (Olberding et al., 2014) and right is Microsoft HoloLens showing Minecraft Demo.

periphery of users' perception and only enter the center of users when they need computing power to support them in solving their tasks (Weiser, 1999; Rogers, 2006). It should be approachable at any time without a "complex jargon." (Weiser, 1991), so that handling and computing of information can happen beneath the conscious awareness of the user. This way, users can utilize their cognitive resources on "application level tasks (e.g., reading, annotating, search)" rather than being distracted by "system level tasks (e.g., view management)" (Andrews et al., 2010).

2.1.2 The Shift to UbiComp Experiences

It is time to rethink UbiComp fundamentally towards computing experiences that exploit users' pre-existing knowledge of everyday life. These computing experiences should account for calm computing (Weiser, 1999; Rogers, 2006) and subtly leverage and extend users' capabilities far beyond reality. For example, by considering tradeoffs between embodied practices and digital power (Jacob et al., 2008).

Current UbiComp Experiences

To get the first impression of current UbiComp experiences, follow Mark Weiser's instruction and "Look around you" (Weiser, 1991)! As one example, Figure 2.7 (page 34) shows my rather messy and cluttered[15] office space, in which I was situated at the moment of writing this paragraph. It is worth to mention that our group[16] recently moved to a new building and modern space and thus actually reflects a

---

[15] I love to quote Albert Einstein to excuse for not having a tidy desk. "If a cluttered desk is a sign of a cluttered mind, of what, then, is an empty desk a sign" (Albert Einstein)

[16] Our group refers to the Human-Computer Interaction Group, University of Konstanz, Germany led by Prof. Dr. Harald Reiterer.

present-day office space. It is very likely that you will find a similar setting in your office. A setting with no "more than 100 tabs, 10 to 20 pads and one or two boards" (Weiser, 1991).

However, most of the devices in this space are still non-digital. Only a desktop computer, a WACOM display, and my smartphone offer digital tool support. All of these devices are located in the same room and are eventually connected to the same logical network. But they are blind to the presence of each other; not to mention that they are agnostic of their relative or absolute spatial location, orientation, and distance. The only way they can communicate and share information with each other is through centrally managed cloud services (Jokela et al., 2015a), portable drives (Dearman and Pierce, 2008), network drives (Jokela et al., 2015a), or by sending documents as emails attachments to personal accounts (Dearman and Pierce, 2008; Cecchinato et al., 2015). The latter approach is a well-documented behavior, even when managing information across someone's multi-device ecosystem (Cecchinato et al., 2016).

Figure 2.7.: A panoramic image taken from my office space. This space probably reflects a common office space with various digital devices used in day-to-day office work and business.

Even this quick sample reveals the still existing gap between Mark Weiser's UbiComp vision and current UbiComp experiences — leaving yet enough room for improvements. In a systematic and research-oriented approach, this thesis dives deep and identifies problems of current UbiComp experiences in the library context (see *Chapter 3 – Context & Analysis*). It presents solutions and insights to these problems, which were developed governed by the following theoretical foundation.

2.2 Theoretical Foundation

Emerging theories, models, and frameworks in HCI can provide a solid theoretical foundation to predict, spark, inspire, and guide the development of novel UbiComp experiences. For instance, theories and frameworks such as Embodied Interaction (Dourish, 1999; Dourish, 2001), Tangible Computing and frameworks for tangible user interfaces (TUI) (Ullmer and Ishii, 2000), Social Computing concerned with the intersection of social science and computing systems (Schuler, 1994), or more recently Blended Interaction (Jetter et al., 2014). Ideally, they provide a fresh look on UbiComp and help to design UbiComp experiences that shift away from “proactive computing to proactive people” (Rogers, 2006). UbiComp experiences that “are designed not to do things for people but to engage them more actively in what they currently do” (Rogers, 2006). They could help to achieve calm computing and eventually fix the broken dream of UbiComp (Aylett and Quigley, 2015).

2.2.1 Interaction with a Model World

It begins with the model of direct manipulation to understand computing experience and user engagement. As explained in *Chapter 1.2.1 – Spatial Navigation*, navigation in virtual information spaces is different and decoupled from the real physical world. Hutchins et al. describe this decoupling as the directness of manipulation[17] and even further differentiate between two aspects of it: *distance* and *engagement* (Hutchins et al., 1985).

Distance is a bi-directional relationship between human and computer and describes the interaction between them. It is (i) the “distance between one’s thoughts” and the required physical actions to operate a computer and (ii) the distance between the form of system output “readily interpretable in terms of the goals of interest to the user” (Hutchins et al., 1985). Hutchins et al., and as presented in Figure 2.8 (page 35), name the first (i) distance “Gulf of Execution” and the second (ii) distance “Gulf of Evaluation”.

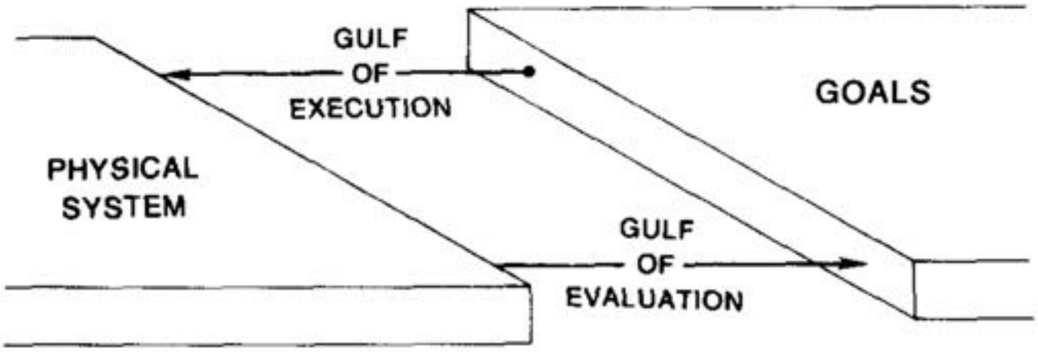


Figure 2.8.: The two gulfs “Gulf of Execution” and “Gulf of Evaluation” metaphorically represent the distances in Hutchins et al. model of direct manipulation. (Source: (Hutchins et al., 1985, p.319))

Engagement is a bit more complex. When interacting with a computer, users need to construct a world model. The construction of a world model is a cognitive process, which creates an understanding of how to interact with the computer. Based on this

---

[17] Direct manipulation was first coined by Shneiderman in (Shneiderman, 1983).

understanding it allows users to *a priori* manipulate this world in their minds. They, then, induce actions that enable them to interact with the computer. This rule-based cognitive process is well-documented in psychology and, for instance, postulated in ACT* theories (Anderson, 1996). Of course, a world model is dynamic and can change over time. Ideally, objects in the user's world model are same or similar to objects represented in a user interface. Thereby, these interface objects "can create the sensation in the user of acting upon the objects of the task domain themselves", rather than interacting with intermediary objects. This sensation or experience, for example, is found in modern WIMP user interfaces where windows immediately grow or shrink when resized according to a user's physical movement of the mouse. Hutchins et al. "call this aspect of directness direct engagement."

2.2.2 Embodied Interaction

The direct manipulation model is great to explain the success of WIMP user interfaces over command line interfaces. However, it solely focusses on cognitive processes and does not encounter any of the physical properties and capabilities of the human body. For instance, memory activated beneath the conscious awareness of users (e.g., muscle memory) (Scott et al., 2001) or their physical constraints (e.g., constraints when moving limbs). Klemmer et al. argue that "less constraining interaction styles are likely to help users think and communicate" (Klemmer et al., 2006). In consequence, direct manipulation is limited to explain the recent success of UbiComp and post-WIMP user interfaces.

Emerging theories include the human body as an essential part of the human-computer interaction and consider it as an instrument to facilitate external cognition and to think in space (Kirsh, 2010). Klemmer et al. puts it as "moving [. . .] in the world helps infants to learn about the physics of the world and consequences of actions, gesture plays a role in pre-linguistic communication for babies as well as aids cognition and fully linguistic communication for adults" (Klemmer et al., 2006). Even further, "evidence supports [. . .] an evolutionary view of human reason, in which reason uses and grows out of bodily capacities" (Lakoff and Johnson, 1999).

Therefore, the embodied interaction theory (Dourish, 2001) proposes a new fresh look on HCI. It grounds in contemporary psychological, sociological, and anthropological perspectives and applies those to computer science. It is concerned with human's cognitive abilities, their behavior in social contexts and environments, and the mutual interdependencies between them. Embodied interaction further connects the two independently evolved strains of tangible computing (Ullmer and Ishii, 2000; Ishii et al., 2012) and social computing (Schuler, 1994). One of the key arguments of embodied interaction is the interdependency between thinking and action, which is rooted in the embodied view of cognition (Dourish, 2001).

2.2.3 Reality-Based Interaction

Reality-Based Interaction (RBI) framework, a descriptive framework, also considers the human mind and human body as mutually affecting entities (Jacob et al., 2008). Jacob et al.'s RBI framework is based on the following four guiding principles, so-called themes: *Naïve Physics*, *Body Awareness & Skills*, *Environment Awareness & Skills*, and *Social Awareness & Skills* (see Figure 2.9, page 37). They "attempt to make computer interaction more like interacting with the real, non-digital world" and thereby "drawing upon these themes of reality, emerging interaction styles often reduce the gulf of execution (Ed. note – gulf of execution refers to direct manipulation (Hutchins et al., 1985))". The four themes of RBI are explained in the following.

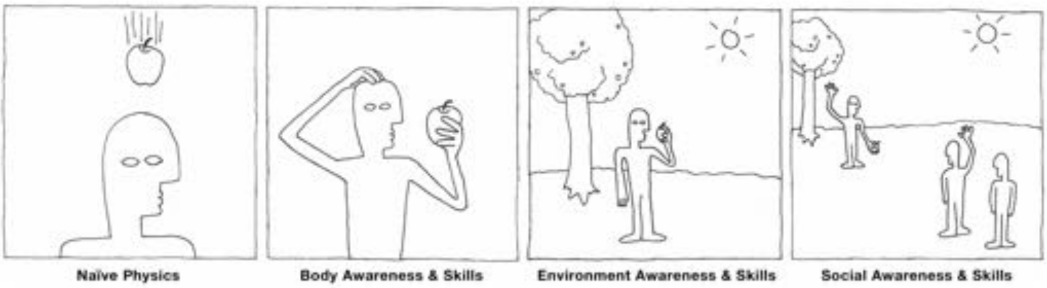


Figure 2.9.: The four themes of Reality-Based Interaction: Naïve Physics, Body Awareness & Skills, Environment Awareness & Skills, and Social Awareness & Skills. (illustration from (Jacob et al., 2008))

Naïve Physics (NP) is the common understanding of underlying physical laws of our non-digital world. "This includes concepts like gravity, friction, velocity, the persistence of objects, and relative scale" (Jacob et al., 2008). Such concepts are usually learned from childhood on. For example, through learning by doing such as experimenting with the concept of containment when successfully sticking smaller objects into bigger containers or failing when doing vice versa. Another example is learning how to ride a bicycle.

Body Awareness & Skills (BAS) describes the perception and understanding of the human body, proprioceptive skills, and the awareness of relative position and orientation of limbs. For instance, awareness of the location of arms, hands, fingers, legs, feet, and toes from an egocentric viewpoint. This understanding is independent of the environment and exploited for everyday activities such as walking, jumping, or climbing.

Environment Awareness & Skills (EAS) is the understanding of a world and the human body embedded in this world. The state of the world and other embedded objects "facilitate our sense of orientation and spatial understanding" (Jacob et al., 2008). Even without having a natural horizon, buildings, streets, and vehicles provide an understanding of the orientation of the horizon. Further, an object partially covered by another object gives a sense of relative distance between these objects. If the (approximate) size of an object is known it further provides information about the relative distance to the own body (e.g., the distance to a car).

Social Awareness & Skills (SAS) addresses social protocols between human-human interaction. It includes verbal and non-verbal communication such as an interaction between group members. In contrast to the other themes, such social rules depend on the culture and personal relationship between peers. For instance, shaking hands is a non-verbal ritual to greet peers in many countries. However, there are subtle differences. Whereas in western countries, a firm handshake communicates self-confidence it is considered as impolite in Asian countries.

These themes guide researchers in the design process (Geyer, 2013). Ideally, the framework helps to identify appropriate real world analogies for design, which leads to an improved user experience. Jacob et al. argue that such real world analogies can create a reality-based sensation for the user so that basing "interaction on pre-existing real world knowledge and skills may reduce the mental effort required to operate a system because users already possess the skills needed" (Jacob et al., 2008).

But the mimicking of reality is often not optimal and comes with a cost; it comes with the cost of losing digital power. Therefore, the framework offers RBI design tradeoffs, as illustrated in Figure 2.10 (page 38). They guide researchers to evaluate intrinsic characteristics of reality and digital power and weigh their tradeoffs.

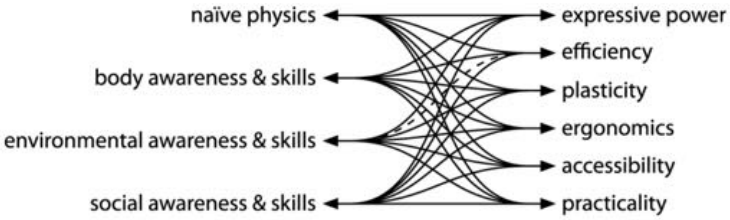


Figure 2.10.: Reality-Based Interaction design tradeoffs between reality and expressive power. (illustration from (Jacob et al., 2008))

These tradeoffs are as follows and entirely adapted from Jacob et al. (Jacob et al., 2008).

- *Expressive Power:* i.e., users can perform a variety of tasks within the application domain
- *Efficiency:* users can perform a task rapidly
- *Versatility:* users can perform many tasks from different application domains
- *Ergonomics:* users can perform a task without physical injury or fatigue
- *Accessibility:* users with a variety of abilities can perform a task
- *Practicality:* the system is practical to develop and produce

All themes of reality and RBI tradeoffs impact this thesis research. They were applied in the design of research prototypes and guided discussion of findings of experimental studies. First, they guided the design of UbiComp systems that leverage users' pre-existing practical knowledge. But at the same time, they help to create systems that equip users with super-powers beyond reality. Presumably, UbiComp experiences designed with respect to RBI design principles intrinsically lower users' cognitive and mental demand during practical use. Second, they help explaining phenomena occurring in human-computer interaction. For example, the sense of the relative location of limbs (BAS) hints towards a proprioceptive memory. This memory or muscle memory can provide additional kinesthetic cues to better memorize object identities and their locations in space.

2.2.4 Blended Interaction

As a ramification of the Reality-Based Interaction framework and the conceptual blending theory explained later (Fauconnier and Turner, 1998), Jetter et al. propose the Blended Interaction framework. They consider their Blended Interaction framework a "conceptual framework that helps to explain when users perceive user interfaces as ”natural” or not" (Jetter et al., 2014). Figure 2.13 (page 44) illustrates the different sources from which the Blended Interaction framework draws rationales to explain why user perceive UIs as "natural." "Blended Interaction is based on blends between concepts from the users' familiar reality, including already well-established digital concepts, and the expressive power of digital computation" (Jetter et al., 2014).

Conceptual Blending

To achieve this, Jetter et al. describe users' cognition and interaction with computing devices on the basis of conceptual integration. Conceptual integration or conceptual

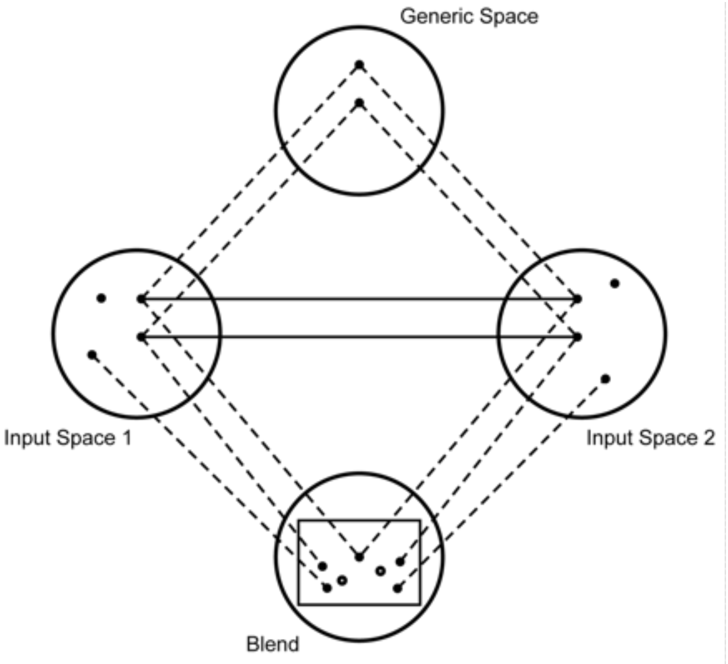


Figure 2.11.: A basic model of conceptual blending with two input spaces, a generic space, and the blend as resulting structure. (Adapted from (Fauconnier and Turner, 1998, p. 143)).

blending is a cognitive science theory first postulated by Fauconnier and Turner (Fauconnier and Turner, 1998). Fauconnier and Turner give examples for conceptual integration in mathematics and how humankind was able to develop an understanding of complex numbers through conceptual blending (Fauconnier and Turner, 1998).

Jetter et al. adopted this theory and applied conceptual blending in HCI. Our knowledge and interaction with computers base on complex and mixed concepts, which Fauconnier and Turner call "blends" (Fauconnier and Turner, 2003). A blend is the result of a conceptual integration of two or more input spaces from different domains. Figure 2.11 (page 40) illustrates a very simplified and basic model of conceptual blending. In this case, it relies on two input spaces: *Input Space 1* and *Input Space 2*. For example, input space 1 could be "one's personal music collection" (see Figure 2.12, page 41) and input space 2 could be a "Discman - a portable CD music player" (see Figure 2.12, page 41). Both input spaces are connected through shared structures indicated by solid lines. Fauconnier and Turner call these connections “cross-space mapping” (Fauconnier and Turner, 1998). They “become possible by means of a generic space. The generic space contains abstract information that is common to both the inputs.”. For instance, a common structure could be "a song title," or "a collection of songs grouped in an album." Each input space, however, has structures that do not exist in the respective other input space. For example, "one's music collection" can contain a vast amount of albums and songs — whereas a CD

Figure 2.12.: One's personal music collection with a magnitude of music CDs (left). A Sony diskman, a portable CD music player (right).

can contain a maximum of approx. 20 songs. A Discman is portable and lets one listen to music anywhere. A blend emerges from conceptually blending the shared structures of both input spaces, as connecting part, and additional non-overlapping structures of each input space. "This results in the blend's emergent structure that is more than a mere "cut-and-paste" combination" (Jetter et al., 2014).

Today, this blend is well-known as Apple iPod. Steve Jobs, the former CEO of Apple, used the following blend to announce the groundbreaking second model of the iPod: He said, an "iPod lets you easily put your entire music collection in your pocket and listen to it anywhere. With the new 10GB iPod, you can listen to your music continuously on six round-trip flights between San Francisco and Tokyo and never hear the same song twice."[18]. Using such a blend was a stroke of genius. It was a marketing clue and possibly one reason why people so quickly understood and adopted the iPod concept.

The iPod concept can be further deconstructed into a myriad of further and more basic-level concepts. For example, it also uses a hierarchical file system to manage songs sorted by artist or album. This blend is explained by Jetter et al. (Jetter et al., 2014). They argue, "the closer a concept is to existing concepts, in particular to the core basic-level bodily, spatial, or social concepts that most of us share since our

---

[18] Statement from Apple Website – `https://www.apple.com/pr/library/2002/03/20Apple-Introduces-10GB-iPod-2-000-Songs-in-Your-Pocket.html` (last accessed: March 21st, 2016)

childhood, the easier it is to integrate and apply" (Jetter et al., 2014). But blends can be arbitrarily complex and constructed of multiple input spaces. They even can be hierarchical and based on blends of blends of blends etc. Recent additions to Jetter et al.'s Blended Interaction framework and further arguments for conceptual blending in HCI argue for dynamics and multiplicity in blends (Bødker and Klokmose, 2016). Bødker and Klokmose state blends as dynamic constructs that change over time, e.g., through learning and experience (Bødker and Klokmose, 2016).

Potentials of Blended Interaction

Beyond Jetter et al.'s thorough review of theories from cognitive science, social sciences, and linguistics and applying the conceptual blending theory to HCI, their article on Blended Interaction is thought provoking in two ways and thus of fundamental relevance for work presented in this thesis. First, they raise the question "why we should bother with theory at all?" Second and most important they provide the four designs of domain *Individual Interaction*, *Social Interaction*, *Workflow*, and *Physical Environment*, which beyond being just a descriptive framework, allows designers to view a problem from different angles before writing any line of code or conducting any usability study.

Why Bother With Theory?

So, why should we "bother with theory at all?" Well, field studies, questionnaires & interviews with stakeholders and end-users, usability testing, and iterative & user-centered design have proven to be great tools for human-computer interaction. And every usability professional or researcher working in the field of usability experience (UX) design would agree that these tools are indeed excellent and above vital for the success of products. However, they only give an explanation to the *What?* and the *How?* Like *What is the problem?* and *How could it be solved?* But they do not explain the *Why?*. The Blended Interaction framework — as other theories — tackles this issue. It eventually provides guidelines that help to explain why interaction with a user interface feels ”natural" to the user. The Blended Interaction framework can also help to improve user experience while designing the interface in first place. Thereby it reduces risk (for companies to go down the wrong path) and in the long-run decreases efforts in usability testing and expensive re-iterations on product design.

Well-Established Digital Concepts

Concerning the second point, although the Blended Interaction framework draws from the Reality-Based Interaction framework there is a significant delimitation between the two. While the RBI framework only considers basic-level concepts from the real physical world, Jetter et al. also allow technologies or digital concepts "entrenched in conceptual structure" of users (Fauconnier and Turner, 1998) to be part of their reality. They call them "well-established digital concepts". These digital concepts become particularly important when technologies have become an integral part of everyday life and work, and have been used for years or even decades. For example, the floppy disk icon and its accompanying *Save* action. The icon design originates from early personal computing when floppy disks were broadly used as persistent storage devices. With the proliferation of new storage hardware and hard disks with storage capacity far beyond of that times imagination, floppy disks vanished, but the icon kept the same. Many users who never got in touch with floppy disks or users of a younger generation who grew up years after floppy disks were banned might not even know about the icon's origin. But they associate the *Save* action with it. Another and more recent example is the burger icon, which is used in many responsive websites to open a navigation menu. Jetter et al. eventually come to the conclusion that real and digital concepts should not be regarded separately from one and another (Jetter et al., 2012c). Instead, digital technologies and "well-established concepts" (Jetter et al., 2014) will be considered as of equal importance and reality when designing new UbiComp experiences.

Domains of Design as Tool for Design

To help with the design of post-WIMP interfaces, the Blended Interaction framework further comes with four domains of design. These domains of design serve as a tool for creating post-WIMP UbiComp experiences and will be respected to develop "natural" user interfaces. The four domains of design *Individual Interaction*, *Social Interaction & Communication*, *Workflow*, and *Physical Environment* are presented in the next paragraphs.

Individual Interaction: No matter of the type of interactive work — either solitary or collaborative work — interaction with a system always composes individual actions like touch input, mouse input, gestures, speech, or even multimodal input (Jetter et al., 2014). In 1998, Gutwin and Greenberg raised concerns when designing interactive groupware. They state that an "ideal solution [. . .] would be to [. . .] support both the needs of individuals and the needs of groups" (Gutwin and Greenberg, 1998). In turn, even if an interactive system is supposed to support co-located

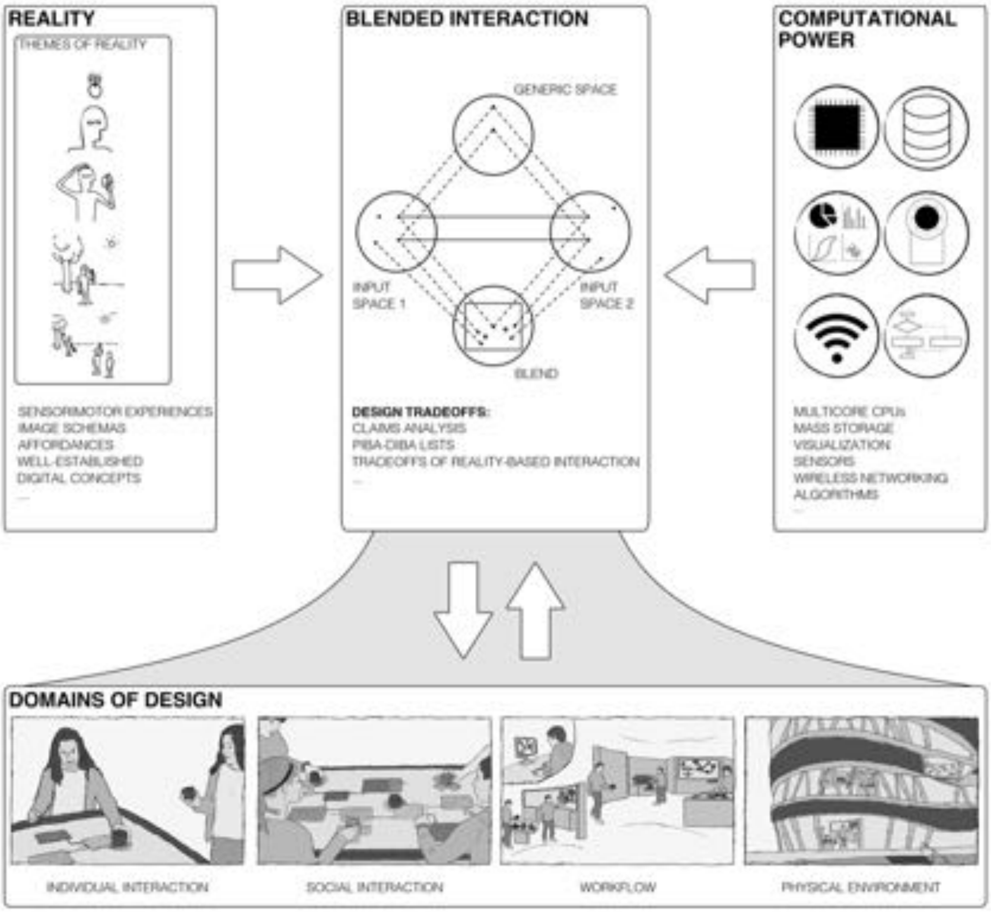


Figure 2.13.: An overview of the Blended Interaction framework. The four domains of design of Blended Interaction (bottom, from left to right) *Individual Interaction*, *Social Interaction (& Communication)*, *Workflow*, and *Physical Environment* (Jetter et al., 2014).

collaboration, the individual interaction should be taken into account, and eventually tradeoffs between power for the individual and overall group awareness should be considered.

Social Interaction & Communication: In addition to supporting individual interaction, it is important to consider group interaction (Gutwin and Greenberg, 1998) and the social aspects and norms applied to mitigate between group members. Ideally, a system has an understanding of these norms and assists groups when appropriate. This aspect emphasizes on Jacob et al.'s theme Social Awareness & Skills (Jacob et al., 2008). To this end, Jetter et al. recommends the consideration and investigation of social processes that take place in groups during the collaboration (Jetter et al., 2012a). For example, Jetter et al. state Tangible User Interfaces (TUI) as a successful interaction approach promoting social interaction and communication (Jetter et al., 2014).

Workflow: Often, research in human-computer interaction focusses on granular interaction techniques looking at their performance isolated from any coherent workflow. While this is tolerable for conducting controlled lab experiments to receive results with high internal validity, such might lead to disruptive interactions when designing an interactive system that by nature needs to support a workflow. Therefore, Jetter et al. consider the process as a whole rather than only focussing on an atomic interaction concept.

The evidence is provided by Sellen and Harper. In their awarded book "*The myth of the paperless office*" (Sellen and Harper, 2001), they report on studies on interactive systems, e.g., systems that re-implemented non-digital workflows in air traffic controlling or police procedures. The interactive systems were meant to replace traditional pen & paper work. However, most systems failed as they introduced confounding factors such as a decreased mutual trust because of a laptop screen suddenly between the police officer and a claimant. Sellen and Harper show that new computing systems with the insufficient support of existing workflows are prone to fail. Most of these systems failed because the traditional workflow was not considered appropriately during systems' design. In consequence, their users returned to the old and non-digital tools. However, they also report on one success story. Thereby, the company employing the digital system changed and adapted to a new workflow rather than sticking to the old non-digital workflow (Sellen and Harper, 2001). Or more recently Greenberg et al. show a very compelling system supporting an entire workflow. Their Proxemic Media Player is a context-aware media player that appropriately mediates between two users concurrently interacting with the same screen (a television screen) (Greenberg et al., 2011).

Physical Environment: The fourth design domain deals with the architecture of the physical space. Architecture and interior design determine and shape interaction with our physical and social environment. Often, the design of objects determines their function and use. For example, physical affordances of chairs and tables evolved over decades and their predetermined form and shape define how we use them (Hajizadehgashti, 2012) for the purpose of eating but also to conduct group meetings. Streitz et al. have discovered the importance of architecture for the design of information spaces. Their i-LAND environment constitutes "several 'roomware' components" (Streitz et al., 1999) for computer-supported individual and collaborative work. These roomware components are integrated into the architecture of a room. For example, the DynaWall®, an "interactive electronic wall" (Streitz et al., 1999) that allows a team or group of people to work in mixed-focus collaboration and seamlessly switch between loosely-coupled parallel work and tightly-coupled collaboration (Tang et al., 2006). Jetter et al. recommend a blend between the nature of objects in the room (e.g., tables and walls) and novel post-WIMP technologies (e.g., the projection of content, tracking of users, or even deformable displays) (Jetter et al., 2014).

2.2.5 Summary of Theories in HCI

All theories, models, and frameworks presented in this section build on the embodied view of cognition and the tight interplay between human mind and body. They implement the theoretical foundation for this thesis and were selected because they best resonate with the vision of UbiComp where computing technologies “weave themselves into the fabric of everyday life” (Weiser, 1991). It is hoped that interaction with new UbiComp systems feels “natural” (Jetter et al., 2014) and becomes a familiar everyday experience (Jacob et al., 2008). This foundation will be heavily used to design new UbiComp experiences. As another important — yet missing — building block we have to understand where users maintain experience and from what source experience is drawn.

2.3 Human Memory

Memory is our mental ability to retain and recall from the experience and is based on mental processes of learning, retention, recall, and retrieval. Through mental processing, we encode a physical stimulus into a form that our brain’s memory system can interpret and use. For example, a stimulus is a sensory information captured by visual or auditorial perception but also from kinesthetic cues through physical activity (Tan et al., 2002; Andrews et al., 2010; Jetter et al., 2011) — as we know from recent psychological advancements and the embodied view of cognition (Scott et al., 2001).

An encoding can be acoustic, visual, or semantic. Acoustic encoding includes sounds or spoken words. Visual encoding can be images or mental snapshots. For instance, a snapshot of your kitchen often used for mnemonic techniques like the method of loci[19]. Semantic encoding contains general meaning like concepts or ideas (e.g., mathematical problems).

Memory takes encodings and stores them as memories in the brain for later recall. Memory can be episodic, procedural, or semantic. An episodic memory encodes

---

[19] The method of loci is also known as memory palace or mind palace. It is based on the assumption that users can best remember familiar places like their homes and objects in it. The technique requires a user to mentally construct a path with waypoints through this location. Their route, then, can be used to mentally associate items to remember to each waypoint by assembling a fictive story.

information about a specific event (e.g., a memory about last birthday). A procedural memory encodes information about how to do things (e.g., how to walk). A semantic memory encodes general concepts. For example, the concept of gravity, a solution to a mathematical problem, or facts about the world. Importantly, stored information can be recalled, which allows us (i) to orient and navigate in a world and (ii) to do higher-order learning (Ohlsson, 1995) or knowledge work.

A rich set of memories and experiences is exceedingly important for research presented in this thesis. First, it is beneficial for conceptual blending. It allows us to perform the act of conceptual integration since memory is the source pool of basic-level and abstract concepts (generic space). Second and as explained later, it enables us to store spatial locations of objects located in the physical world and quickly recall them from so-called spatial memory instead of requiring us to visually scan our environment.

2.3.1 Models of Human Memory

Human memory has been a source of research for centuries. Only recently cognitive psychologists and neuroscientists began to understand how parts are reassembled into a whole. They proposed and tested various models on human memory and critically reflected and discussed their strengths and limitations in the scientific community. For example, Atkinson and Shiffrin's multi-store model (Atkinson and Shiffrin, 1968), Craik and Lockhart's model of levels-of-processing (Craik and Lockhart, 1972), or Baddeley and Hitch's model of working memory (Baddeley and Hitch, 1974).

Working Memory Model

This work relies on Baddeley and Hitch's working memory model, which is one of the well-established models for human memory. In their opinion, the multi-store model (Atkinson and Shiffrin, 1968) was far too simple as it describes memory stores as unitary systems not allowing for parallel processing of sensory information. For example, the multi-store model would not allow for parallel processing of multimodal input such as listening to a talk while at the same time visually perceiving images on slides of the very same talk.

Therefore, Baddeley and Hitch argue for a modification and subdividing stores in further independent and parallel working systems, which they propose in their model of *working memory* (see Figure 2.14). The working memory model, in contrast to most other human memory models, is based on a solid foundation of clinical and

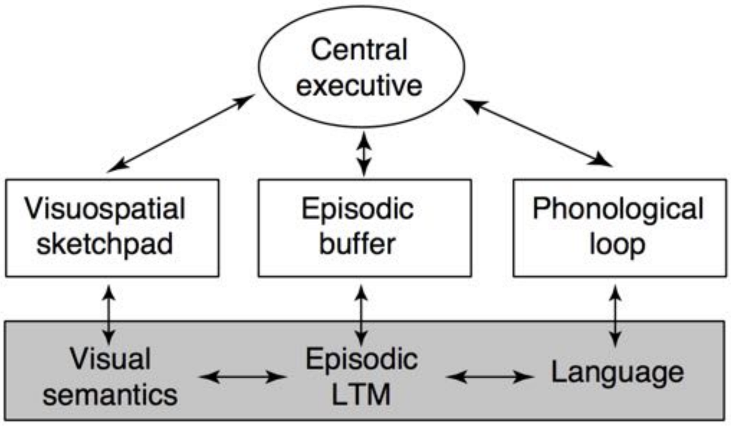


Figure 2.14.: Working memory model with all four components: central executive, visiospatial sketchpad, episodic buffer, and phonological loop (Baddeley, 2000).

experimental evidence gathered throughout the 1960s and 1970s. Clinical studies provide a high ecological validity and allow for generalization of results whereas the experimental results have a high internal validity tested through very controlled studies.

They found in various experiments that multiple systems must be involved in order to allow concurrent processing of sensory input, which happens on a daily basis, e.g. when watching a movie on TV. In this example, one perceives a sequence of images and at the same time listens to audio. A unitary system like the multi-store model would only allow for sequential processing. In the working memory model, however, there are different systems for different types of information. Working memory consists of a *central executive* which manages and coordinates three subsidiary and in parallel working slave systems: the *visuospatial sketchpad*, the *episodic buffer*, and the *phonological loop*.

The TV example above involves all three subsystems. The phonological loop recognizes phonemes, assembles a sequence of them into words, and finds their meaning in long-term memory. At the same time, the visuospatial sketchpad associates seen images with concepts and retrieves their meaning also from long-term memory. The episodic buffer assembles both information into a coherent story, which ideally adheres to the movie plot. Important for work in this thesis is the visuospatial sketchpad as in addition to process visual information, it also processes spatial information.

2.3.2 Spatial Memory

The visuospatial sketchpad, sometimes referred to as the "*inner eye*", “is assumed to hold visuospatial information, to be fractionable into separate visual, spatial and possibly kinaesthetic components” (Baddeley and Hitch, 1974). This visuospatial information is important for navigation. For instance, when finding a way to rescue oneself from a building on fire. Jeanne Sholl calls this a “navigational tracking device in humans” but often and henceforth called spatial memory.

Spatial memory is built by path integration or dead reckoning, which is the constant “process of updating the body’s location and heading on the basis of the velocity and acceleration signals produced by self movement” (Jeanne Sholl, 2001). Such “self-locomotion and spatial exploration” (Oudgenoeg-Paz et al., 2014) is “thought to be especially important for advances in spatial cognition” (Oudgenoeg-Paz et al., 2014). Humans “learn the locations of objects in a new environment, they interpret the spatial structure of the layout in terms of a spatial reference system” (McNamara and Valiquette, 2004).

Sholl and Nolin differentiate between two referential modes of coding spatial relations: “object-to-object relations in environment-centered coordinates” (Sholl and Nolin, 1997) (allocentric) and “self-to-object relations in body-centered coordinates” (Sholl and Nolin, 1997) (egocentric). Scarr et al. further differentiate between three different frames of reference for spatial information: *inside objects’s boundaries*, *side-connected*, or *landmarks*. Spatial information “can be encoded in object space, for example, relative to the object’s envelope (inside the object’s boundary), its principal axis (side-connected), or to other objects and external reference points (landmarks)” (Scarr et al., 2013).

Spatial memory allows us to navigate in well-known areas without deciphering cartographic maps, which otherwise would occupy valuable cognitive resources. Spatial memory assists humans to navigate in their homes, even in the dark when only kinesthetic cues are available for navigation. However, in this example, input modalities are reduced, and as commonly known, navigation in the dark often happens with both cautious and conscious movements actively acquiring many cognitive resources. In contrast, navigation in light often happens beneath humans’ conscious awareness. For a similar reason, human-computer interaction stimulating multiple human modalities “minimizes users’ cognitive load, which effectively frees up mental resources for performing better while also remaining more attuned to the world around them” (Dumas et al., 2009). Presumably, such unallocated mental resources can be either used for other tasks or stay unused, which eventually reduces effort or frustration.

In HCI, spatial memory provides similar advantages as in the real physical world (Scarr et al., 2013). A “strong spatial knowledge of interface layouts and control locations, particularly in graphical user interfaces, allows users to substantially reduce the cognitive and physical effort required for interaction” (Scarr et al., 2013).

2.3.3 Studies of Spatial Memory in HCI

An early and well-known study of spatial memory in HCI is Data Mountain that concentrated on the effect that visualizations had on users’ spatial memory performance in a document management task (Robertson et al., 1998). Robertson et al. compared the management of browser bookmarks with a traditional browser (Microsoft Internet Explorer 4) and the Data Mountain interface that allowed users to arrange thumbnails of bookmarks in a virtual 2.5D space (see Figure 2.15, page 50). The results show that, on average, users recalling bookmarks quicker and more reliable when using the spatial interface and hinted at the importance of designing for better spatial memory in HCI.

Figure 2.15.: Data mountain interface. A 3D visualization to spatially manage browser bookmarks (Robertson et al., 1998).

According to embodied views of cognition, memory performance can be influenced by body movements and recent “findings from Embodied Cognition reveal strong effects of arm and hand movement on spatial memory” (Jetter et al., 2012c). For example, a study by Tan et al. (Tan et al., 2002) reported a 19% better spatial memory performance in favor of touch input compared to mouse input when recalling objects that have been positioned by the users in a prior task. Tan et al. reason that the increase of precision in the spatial recall task is attributable to “kinesthetic cues” (Tan et al., 2002).

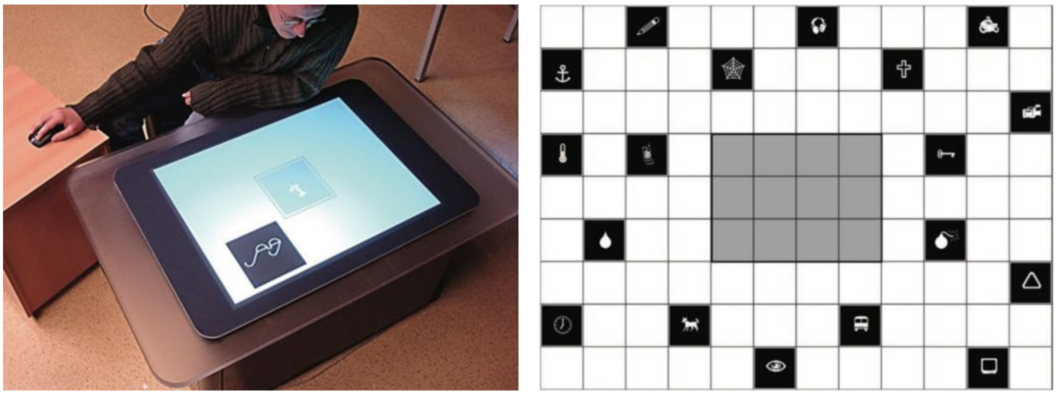


Figure 2.16.: Study setup of Jetter et al.'s comparative study on spatial memory in mouse condition (left). The large information space with memory cards participants had to find and navigate to (right). (both images adapted from (Jetter et al., 2012c, p. 85))

Inspired by these findings, Jetter et al. conducted two sequential experiments in 2012 that investigated the effect of touch vs. mouse input on navigation performance and users' spatial memory for panning user interfaces (panning-only UI) and zoomable user interfaces (ZUI) (see Figure 2.16, page 51). In their first experiment, they found that multi-touch interaction for a panning-only UI on a tabletop leads to better navigation performance as well as better spatial memory compared to traditional interaction with a mouse. Conversely, the second experiment does not show such significant differences for a ZUI. The navigation performance is even worse than with a mouse (Jetter et al., 2012c).

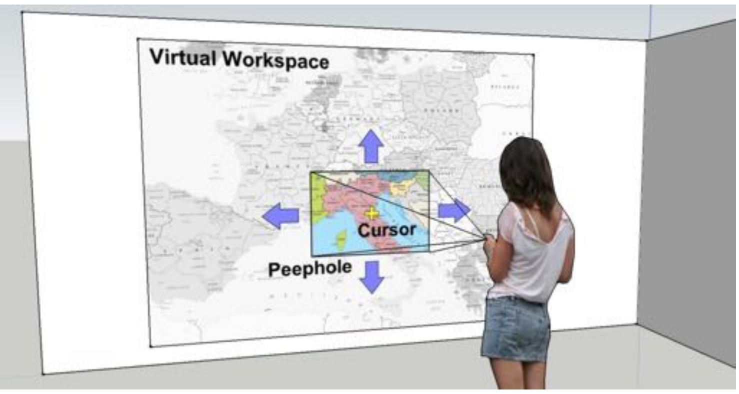


Figure 2.17.: Projector-phone peephole where physical movement of a phone with an integrated projector reveals content of a situated information space projected on a wall (Kaufmann and Ahlström, 2013, p. 3173).

Kaufmann and Ahlström take Jetter et al.'s concept even further. Kaufmann and Ahlström (Kaufmann and Ahlström, 2013) conducted a similar study on navigation performance and spatial memory on projector phones with physical peephole navigation (see Figure 2.17, page 51). Peephole navigation with a handheld and spatially-aware device was originally conceived by Fitzmaurice in 1993 (Fitzmaurice, 1993) and will be explained in the next chapter.

Figure 2.18.: Situated information spaces like cartographic maps explorable by spatially-aware palmtop computers (Fitzmaurice, 1993, p. 45).

Important here is that research conducted by Kaufmann and Ahlström shows great potential for spatial navigation in terms of recalling yet invisible object locations from memory. Their findings are inline with the assumption that physical exploration and navigation through body moves "may trigger automatic self-to-object updating" (Sholl and Nolin, 1997) and "naturally" build spatial memory. Research presented in this thesis relies on this assumption and previous related work. It sheds light into a subconscious use of spatial memory also indirectly impacting users' subjective workload like mental demand, frustration, and effort.

2.4 Summary

This chapter narrows down the research context to HCI, in particular, research in the context of UbiComp. It elaborates on contradicting opinions of HCI researchers and thereby unfolds opportunities for further research with a focus on novel UbiComp experiences. Recent and emerging theories, models, and frameworks in HCI were presented. They build the theoretical foundation of this thesis and will guide the design of novel UbiComp experiences. Hereby, they particularly provide guidance exploiting users' pre-existing knowledge of everyday life and other inherent (cognitive) capabilities such as spatial memory. For instance, by applying conceptual blending theory with a strong focus on Reality-Based Interaction tradeoffs or by considering the four domains of design from Blended Interaction.

Parts of the next Chapter 3 appear in the following publications:

Rädle, R. (2013). “Design and evaluation of proxemics-aware environments to support navigation in large information spaces”. In: *CHI ’13 Extended Abstracts on Human Factors in Computing Systems on - CHI EA ’13*. New York, NY, USA: ACM, p. 1957. DOI: 10.1145/2468356.2468710

Reiterer, H. **Rädle, R.** Butscher, S. Müller, J. (2016). “Blended Library – neue Zugangswege zu den Inhalten wissenschaftlicher und öffentlicher Bibliotheken”. In:[20] *Bibliothek Forschung und Praxis* 40.1. DOI: 10.1515/bfp-2016-0010

Kleiner, E. **Rädle, R.** Reiterer, H. (2013). “Blended Shelf: Reality-based Presentation and Exploration of Library Collections”. In: *CHI ’13 Extended Abstracts on Human*[21] *Factors in Computing Systems on - CHI EA ’13*. New York, New York, USA: ACM Press, p. 577. DOI: 10.1145/2468356.2468458

Gebhardt, C. **Rädle, R.** Reiterer, H. (2014c). “Integrative workplace: studying the effect of digital desks on users’ working practices”. In: *Proceedings of the extended abstracts of the 32nd annual ACM conference on Human factors in computing*[22] *systems - CHI EA ’14*. New York, New York, USA: ACM Press, pp. 2155–2160. DOI: 10.1145/2559206.2581186

---

[20] The responsibilities for this joint publication were divided as follows: Harald Reiterer spearheaded the writing. I contributed to writing and the report of qualitative findings resulting from questionnaire and interviews as well as writing of system descriptions for Blended Shelf and TwisterSearch. Simon Butscher and Jens Mueller contributed to writing on systems for public libraries.

[21] The responsibilities of this joint publication were divided as follows: Eike Kleiner spearheaded the writing of the paper and implemented the prototype. I contributed to the writing and implemented the data backend. Harald Reiterer and I gave advise and contributed to the concept.

[22] The responsibilities of this joint publication were divided as follows: Christoph Gebhardt spearheaded the writing of the paper, implemented the prototype, and conducted the user study. I contributed to the writing. Harald Reiterer and I contributed to the study design and supervised the project

Context & Analysis

3

> *I never guess. It is a capital mistake to theorize before one has data. Insensibly one begins to twist facts to suit theories, instead of theories to suit facts.*
>
> — **Sir Arthur Conan Doyle**
> (Author of Sherlock Holmes stories)

Contents

This chapter narrows down the application context to academic libraries. Therein, it is concerned with knowledge work and activities surrounding it. Qualitative research methods were applied to gain a deeper understanding of knowledge work activities and to identify real-world problems of knowledge work within academic libraries.

These problems inspired concepts and implementation of two research prototypes *Blended Shelf* and *Integrative Workplace*. Both research prototypes were evaluated in "the wild" to refine the problem space and derive specific research questions.

The theoretical foundation of the previous chapter and the qualitative research presented in this chapter build the scaffolding of this thesis. They also inform the study designs of research presented in Chapters 4-7.

3.1 Libraries

Digital libraries are growing rapidly, and literature research is transitioning more and more to the world wide web. Physical libraries, like public, academic, or special libraries, cannot compete with the internet and its massive amount of information being pervasively available at users' fingertips. Their primary reason for existence as such — being a pure information storage — is fading away.

John Welford, a librarian, publicly criticizes this trend. He argues that libraries "are under threat, mainly because the people who fund them are under the mistaken impression that they are no longer needed in the age of the Internet."[1] He further writes: "I used to be a full-time librarian, but I lost my job in 2002 for that very reason. The company that employed me took the view that because it was "all on the Internet" there was no reason why they should employ somebody to do what everybody could do for themselves from their desktop."[1]

A few libraries are more opportunistic in that respect. They see this trend as a chance for themselves to change. An opportunity to clear out the dust traps from archaic perspectives of what a library is. They discovered the trend early enough and proactively shift away from being an information provider. These libraries desire new qualities and prospect the place library as a place to learn, to meet & greet, and importantly a happy place to joyfully carry out information literacy. Q-thek and Blended Library, as two example projects, make the effort moving libraries from pure information storages to places for people with qualities as mentioned earlier.

3.1.1 Q-thek for Public Libraries

The project "*Lernort Bibliothek - zwischen Wunsch und Wirklichkeit*"[2] (English Translation: Learning Place Library - between desire and reality) is an initiative of the state North Rhine-Westphalia (NRW). It is centered around *learning*, learning as a lifelong experience. Together with eight public libraries in NRW and the architects

---

[1] Blogpost by John Welford on Hubpages – `http://hubpages.com/literature/The-three-main-types-of-library`

[2] Lernort Bibliothek - Q-thek Concept – `http://www.brd.nrw.de/schule/privatschulen_sonstiges/pdf/Lernort_Bibliothek_Q-thek_-_innovative_Bibliotheksr__ume_2011_04_20.pdf` (last accessed: March 16th, 2016)

Reich & Wamser GbR, they designed an architectural concept that supports the five aspects *presentation*, *relax*, *learn*, *communication*, and *inform*. Among new library services such as seminars, the focus for Q-thek was primarily put on presentation of library collections and multi-functional furniture. For example, the guard post in Figure 3.1 (page 59) (left) sets new publications on stage.

Figure 3.1.: Elements of Q-thek. A guard post that sets new publications on stage (left) and a space for group work (right).

In the core vision of Q-thek, the library becomes a multi-functional place where visitors can browse for literature and quietly read books but also work in groups (e.g., school work), surf the internet, or even relax. As an addition to the furniture, the libraries also extended their service portfolio, to train non-digital natives (e.g., students and elderly people) in “literacy technology” (Weiser, 1991) by exposing them to and teaching them how to handle state-of-the-art information and communication technology.

3.1.2 Blended Library for Academic Libraries

The project *Blended Library*[3], as another example, aims at the development of new concepts for supporting research and knowledge work processes in academic libraries. Similar to Q-thek, Blended Library considers the environment and architecture as important aspect to facilitate both, individual and group work. Figure 3.2 illustrates a room concept with interactive tabletops for information & literature search supporting mixed-focus collaboration (Tang et al., 2006) (e.g., (Jetter et al., 2011; Rädle et al., 2012a; Rädle et al., 2013b)), a work desk to interchangeably work with digital and analogue media (e.g., (Gebhardt et al., 2014a; Gebhardt et al., 2014c)), a large wall-sized display for brainstorming and group-discussion, and a place to relax and conformably read and reflect (e.g., (Rädle et al., 2011b)) (Scan QR Code #1 or enter “1” in the *MediaBrowser* application).

---

[3] Blended Library project website – `http://www.blendedlibrary.org` (last accessed: March 16th, 2016)

Video

Code #1

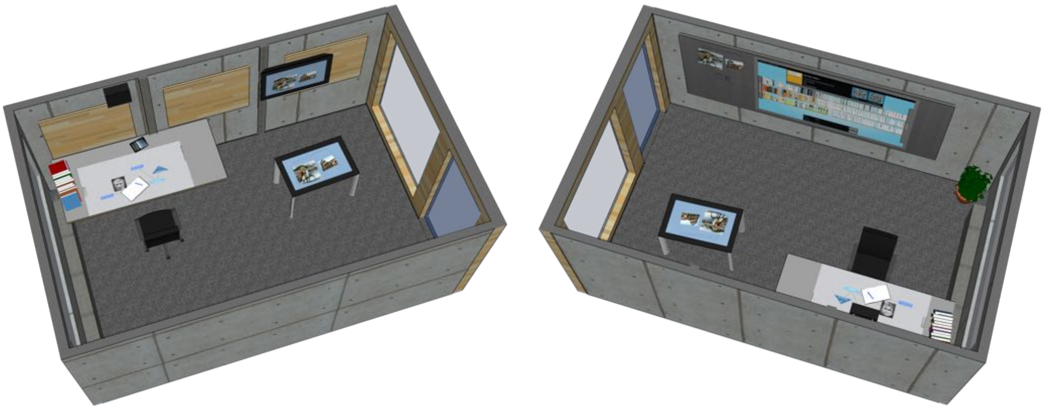


Figure 3.2.: Sketched 3D model of a physical environment that supports a variety of academic work activities, which range from individual to group work.

In contrast to Q-thek and as a second pillar, Blended Library[4] also focusses on human-computer interaction and the development of novel interactive systems and library services to better support academic work activities. Because nowadays, library users are still faced with traditional terminals and Online Public Access Catalogues (OPAC) to start literature research. For relevant printed literature, they often need to transcribe call numbers[5] from terminal screens to paper. For reasons of media continuity, they have to switch and find the book at its physical location on the shelf. Even worse, making annotations while reading a book is strictly prohibited by library rules, so the reflection on book contents and sensemaking is often disrupted and happens on different media. When sensemaking is not carefully practiced, sharing such information with co-workers is worthless because it is hardly possible to trace back to the source and verify a knowledge worker's thoughts and arguments.

The two examples are seed motivation for improving knowledge work in academic libraries. They clearly pinpoint two current problems in knowledge work, namely literature search and cross-document referencing. At the same time, they narrow down the application context of this thesis to academic libraries and therein to academic work activities.

To understand the problem space in more detail, we conducted several qualitative research studies to shed light into current knowledge work practices in academic libraries and to find potentials for improving them. Of course with a special focus on literature search and cross-document referencing.

---

[4] Libros, a follow-up project of Blended Library, similarly focusses on HCI in libraries but for public libraries.

[5] A call number is an ID that also often encodes the physical location of media in a (book)shelf.

3.2 Field Studies and Problem Space

We applied qualitative research methods to gain a deeper understanding of knowledge work activities in an academic context. This understanding includes insights on users' current working practices and to help identify therein prevalent problems. Such research is an essential part of the *Analyze* phase of almost any kind of usability lifecycle (Hartson and Pyla, 2012, p. 53) (see also (Mayhew, 1999; Rosson and Carroll, 2002; Deutsche Akkreditierungsstelle Technik, 2009)). Rogers et al. state "the purpose of data gathering is to collect sufficient, accurate, and relevant data so that a set of stable requirements can be produced" (Rogers et al., 2011, p. 222).

Since no data gathering technique is perfect and has its strengths and weaknesses (Butz and Krüger, 2014, pp. 109-120), we applied triangulation. Triangulation means to investigate in a phenomenon from multiple perspectives to sharpen understanding and rely on emerging problems (Lazar et al., 2010). Rogers et al. enumerate four different types of triangulation (Rogers et al., 2011):

1. Triangulation of data means that data is drawn from different sources at different times, in different places, or from different people (possibly by using a different sampling technique).
2. Investigator triangulation means that different researchers (observers, interviewers, etc.) have been used to collect and interpret the data.
3. Triangulation of theories means the use of different theoretical frameworks through which to view the data or findings.
4. Methodological triangulation means to employ different data gathering techniques.

Data triangulation applies to all empirical findings presented in this chapter. For instance, gathering qualitative data through different methods like online & paper questionnaire, interviews, focus groups, and observations. Studies were conducted with students and academic staff of the University of Konstanz as representatives for academic knowledge workers. Triangulation of theories is mainly applied in research Phase 2 (see Chapter 1.4, page 13) in discussion sections of experimental findings.

In the following, we first present our procedures for each data gathering technique and then conflate findings altogether in the next section *Chapter 3.2.4 – Empirical Findings*. As part of the analysis and based on findings of empirical data, we implemented two research prototypes *Blended Shelf* and *Integrative Workplace*. The prototypes are fully functional systems deployed and tested with knowledge workers. They helped to identify particular problems and operationalize research, which will be part of research presented in Chapters 4-7. Figure 3.3 provides a structured overview of the empirical research presented in the following sections.

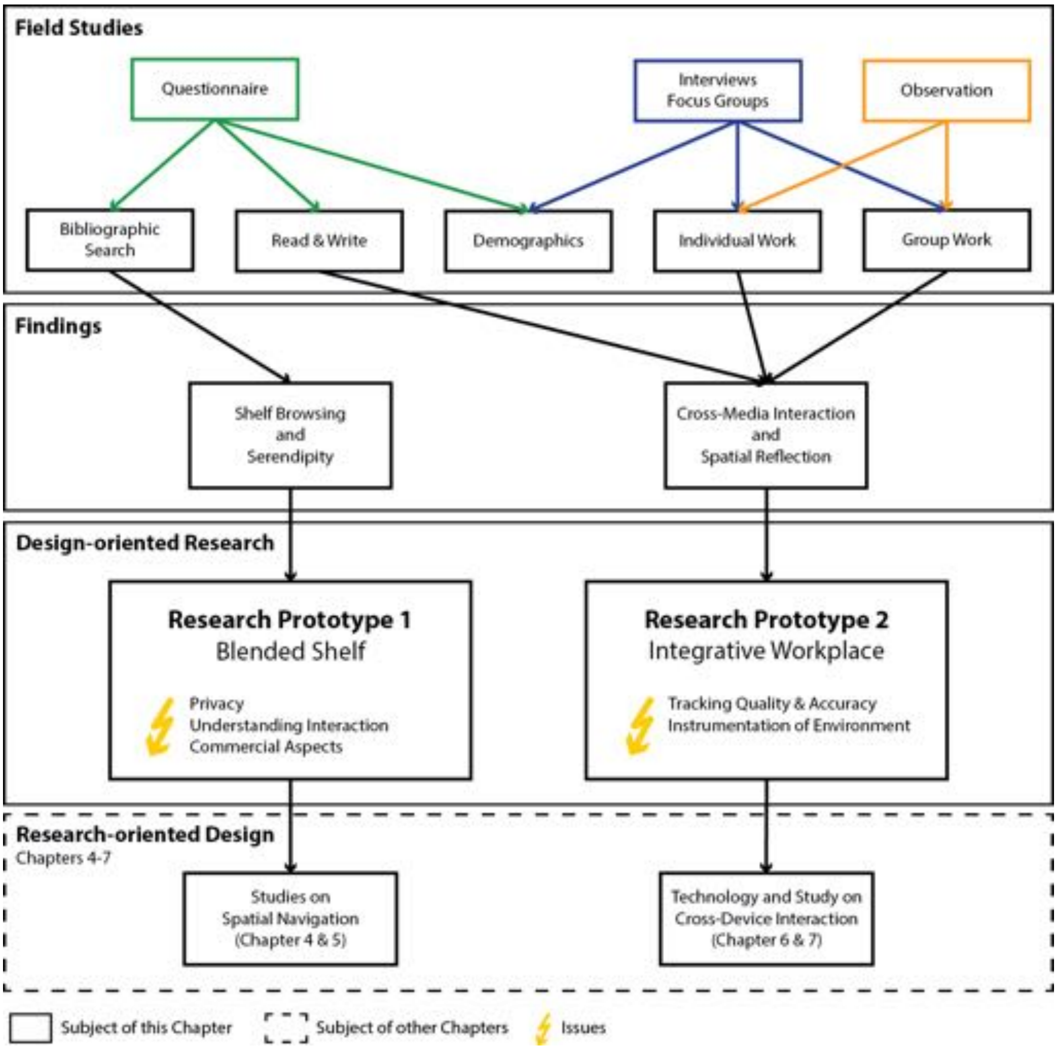


Figure 3.3.: Overview of empirical research consisting of field studies and their findings. The findings inspired the design and implementation of the two research prototypes Blended Shelf and Integrative Workplace. In a design-oriented research, the two prototypes helped to identify particular research questions, which are answered in controlled experiments and through implementation of enabling technology in Chapters 4-7.

3.2.1 Online & Paper Questionnaire

As a starting point, we distributed a questionnaire to students and researchers from the University of Konstanz to gain an overview of tasks and goals that occur during knowledge work activities. The questionnaire was available for eight weeks (57 days) from December 22, 2011, to February 16th, 2012. We deliberately chose this period, which included teaching time, exam time, as well as a study break. This mix covers a great variety of different academic tasks such as solving class assignments (teaching time), preparation for exams and presentations (exam time), as well as writing essays (study break). The questionnaire was provided online and also "offline" as paper questionnaire and was available in German and English language. To increase the likelihood of reaching not just technology-savvy but also technology-averse users, we also deployed a paper questionnaire. The paper questionnaire was also available in both languages. We provided cardboard boxes for anonymous submissions at the entrances and exits of all library buildings. Having both — technology-savvy and technology-averse users — participating in the questionnaire eventually provide a broader spectrum of insights and issues surrounding UbiComp technology. This approach also helps to reveal use cases that are currently neglecting technology. As an incentive to fill out the questionnaire, participants could also take part in a lottery and win one of three BIBBAGs[6] or one of 25 coffee vouchers.

Questionnaire Structure

The questionnaire consisted of 34 questions and was structured in three parts.

1. Demographics (10 questions)
2. Literature & Bibliographic Search (8 questions)
3. Academic Reading & Writing (16 questions)

Its structure is closely related to O'Hara et al.'s model of document related activities of library users (O'Hara et al., 1998). However, O'Hara et al.'s report on document-related activities dates back to 1998. We conducted our questionnaire for two reasons. First, to have a fresh look on library user activities. Second as presented in *Chapter 2.1 – The UbiComp Trend*, the computing device landscape changed since then, which might have had an impact on how knowledge work is conducted nowadays.

In the literature & bibliographic part, we were interested in the media types that are used for academic work and the sources and tools that are used to acquire media. In the reading & writing part, we were interested in the tools used for academic work. Also in this part, we were interested why particular tools are used for certain tasks. We had participants to reflect on benefits over other tools or possible drawbacks they face during use. Also, participants could also comment on tasks that were not listed in the questionnaire. At the end of the questionnaire, they were also

---

[6] BIBBAGs are transparent bags allowed in the Library of the University of Konstanz. Due to book theft protection and because of their transparent material library staff can quickly inspect bag content. – `http://www.bibbag.de` (last accessed: March 9th, 2016)

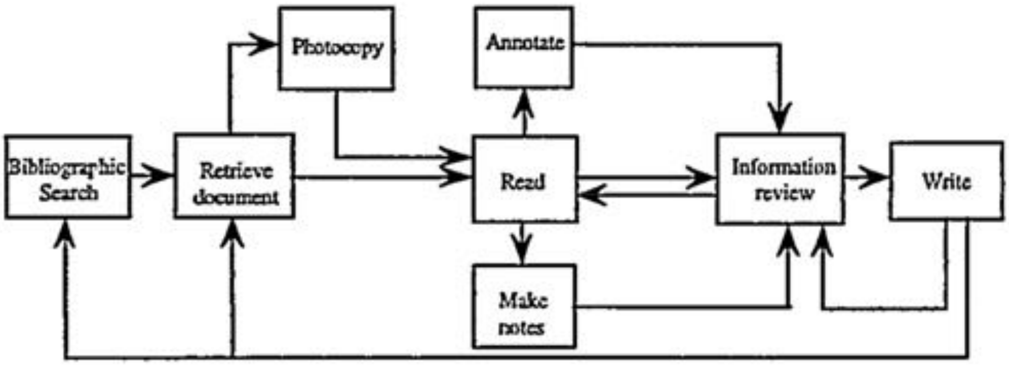


Figure 3.4.: O'Hara et al.'s model of document related activities of library users. (adapted from (O'Hara et al., 1998, p. 234))

encouraged to propose ideas and thoughts on task improvements or tools to better fit their current needs and requirements. The complete questionnaire is in *Appendix B – Paper Questionnaire*.

Validity & Respondence Rate

A total of 746 people responded to the questionnaire whereas 682 count as valid responses after pre-processing data (599 students, 83 academic staff). The control sample achieved 5.48% (students only[7]) when compared to enrolled students for the given study timeframe. The response rate is a great achievement compared to usual respondent rates of 1% to 2% (Butz and Krüger, 2014, p. 115). The sample size is furthermore representative when control sample and total population is divided into and compared to all 13 departments of the University of Konstanz (see Figure 3.5, page 65) or the students' respective targeted degree (see Figure 3.6, page 66). An illustration of gender comparison is omitted but reflects similar gender distribution when compared to the total population. Participants mean age is 24.5 years ($\tilde{x}$=23, SD=6.48).

---

[7] The University of Konstanz only publishes statistics about students. Statistics about the academic staff was not available at the time of writing this thesis.

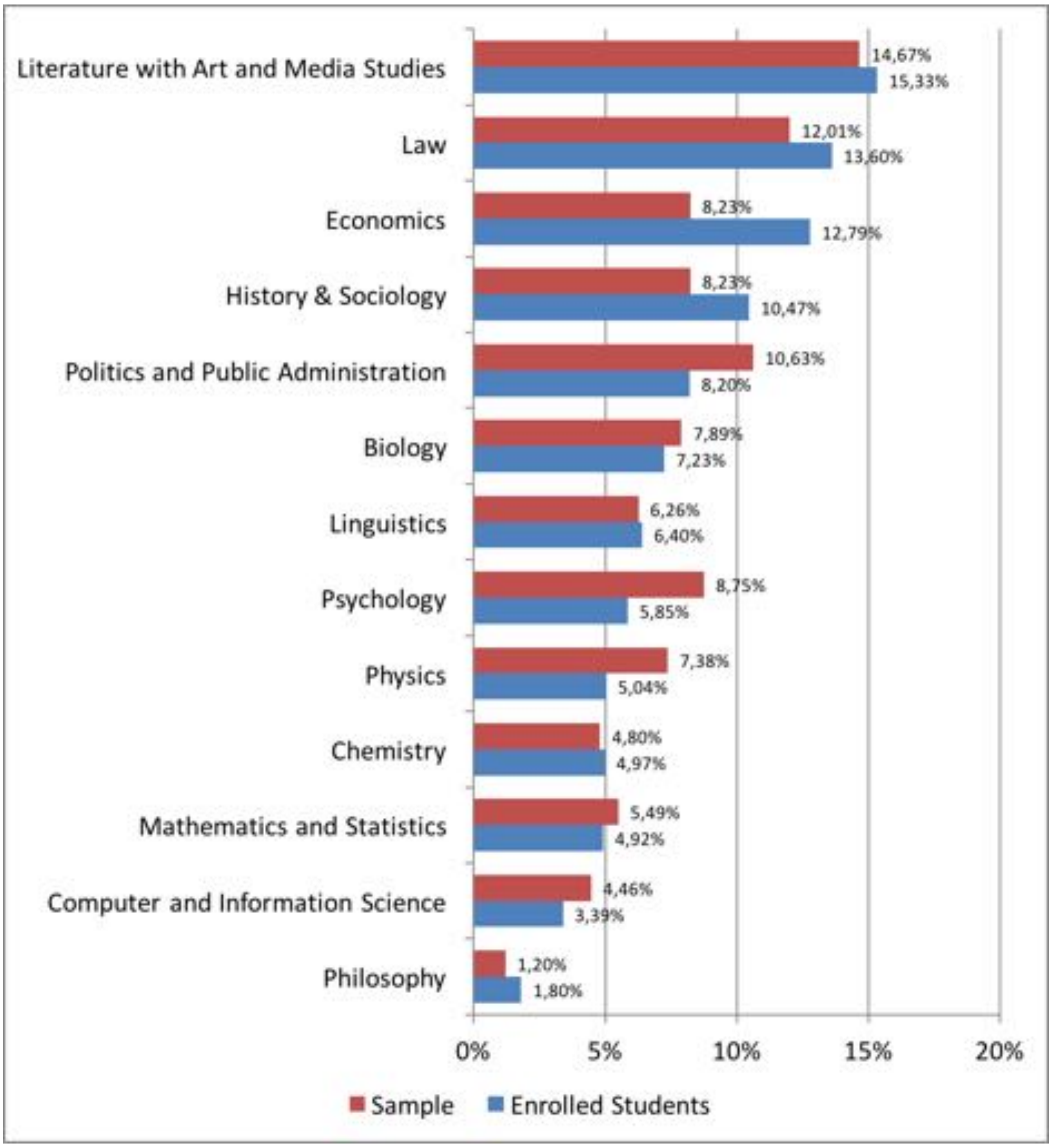


Figure 3.5.: Comparison of achieved sample size and target population divided into all 13 departments of the University of Konstanz.

Analysis of Open Questions

We first analyzed open questions using a topic analyzer[8] to cope with the vast amount of responses. The topic analyzer is based on MALLET[9], which implements a sampling-based version of LDA[10]. However, results of the LDA and returned topics were unsatisfying due to mostly short sentences that were given in response. So we

---

[8] Topic Analyzer – `http://analysis.blendedlibrary.org` (last accessed: March 30th, 2016)

[9] MALLET is a toolkit for statistical natural language processing, document classification, clustering, topic modeling, and information extraction. – `http://mallet.cs.umass.edu` (last accessed: March 30th, 2016)

[10] Latent Dirichlet Allocation (LDA) is based on a statistical approach to extract topics from a given corpus.

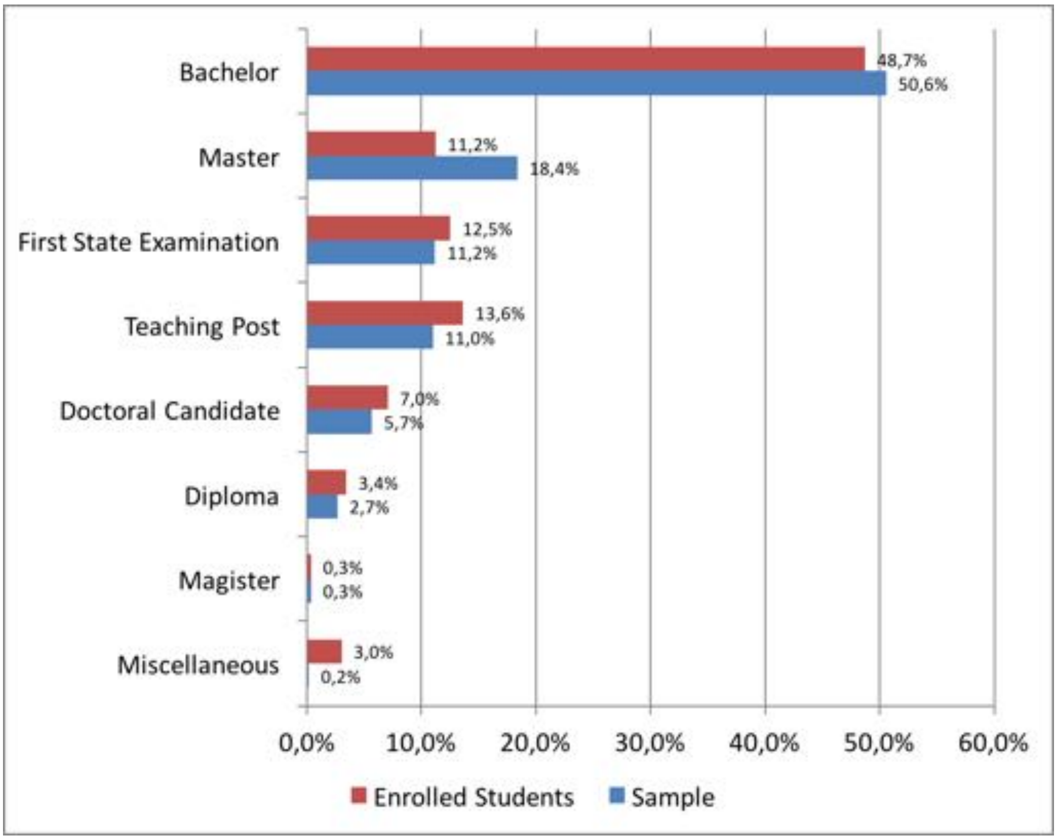


Figure 3.6.: Comparison of achieved sample size and target population divided into participants' targeted degree.

decided for a laborious but reliable manual analysis. Thereby, all answers to open questions were printed on index cards and manually clustered in categories. Later, we put focus primarily on the following eight open questions. The final analysis resulted in a total of 5456 index cards (682 participants × 8 open questions).

Questions related to pen & paper use

- In what kind of circumstances and why do you work with pen & paper?
- Which problems do you encounter and what bothers you?

Questions related to personal computer use

- In what kind of circumstances and why do you work with personal computer?
- Which programs/applications are you using on your personal computer for academic work?
- Which problems do you encounter and what bothers you?

Questions related to tablet PC & smartphone use

- In what kind of circumstances and why do you work with tablet PCs / Smartphones?
- Which programs/applications are you using on your tablet PC / Smartphone for academic work?
- Which problems do you encounter and what bothers you?

Figure 3.7 (page 68) shows the final outcome of the clustering process (clusters labels are in German). Findings will be discussed together with results from interviews, focus groups, and observations in *Chapter 3.2.4 – Empirical Findings*.

3.2.2 Interviews

While questionnaires are an excellent research tool to reach a large percentage of the target population (Butz and Krüger, 2014, p. 115), it does not allow for follow-up inquiries for disambiguation (Rogers et al., 2011, p. 238) in case questions or given answers are unclear. To compensate for this, we additionally interviewed 20 individual students and 18 dyads[11] of students. Inspired by the Individual Interaction design domain (Blended Interaction), the in-depth one-to-one interviews allowed us to specifically investigate into academic work activities and issues related to individual work. The group interviews emphasized on cooperative work. In both, interviews, we only interviewed students of a higher semester ($\geq$ 3. semester) because they usually have gained enough experience around academic working practices.

Both interview types followed a semi-structured approach. The semi-structured approach was necessary because interviews were conducted with two research teams in parallel. Over and above, it employs the 2. type of data triangulation. Each team consisted of an interviewer[12] and a note taker. An interview took approximately 30 minutes.

3.2.3 Focus Groups

To further highlight on cooperative work in academic activities, we conducted three focus groups with four students per group (see Figure 3.8). Focus groups are intended to provoke discussion among participants rather than on a bi-directional communication between interviewee and interviewer. Their nature is particularly

---

[11] "Something that consists of two elements or parts" (Oxford Dictionary). In research studies, a dyad often refers to a group of two participants.

[12] I was the interviewer in one of the two research teams

Figure 3.7.: Analysis of open questions: "*In which situations are you working with 1) the computer, 2) tablet & smartphone, and 3) pen & paper and why are you working with it?*" *Every response was manually assigned to a sub-category. Multiple sub-categories were clustered in general categories like research, accessibility, management, annotation, etc. (Note: the image blur results from stitching together multiple source images)*

Figure 3.8.: Focus group with four participants.

suitable to elaborate on aspects of collaboration and, therefore in HCI, focus groups often reveal, otherwise missed, insights socio-technical issues (Rogers et al., 2011, p. 232).

Each focus group had a facilitator[13] who sparked discussion in case the group got stuck. I was present in all focus groups to take notes and ask questions when appropriate. The facilitator was provided with a guideline and a set of potential and pre-defined questions (see Appendix A for interview guidelines, page 219). In addition to notes, each session was video recorded and lasted about 90 minutes.

3.2.4 Empirical Findings

This section discusses findings and issues emerged from conducted qualitative research. The analysis process of the questionnaire, interviews, and focus groups revealed issues with current academic work activities and surrounding practices. They will be revamped into hidden potentials with the help of theories presented in *Chapter 2.2 – Theoretical Foundation*.

The first finding is around searching & browsing for literature. It argues for reality-based and spatial interaction techniques to navigate in virtual spaces. It relates to

---

[13] Michael Schubert, researcher from the Knowledge Media Research Center in Tübingen, was the facilitator

Spatial Navigation (RO1) and will be explained and discussed in the section *Shelf Browsing & Serendipity*.

The second finding is around reading and the use of space to spatially reflect current thoughts and arguments. It directly relates to Cross-Device Interaction (RO2) and will be motivated and discussed in section *Fluid Configuration of Work Artifacts*.

Shelf Browsing & Serendipity

The Library of the University of Konstanz is one of the best academic libraries in Germany. It is frequently ranked as number one library in the nationwide *Bibliotheksindex BIX*[14]. Among many other reasons for this top ranking, it provides shelf browsing to their users. Shelf browsing is a truly valued experience as found in our study.

Most digital libraries such as search engines or OPAC allow users to find media through convergent search strategies only. For example, through a keyword search or facetted search. In contrast to digital libraries, bookshelves allow for divergent search strategies (Palfrey, 2015). No doubt, bookshelves are still popular for the literature search. This popularity is revealed in the analysis of our questionnaire in the multiple choice question "*What tools do you use for literature search?*". Interestingly, 56% of library users still search and browse for literature at shelves directly and bypass search terminals and OPAC (see Figure 3.9, multiple choice).

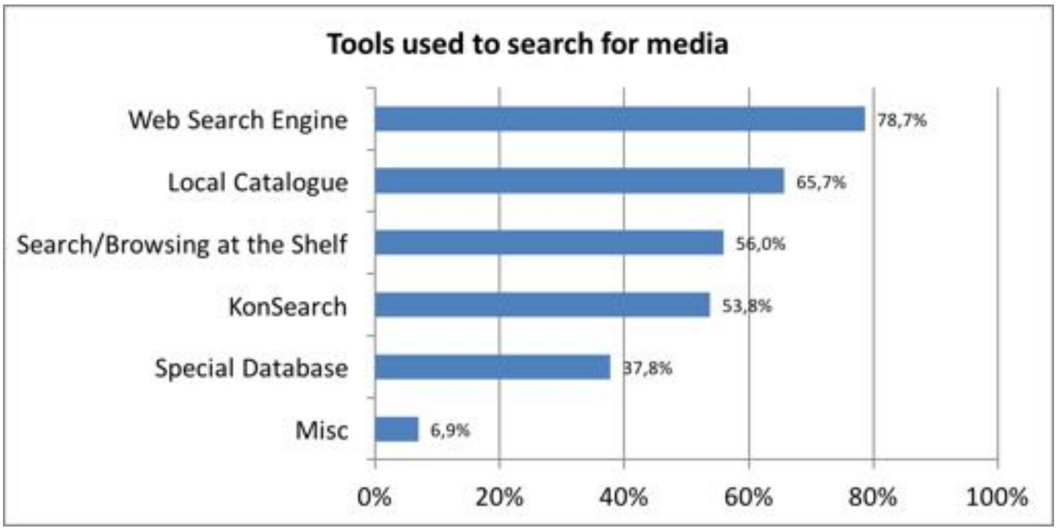


Figure 3.9.: Analysis of the question "*In which catalogs and search engines do you search for further literature?*" (multiple choice)

---

[14] Press release University of Konstanz – `http://www.aktuelles.uni-konstanz.de/presseinformationen/2011/80/` (last accessed: March 16th, 2016)

The "Search/Browsing at the Shelf" was the third-most answer. Even before the Resource Discovery System (RDS)[15] KonSearch[16] and specialized and subject-related databases such as Juris[17], PsycINFO[18], and ACM Digital Library[19]. Obviously, searching at the shelf is an integral part of the bibliographic search. Participants consider printed media as very "fast," "intuitive," and "completely free manipulation of all haptic media." Head and Eisenberg report similarly in their article "How College Students Seek Information in the Digital Age" (Head and Eisenberg, 2009). They conducted a large survey with 2,318 participants across six U.S. campuses and report that 55% of respondents browse the shelf when they seek for information (see Figure 3.10, page 71).

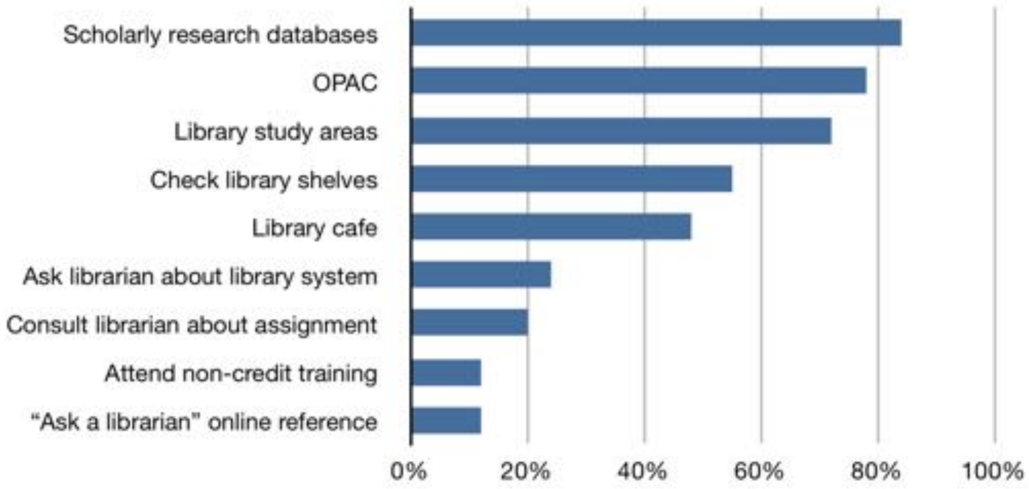


Figure 3.10.: Usage of Libraries for Course-Related Research (Head and Eisenberg, 2009). (illustration from (Head and Eisenberg, 2009))

Why is free access to shelves so important? Umberto Eco, an Italian novelist, and semiotician enunciates it to the point.

"In reality, it often happens that you go to a library because you want a book whose title you do know, but the principal function of the library, at least the function of the library in my house and that of any friend we may chance to visit, is to discover books whose existence we never suspected, only to discover that they are

---

[15] A resource discovery system unifies metadata from various digital libraries & databases to provide a single point of search experience–much like Google Search.

[16] KonSearch is the literature search engine of the University of Konstanz, which currently indexes over 300 million essays, books, papers and other media purchased by the Library of the University of Konstanz or that is freely accessible on the Internet. – `http://konstanz.summon.serialssolutions.com/` (last accessed: March 30th, 2016)

[17] Juris is a database for legal work – `http://www.juris.de/` (last accessed: March 30th, 2016)

[18] PsycINFO is a literature database for behavioral and social sciences – `http://www.apa.org/pubs/databases/psycinfo/` (last accessed: March 30th, 2016)

[19] ACM Digital Library containing full-text collection of all ACM publications. – `http://dl.acm.org/` (last accessed: March 30th, 2016)

of extreme importance to us. Of course, it's true that this discovery can be made by leafing through a catalogue, but there's nothing more revealing and exciting than exploring the shelves that perhaps contain a collection of all the books on a certain subject — something that you wouldn't be able to discover in a catalogue ordered by authors' names — and to find another book beside the book you went to find, one that you weren't looking for but that emerges as being of fundamental importance. In other words, the ideal function of the library is to be a bit like a secondhand bookseller's stall, a place where you might make a lucky find, and this function can only be fulfilled through free access to the aisles lined with shelves." (Umberto Eco, Translation from the Italian Alastair McEwen)

Many libraries follow Umberto Eco's opinion. In Konstanz, for instance, a unique systematic and manual classification by topic and free access to the media allows clients to physically search and browse shelves for interesting findings. All without knowing book titles, call numbers, or other metadata in advance. This is particularly expedient once visitors memorized the general physical location of classified topics. Based on this spatial memory, they can search and find books without the need of any computing system or device. Because of the systematic classification and arrangement, the Library of the University of Konstanz offers shelf browsing as an inherent quality. It allows clients to browse further literature by looking at the immediate vicinity of a single relevant book. Thereby, clients eventually make a serendipitous discovery. Serendipity[20] is "the occurrence and development of events by chance in a happy or beneficial way" (Oxford English Dictionary).

Lennart Björneborn[21] postulates ten dimensions, which may affect possibilities for serendipity in physical libraries (Bjorneborn, 2008). He reports on convergent (goal-directed) and divergent (explorative) information behavior and dimensions affecting serendipity. Table 3.1 (page 73) shows the ten serendipity dimensions, which were derived from observational field studies and qualitative interviews conducted in two Danish public libraries. In all 113 interviews, participants were asked about "*What did they intend to find?*", "*What did they actually find?*", and "*How did they find it?*".

Dimensions like unhampered access, imperfection, explorability, and stopability are intrinsic characteristics of freely accessible bookshelves. In contrast to OPACs and their support for convergent (goal-directed) search behavior, bookshelves fulfill needs for divergent (explorative) information behavior (Bjorneborn, 2008). If curation of books and other media like DVD and CD in shelves is systematic, like in the Library

---

[20] "1754: coined by Horace Walpole, suggested by The Three Princes of Serendip, the title of a fairy tale in which the heroes "were always making discoveries, by accidents and sagacity, of things they were not in quest of."" (Oxford English dictionary

[21] Lennart Björneborn is a professor at the Royal School of Library and Information Science at the University of Copenhagen.

Serendipity dimensions	**Explanation**
Unhampered access	Unhampered direct access to information resources
Diversity	Rich and dense variety of topics, genres, resources, activities, sections
Display	Curiosity-teasing mediation of information resources
Contrasts	Eye-catching differentiation including quiet zones and display zones
Pointers	Distinct signage, maps, markers, etc., may trigger users' interest spaces
Imperfection	Imperfect 'cracks' and 'loopholes' in library interfaces
Cross contacts	Contact surfaces across different topics, genres, resources, activities, sections
Multi-reachability	Many different access routes across library interfaces
Explorability	Library interface invites users to move, explore and browse
Stopability	Library interface invites users to stop, touch and assess found materials

Table 3.1.: Identified dimensions in library interface for stimulating serendipitous findings (Bjorneborn, 2008).

of the University of Konstanz, it covers further dimensions like diversity, display, and pointers.

Analysis of our questionnaire clearly hints at advantages of freely accessible bookshelves. This finding possibly emerged from an unfortunate incident. In November 2010, the Library of the University of Konstanz literally closed large parts of the library overnight. Asbestos fibers were found deposited on some shelves in the library and eventually got distributed across most library buildings through the air ventilation system. Many print media that were freely accessible through open shelves the day before were suddenly unavailable and kept under tight wraps. Only after careful and tedious page-by-page vacuum cleaning, books were gradually available again. However, the library space was still contaminated, and books could only be ordered from storage, a closed-stack library.

So it is unsurprising that many study participants wished a "library with bookshelves to browse"[22], "literature search directly at the shelf"[23], or "workplaces in the library, so you can quickly go to the shelf, when you need information"[24]. They even

---

[22] Translated from German "Bibliothek mit Bücherregalen zum stöbern"
[23] Translated from German "Literatursuche etc. am Regal"
[24] Translated from German "Arbeitsplätze direkt in der Bibliothek, sodass man 'mal eben' ans Regal gehen kann, wenn man noch Informationen braucht"

Figure 3.11.: An employee from a company specialized in asbestos removal retrieving ordered books before they are getting cleaned page-by-page.

acknowledged the "overview, which you can obtain when you stand in front of a shelf in the library"[25]. Strengthened by the asbestos incident, one participant clearly misses "browsing the shelves of the Library of the University of Konstanz (asbestos renovation), which is one of [his] favorite research methods"[26]. Another participant proactively suggested "a search program that allows finding media divided by topics"[27].

Fluid Configuration of Work Artifacts

With the help of the clustering of answers (see Figure 3.7, page 68), we could identify general themes of work desk related issues. Alternating between both, digital and analog media sources often disrupts higher-order learning activities (Ohlsson, 1995). Participants complained about the costly process of digitizing quotes from printed sources or handwritten notes. A participant, for instance, stated: "ultimately everything is digitally written." For this reason, many participants require

---

[25] Translated from German "Überblick, den man sich verschaffen kann, wenn man vor einem Regal in der Bibliothek steht"

[26] Translated from German "Im Moment vor allem das Stöbern in den Regalen der Konstanzer Unibibliothek (Asbestrenovierung), was zu meinen bevorzugten Recherchemethoden gehört."

[27] Translated from German "ein Suchprogramm, in dem man Medien unterteilt in inhaltliche Gruppen finden kann"

a "full digitization of all library texts." One participant summarized the advantages of working with digital texts on electronic devices as "clean working, fast search and easily shareable with others." In contrast, other participants highlighted the importance of printed documents for their working practices. In their opinion, it is easier to use physical space and compare the contents of several books spread on a desk rather than on a single digital device. Resonating with this, one participant reported "I keep my research texts open on my tablet to free up my computer screen for other work".

The following real-world case, documented through observation, exemplifies the use of space as an important tool during higher-order learning activities (Ohlsson, 1995). Figure 3.12 (page 75) (left) illustrates a typical use of a work desk in academic libraries. This picture was taken in the Library of the University of Konstanz during an interview with students. In this particular case, two law students are working individually on their assigned legal cases. Their individual workspaces are highlighted in yellow and blue in Figure 3.12 (page 75) (right). Important to notice is that both students work with printed analog media as well as with laptop computers. The computing device defines the center of each workspace. The space surrounding this center area is used to spatially reflect the process of literature review and for note taking.

Figure 3.12.: A workspace in a library occupied by two students working independently on legal cases. *(Photo credit: Christoph Gebhardt)*

In HCI, this type of work practice is summarized under the term *external cognition*. Rogers puts it as follows: "At the highest conceptual level, external cognition refers to the interaction between internal and external representations when performing cognitive tasks (e.g., learning)" (Rogers, 2005). For instance, stacked books define the chronological sequence of a search path from a legal text to legal commentaries. The lowermost book is the initial source, and the uppermost book represents the latest acquired source. Lined up books at the far end are currently unused and are information resources shared by both students. By doing so, the students keep an external representation of the search paths to free their working memory for other cognitive tasks.

This use of space is very similar to Sellen and Harper's hot, warm, and cold documents metaphor (Sellen and Harper, 2001). Hot documents are resources with which a user currently works. Examples are the laptop computers or piled books. Warm documents are documents that will attain a user's attention in the near future. For instance, documents put aside or to the corner of a desk considered relevant but for later use. In this example, the lined up books are warm documents. Cold documents are resources that may be important in distant future. Usually, these documents are stored in filing cabinets, drawers, or in bookshelves.

Recent efforts on post-WIMP user interfaces, in particular tangible of user interfaces, build on the principle of external cognition allowing for physical spatial arrangements and fluid configuration of interface artifacts. For example, Geyer et al. present IdeaVis (Geyer et al., 2012) for collaborative sketching and AffinityTable (Geyer et al., 2011a), a computer-supported hybrid surface for affinity diagramming. Examples of other researchers include tools for ideation in creative design and sketching (Vyas et al., 2009; Klemmer et al., 2001; Harboe and Huang, 2015), sensemaking and analysis of data (Isenberg et al., 2012; Morris et al., 2010; Andrews et al., 2010; Wallace et al., 2013; Haber et al., 2014), and collaborative search (Jetter et al., 2011; Klum et al., 2012; Rädle et al., 2013b).

The *use of* and *interaction in space* is an important cognitive resource during work desk activities (Kirsh, 2010). For instance, to externalize and spatially reflect current thoughts or to provide a shared space for in-group discussion. Another important aspect of supporting knowledge work activities is the need to interact with and across documents. For instance, "laying out documents in space for reading, in order to read and write across documents is crucial" (Sellen and Harper, 2001). As illustrated in *Chapter 1.2.2 – Cross-Device Interaction*, interaction across multiple mobile devices can be quite difficult. Resonating with Kirsh and Sellen and Harper, Kidd also documented a similar behavior for knowledge work where "knowledge workers use physical space, such as desks or floors, as temporary holding pattern for inputs and ideas which they cannot yet categorise or even decide how they might use it" (Kidd, 1994). For collaborative work, Vyas et al. "believe that in order to better support [. . .] work and to develop ubiquitous collaborative technologies, we need to understand how spatial aspects support cooperative practices [. . .] and what role they play in supporting creativity in the everyday work" (Vyas et al., 2009).

3.2.5 Summary of Findings

Qualitative research conducted at the Library of the University of Konstanz revealed two prevalent knowledge work practices in libraries.

First, shelf browsing to search for literature is still practiced by 56% of library users. The shelf browsing experience and its inherent quality providing serendipitous discoveries are valued and wanted by most library users but only available in open-stack libraries.

Second, a fluid configuration of work artifacts in physical space allows knowledge workers to externalize thoughts and manifest research trajectories in a tangible graspable way. This externalization frees resources in working memory, which can be used for higher-order learning such as arguing, critiquing, or explicating (Ohlsson, 1995).

However, both knowledge work practices are not entirely supported by computing technology. The following section will explore this problem space through the design of two research prototypes. The practical exploration leads to particular research questions, which are subject to study in Chapters 4-7.

3.3 Research through Design

In recent years, HCI has transitioned from "commitment to users towards a more exploratory take-it-or-leave-it approach" (Bødker, 2006). Resonating with this change, "usability has been re-operationalized, in terms of a range of user experience goals (e.g., aesthetically pleasing, motivating) in addition to the traditional set of efficiency goals" (Rogers, 2012).

As a consequence, new approaches found their way into HCI such as *design-oriented research* and *research-oriented design* (Zimmerman et al., 2007; Fallman, 2007). Figure 3.13 (page 78) illustrates the spectrum of design in research with both extremes design-oriented research and research-oriented design. The research-oriented design is targeting towards a product that is deployed and tested with clients. It aims for "'real' things, such as commercial aspects, cost, time to market, sales figures, political interest, or user preference" (Fallman, 2007). The making process is intuitive with a strong focus on aesthetics and user experience. Design-oriented research emphasizes on rational and design often grounds on theoretical pillars and thus truth, which "is not necessarily what is 'real'" (Fallman, 2007). The process of making happens through problem analysis, discussion and reflection with peers.

"Yet, there is currently a general lack of informed design" (Buxton, 2007). Still, "many HCI researchers commonly view design as providing surface structure or decoration" (Zimmerman et al., 2007). However, Zimmerman et al. argue for design and "making as a method of inquiry in order to address wicked problems" (Zimmerman et al.,

Figure 3.13.: Continuum of design in research with design-oriented research at the far left end and research-oriented design at the far right end of the spectrum. (adapted from (Fallman, 2007))

2007). They “intend the term design research to mean an intention to produce knowledge and not the work to more immediately inform the development of a commercial product” (Zimmerman et al., 2007). Similar, Buxton believes that research through design is “a good start” (Buxton, 2007) and thus an important tool to understand “how to take the larger ecological, contextual, and experiential aspects of "the wild" into account” (Buxton, 2007).

3.3.1 Exploring Problem Space through Design

Both design-oriented research and research-oriented design have their right to exist. I believe that both can go hand in hand and research in HCI can benefit from their mutual influence. My beliefs are in line with other HCI researchers (Buxton, 2007; Zimmerman et al., 2007; Fallman, 2007).

Therefore, we further explored the problem space in a research-oriented design approach. Based on empirical findings of previous sections, we implemented two research prototypes. *Blended Shelf* (research prototype 1) to tackle *bibliographic search & browsing* and *Integrative Workplace* (research prototype 2) to tackle *reading & writing across documents*. Both prototypes were deployed in “the wild” (Buxton, 2007) to evaluate them under realistic constraints (Fallman, 2007) and further refine the problem space.

Then, we traveled along the continuum and transitioned from the research-oriented design towards the design-oriented research. In the design-oriented research, we took problems identified during user testing and transformed them into research questions and hypothesis. This approach allowed us to operationalize particular HCI problems and to do further experimental research. As part of this research process, hypothesis generation was massively influenced by the UbiComp vision and theoretical foundations from *Chapter 2.2 – Theoretical Foundation*.

3.3.2 Research Prototype 1: Blended Shelf

As an initial stab to satisfy participants needs and wishes and bring back the qualities of shelf browsing, we designed and implemented Blended Shelf (Kleiner et al., 2013). Blended Shelf is a user interface, which provides a shelf-like browsing experience combined with OPAC search & facetted navigation. It enables the exploration of library collections through multi-touch interaction and a familiar 3D shelf visualization (see Figure 3.14, page 80). We deliberately chose a 3D visualization of library collections to reflect real-world attributes like dimensions and color of books, which was inspired by libViewer (Rauber and Bina, 1999). Beyond Rauber and Bina's work, books are initially arranged by library's classification to emulate shelf-like experiences and spark serendipitous discoveries.

But instead of implementing a digital shelf only, we thoroughly considered Reality-Based Interaction design tradeoffs and respectively followed the four domains of design of the Blended Interaction framework. Blended Shelf was also inspired by Bohemian Bookshelf and its multiple coordinated visualizations through linking & brushing[28]. It provides various visual access points to explore book collections and thereby sparks serendipitous discoveries (Thudt et al., 2012). Bohemian Bookshelf was evaluated at the University of Calgary library. Results show the great potential of interactive exploration of book collections and even some users "wanted to check them (books – Ed. notes) out from the library" (Thudt et al., 2012, p. 1468). However, book checkouts were impossible due to the lack call numbers and their system indexed a collection of 250 books only (Thudt et al., 2012).

Blended Shelf indexes over two million books. It purposefully supports an entire workflow eventually beginning with a search, then browsing for related literature until getting a book's call number, and finally retrieving the book from the library. As such Blended Shelf supports two additional serendipity dimensions, cross contacts, and multi-reachability (see Table 3.1, page 73) and is therefore truly superior to a traditional physical bookshelf. Users can re-arrange books on-the-fly or search within whole library collections. By combining qualities of shelf browsing with functions of OPAC, it thereby unifies both convergent (goal-directed) and divergent (explorative)

---

[28] Linking & Brushing: "The idea of linking and brushing is to combine different visualization methods to overcome the shortcomings of single techniques. Interactive changes made in one visualization are automatically reflected in the other visualizations. Note that connecting multiple visualizations through interactive linking and brushing provides more information than considering the component visualizations independently." (Keim, 2002)

information behavior in a single system (Bjorneborn, 2008). This reality-based approach allows users to apply their knowledge and habits from library browsing in the digital domain and profit from the advantages of the indexed and classified information space library (Kleiner et al., 2013).

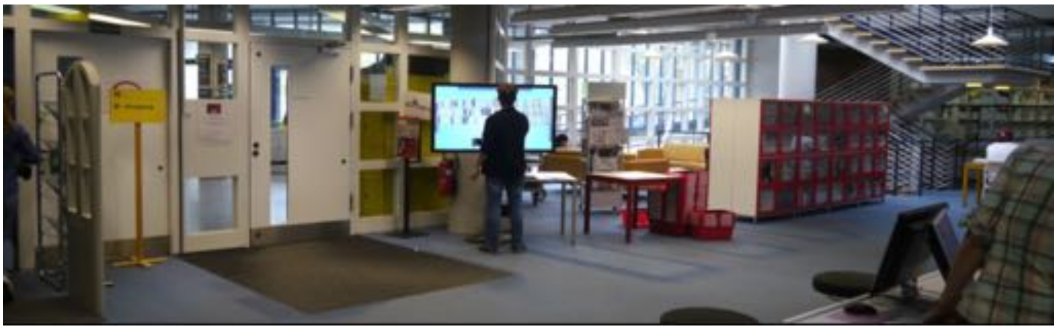
(a) 3D shelf visualization

(b) Loan status of books.

Figure 3.14.: Blended Shelf is a user interface for reality-based exploration of library collections. (a) Reality-based shelf visualization view with books, a search field (bottom), and an expanded sorting option (left). (b) Loaned books are presented semi-transparently in Blended Shelf and thus can be found during search or browsing. Opaque books are available.

In his master thesis, Kleiner evaluated Blended Shelf in a comprehensive field study with actual users of the Library of the University of Konstanz (see Figure 3.15, page 81) (Scan QR Code #2 or enter "2" in the *MediaBrowser* application). Apart from quantitive research methods such as data logging, he conducted questionnaires and interviews (N=16) and observed users during interaction with Blended Shelf. In addition, he also conducted expert interviews with librarians (N=6) to gain insights from professionals working in digital libraries & media archiving, system administration & maintenance, and classification of literature. The study revealed valuable information about users' experience with the system and whether they conceive it as usable and useful. An elaborate report of study results can be found in Kleiner's master thesis (Kleiner, 2013).

Participants conceived Blended Shelf as a useful system to explore digital library collections. They valued the design, aesthetic, and graphical representation of the



Video

Code #2

Figure 3.15.: Blended Shelf study setting in the Library of the University of Konstanz. The display was put at the entrance/exit of the N-building.

content. Moreover, they find it a simple and clear user interface to browse library collections (Kleiner, 2013, p. 124). However, the study also revealed weaknesses. Thus, most participants consider Blended Shelf as a complement to the physical exploration of bookshelves rather than a replacement (Kleiner, 2013, p. 116). Three potential reasons for this are: participants raised (i) *privacy* concerns when using the system and they often had problems (ii) *understanding interaction* (e.g, implemented touch gestures). Concerns raised by library experts are (iii) *commercial aspects* and costs to maintain such new services.

(i) Privacy: Participants consider "the display is too big"[29] (Kleiner, 2013, p. 116) and the setting as being too public. In turn, many of them felt monitored by others (e.g., by-passers). They felt, especially, uncomfortable when searching for socially inept or sensitive literature. For instance, students from social sciences searching for literature on sexual studies or literature about historically delicate incidents (e.g., terrorism). Because of this, some participants wished Blended Shelf to be more private and ideally available on tablet devices.

(ii) Understanding Interaction: Participants did not understand certain touch gestures. This was revealed during participant observation and consecutive interviews. Participants had in particular problems identifying the two fingers horizontal swipe gesture to change the perspective of the rendered bookshelf. This suggests that implemented touch gestures use a "complex jargon" (Weiser, 1991), which is "unnatural" to users and farther away from "a user's already existing concepts and basic-level experiences" (Jetter et al., 2014). In consequence, users need to learn these gestures from scratch. Of course, some touch gestures like swiping and pinching are known concepts implemented in many nowadays mobile applications. Thus, they constitute basic-level experiences of a digital native. However, interaction in 3D virtual spaces adds a new complexity, and a horizontal two-finger swipe seems yet to be unnatural.

---

[29] Translated from German "Das Display ist zu groß. Man benötigt viel Abstand. Es ist erschlagend." (Kleiner, 2013, p. 116)

Also, the system was developed for a single user only and led to unintended zooming operations when two users interacted simultaneously (Kleiner, 2013, p. 118).

(iii) Commercial Aspects: In addition to user experience issues, Blended Shelf also has disadvantages from an economical point of view. A single multi-touch display engineered to work 24/7 is expensive to buy and expensive to maintain. Libraries eventually need to hire additional personnel, IT experts, to service these devices.

Potentials of Physical Shelf-Browsing

For these reasons, it is still a challenging task to achieve a shelf-browsing experience that is at least equal or superior to an exploration of physical bookshelves and is cost neutral for libraries.

One opportunity — yet under-explored — is exploiting whole body interactions to navigate in virtual information spaces. Whole body interactions feature two potentials: *physical navigation* and *spatial memory*. Figure 3.16 (page 82) shows a potential next iteration of the Blended Shelf (Scan QR Code #3 or enter "3" in the *MediaBrowser* application). It illustrates ShelfHole, a concept that uses a situated and spatially-aware display (Fitzmaurice, 1993), also often referred to as peephole interaction (Yee, 2003; Mehra et al., 2006).

Video

Code #3

Figure 3.16.: The ShelfHole concept. Spatially-ware displays (peephole interaction) for reality-based exploration of library collections.

Peephole interaction or dynamic peephole navigation (Mehra et al., 2006) is an increasingly popular technique for navigating large information spaces using small, spatially aware displays (Fitzmaurice, 1993). Typically, the display of a handheld computer (Fitzmaurice, 1993; Morrison et al., 2009; Tsang et al., 2002; Yee, 2003), mobile phone (Morrison et al., 2009; Pahud et al., 2013; Rohs and Essl, 2006; Rohs et al., 2007), tangible display (Spindler et al., 2009), or handheld projector (Cao and Balakrishnan, 2006; Kaufmann and Ahlström, 2012; Kaufmann and Ahlström, 2013; Löchtefeld et al., 2011; Rukzio et al., 2012) acts as a window or peephole to a much larger information space, such as a map (Kaufmann and Ahlström, 2013; Morrison et al., 2009; Rohs et al., 2007; Yee, 2003).

Figure 3.17.: Example of peephole interaction. Movements of a mobile display in physical space pans and zooms virtual content displayed on the mobile device's screen.

Users can control the mobile display's content by physically moving it up, down, and sideways. By this, they can pan their view to move invisible off-screen content into the display and access the entire information space as if it was situated in physical space (see Figure 3.17, page 83). This physical way of navigation provides users with more proprioceptive cues which are assumed to improve their orientation and understanding of the information space (Fitzmaurice, 1993) and their spatial memory (Kaufmann and Ahlström, 2013). Ideally, peephole users can navigate quickly (short navigation time) and directly (short traveled path length) from their current location to any destination in the information space without an extensive task load, even if the location is off-screen or yet unknown.

For example, peephole navigation was used in 2002 and 2003 for navigation and pen interaction in 3D (Tsang et al., 2002) and 2D (Yee, 2003). In 2004, Rapp et al. transferred this concept to handheld projectors for navigating the content of a general purpose UI (e.g., calendars, emails) (Rapp et al., 2004). From then on, many more peephole designs and systems were created, including augmented reality maps (Henze and Boll, 2010; Morrison et al., 2009; Rohs et al., 2007), peepholes using handheld projectors or projector phones (Cao and Balakrishnan, 2006; Kaufmann and Ahlström, 2012; Kaufmann and Ahlström, 2013; Löchtefeld et al., 2011; Rukzio et al., 2012), peephole navigation with smartphones, tablets, and tangible displays (Henze and Boll, 2010; Pahud et al., 2013; Spindler et al., 2009). More recently, Spindler et al. conducted a comprehensive user study to contrast traditional pinch-drag-flick interaction on mobile devices to a spatial interaction or peephole interaction. Their result shows that users, on average, were 35% faster with the spatial interaction than with pinch-drag-flick (Spindler et al., 2014).

Peephole interaction is a promising interaction technique and recent work in HCI indicates potential advantages of physical navigation over virtual navigation in large information spaces. Physical navigation in space is an everyday task of humans. Early on in our lives we observe the world and imitate physical actions by kinesthetic learning. Gopnik and Meltzoff formulate it as follows: “we innately map the visually perceived motions of others onto our own kinesthetic sensations” (Gopnik and Meltzoff, 1997). Ultimately and as Jeanne Sholl argues, spatial memory is built by path integration. As introduced in the previous chapter *Chapter 2.3.2 – Spatial Memory*, spatial memory is an inherent function of human memory and allows us to, often subconsciously, navigate in space.

Also, a mobile display as an interactive device is considerably cheaper than large interactive screens. Ideally, users bring their own tablets, which further reduces the cost for libraries to only maintain software services.

Operationalization and Prospect of Research

We employed aforementioned ShelfHole concept to operationalize research in Chapters 4 and 5. The two chapters specifically address research question **RQ1** (see Chapter 1.3, page 12).

RQ1 Does an egocentric spatial navigation improve users’ navigation performance and increase their ability to recall information from memory?

In detail, research presented in *Chapter 4 – Spatial Navigation – Peephole Size & Navigation Behavior* explores the optimal size of display ("sweet-spot") for peephole interaction. It shows that a tablet can already be a good fit to explore large virtual information spaces.

In *Chapter 5 – Spatial Navigation – Spatial Memory*, egocentric peephole interaction is compared with traditional multi-touch interaction. In addition, it focusses on users' navigation performance regarding spatial memory and subjective workload.

Results of both contribute to and foster understanding of peephole interaction. They prove peephole interaction as a promising and low-cost alternative for searching & browsing library collections.

3.3.3 Research Prototype 2: Integrative Workplace

To accommodate for the need of fluid configuration of work artifacts, we designed and implemented *Integrative Workplace*. Integrative Workplace is a digital desk, which blends the qualities of digital and physical work artifacts. It allows the user to search and excerpt content from both digital and printed documents using the same interaction techniques (see Figure 3.18, page 86) (Scan QR Code #4 or enter "4" in the *MediaBrowser* application).

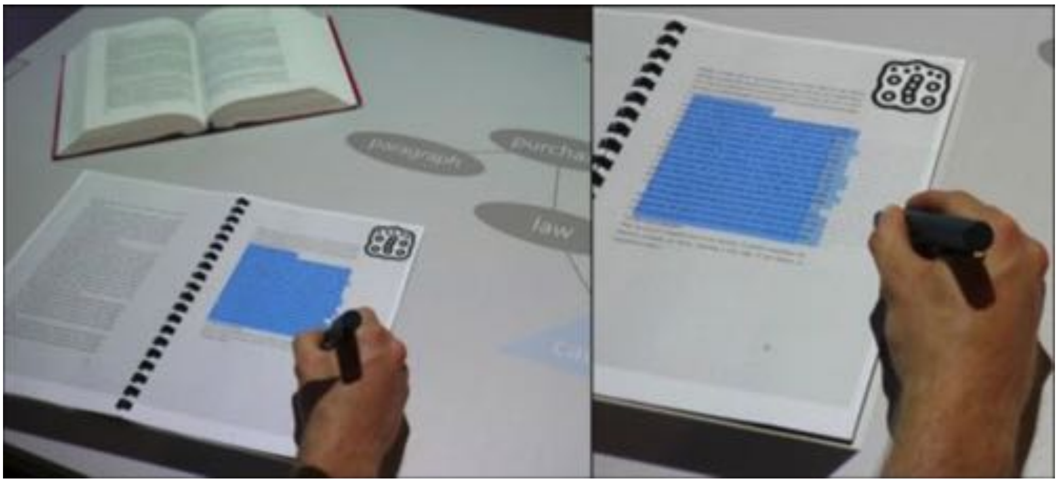
It particularly supports spatial layout of materials as argued by Sellen and Harper (Sellen and Harper, 2001), Kidd (Kidd, 1994), and Kirsh (Kirsh, 2010). For this purpose, the Integrative Workplace provides a sufficiently large surface of 100″ that be interactive by multi-touch and pen. Figure 3.19 (page 87) shows the physical setup. The setup consists of a projector to digitally augment the table and documents and a camera to track the documents put on the table. Users can spatially arrange analogue and digital contents alike, digitally excerpt from analogue media (see Figure 3.18, page 86), and cross-reference between them (similar to FACT by (Liao et al., 2010)).

Integrative Workplace was mainly inspired by Pierre Wellner's DigitalDesk. The DigitalDesk "enables people to interact with ordinary paper documents in ways only possible with electronic documents on workstation screens" (Wellner, 1993). Ever since many HCI researchers built same or alike systems that tie in with Wellner's pioneering idea "to give the physical desk electronic properties and merge the two desktops into one" (Wellner, 1993).

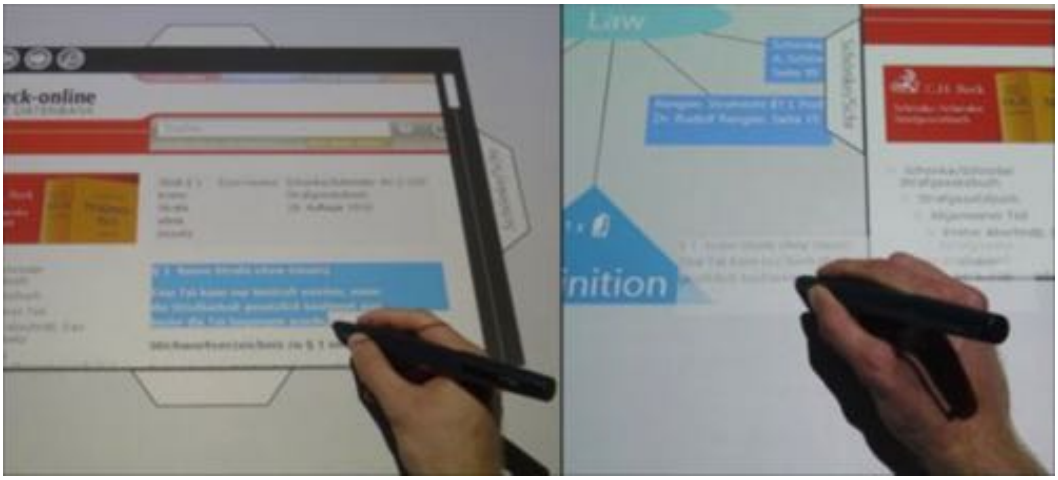
Koike et al.'s EnhancedDesk similarly integrates paper and electronic information on a desk (Koike et al., 2001). In addition to DigitalDesk, EnhancedDesk digitizes content from real objects like books (Koike et al., 2001), which is then manipulatable by touch interaction (Koike et al., 2000; Koike et al., 2001). WikiTUI concept by Wu et al. bridges the gap between printed media and hyperlinking. It "allows readers to access the digital world through fingertip interactions on books" (Wu et al., 2007; Wu et al., 2008). They evaluated their concept using a Wizard of Oz technique. Liao

Video

Code #4



(a) Excerption from a book.



(b) Excerption from a web page.

Figure 3.18.: The realization of the excerption of text. (adapted from (Gebhardt, 2013, p. 53))

et al. presents FACT, “an interactive paper system for fine-grained interaction with documents across the boundary between paper and computers” (Liao et al., 2010). It allows for fine-grain selection of content on printed documents (e.g., text or image selection and similar to (Arai et al., 1997)) and provides digital functions like copy & paste to a nearby laptop or web search using selected content as search input (Liao et al., 2010). Similarly, Dachselt and Sarmad envision *Projective Augmented Books* (PAB), a digital device in a form factor of a reading lamp enabling users to perform a variety of digital functions. PAB allows for marking and making annotations to printed media when attached to them (Dachselt and Sarmad, 2011).

Integrative Workplace combines many of the concepts presented aforementioned related work. While these systems were technically evaluated (e.g., accuracy of pen tracking or detection of hand and finger touches), none of them were tested with real users. Integrative Workplace goes beyond a mere technical evaluation. We wanted to

Figure 3.19.: Technical setup of Integrative Workplace consisting of a 100″ interactive tabletop (pen & touch enabled), a camera to track documents put on the table surface and a projector to augmented the table and documents likewise. (adapted from (Gebhardt, 2013, p. 43))

test Wellner's DigitalDesk concept with real users under realistic conditions providing an ecologically valid task.

We chose the use case of law students building a solution sketch for a legal record (as observed in *Chapter 3.2.4 – Fluid Configuration of Work Artifacts*). A solution sketch describes the strategy to solve a legal case. Law students create this document based on a literature review before they start writing a legal opinion. In a formative evaluation, participants used Integrative Workplace to solve a real-world legal record by building a solution sketch Figure 3.20 (page 88). After completing the task, they had to answer the System Usability Scale (SUS) (Brooke, 1996). In an additional questionnaire, they also gave feedback concerning the support of juristic working practices and the interaction design. This questionnaire used a five-point Likert scale ranging from 1 for "I agree" to 5 for "I disagree". Last, participants were interviewed to go deep on the questionnaire answers and to get ideas for improvements. Nine law students (3 female, 6 male) were recruited as participants. A screening ensured that participants had sufficient skills to solve a legal record (e.g., it required them to have at least written two seminar papers that include solving a legal record).

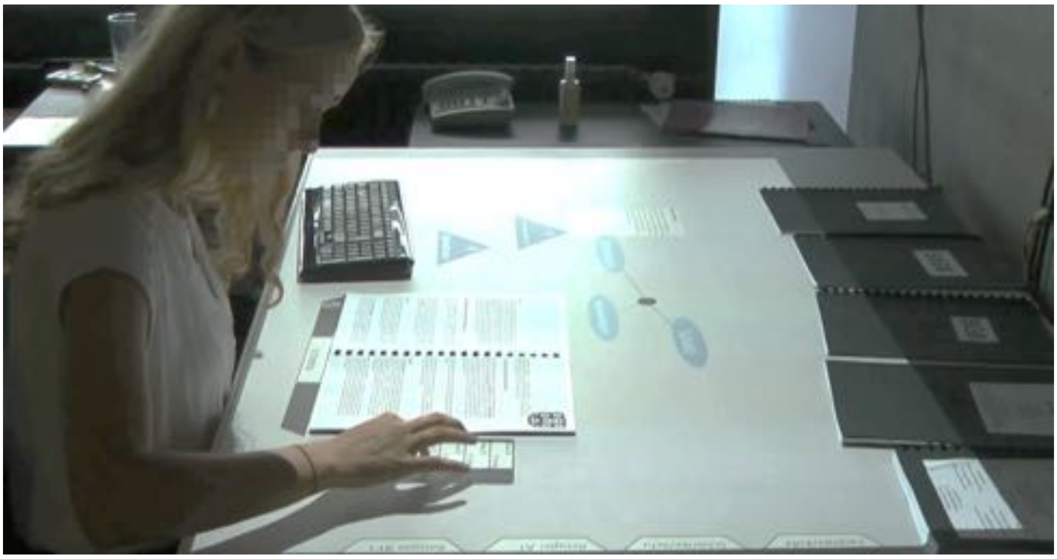


Figure 3.20.: A participant working on a legal case with help of Integrative Workplace. (adapted from (Gebhardt, 2013, p. 79))

The participants rated the system with an SUS Score of 66.94 (SD=13.22). This classifies the system's adjective rating between OK and GOOD and the system's acceptability as marginally acceptable (according to (Bangor et al., 2008)). It demonstrates that the system can be considered usable in the context of solving legal records, though the SUS Score, and therefore usability must be improved. This result is supported by the additional questionnaire where, in one question, participants rated Integrative Workplace with an average Likert value of 1.78 as useful to solve legal records. In the interview, participants most frequently mentioned “saving all excerpts at one place, no matter if excerpted from digital or analog sources” as an advantage of Integrative Workplace. Moreover, the students stated that the reference backtracking and the digital full-text search in printed documents were very useful for juristic work.

A full system description and analysis of the case study can be found in Gebhardt's master thesis (Gebhardt, 2013). Therein, Gebhardt reports on two technical limitations of Integrative Workplace, summarized in: (i) unreliable and often inaccurate *tracking quality & accuracy* and (ii) *a priori instrumentation of environment*.

(i) Tracking Quality & Accuracy: Integrative Workplace uses reacTIVision (Kaltenbrunner and Bencina, 2007) fiducial markers printed on pages to detect location and orientation of documents. While reacTIVision works dependably good for interactive and temporary music performances (Jordà et al., 2006) where lighting conditions can be controlled as part of the performance, detection of markers is often unreliable and impractical for day-to-day use. Such systems often need to be re-calibrated to map and align image projection of a beamer with reacTIVision marker tracking. Such systems are usually prone to occlusion by other objects or users' limbs. Also, system engineers have to find a perfect tradeoff for fiducial marker size. Fiducial markers are visible to the human eye and, if too big, they might overlap with document content. On the other side, fiducial markers are not detected if they are printed on page margins and therefore too small for reacTIVision image processing. Not to mention, that most documents, especially books, do not have fiducial markers printed innately on their pages (see Figure 3.18a, page 86).

(ii) Instrumentation of Environment: While a camera for fiducial tracking is to some extent movable and can be mounted above a working place and within a few minutes, a large touch-enabled surface is immobile and often permanently installed in rooms. In addition, most interactive surfaces are based on optical touch tracking and therefore often unreliable when lighting conditions change. An alternative is a capacitive sensing foil, as used for Integrative Workplace, which is available in sizes larger than 100″. However, their accuracy is often insufficient for fine-grain interaction as needed when selecting document content (e.g., selecting a single word or even particular characters).

Potentials for Fluid Device Configurations

Instead of heavily instrumenting environments, a recent approach in HCI is to join multiple mobile devices spontaneously to "community of devices" (Jetter and Rädle, 2013) or "symphony of devices" (Hamilton and Wigdor, 2014). Thereby, devices can join their individual capabilities (e.g., precise pen & touch input or photo camera) and share them for parallel or sequential use (Jokela et al., 2015a). Such a fluid configuration of multiple mobile devices allows for ad hoc adaptation of the user interface to a user's or even multiple users' needs. Ideally, users can seamlessly add or remove devices from the community in an ad-hoc fashion without prior setup of software, explicit pairing, or other additional device association (Chong et al., 2014). Instead, this happens implicitly as a by-product of natural use in space, for example, by bringing multiple devices to the same room, placing them side-by-side on a table or desk, and moving them around as needed. Ideally, users will experience these co-located cooperating devices and reconfigurable displays as one seamless and natural UI for ad-hoc co-located collaboration. Findings from a recent study show that such “lightweight relationships among an existing ecosystem of co-proximate devices can increase their cumulative value” (Chattopadhyay et al., 2016).

Operationalization and Prospect of Research

In Chapters 6 and 7, we first tackle open issues from a technological point of view to see how far we can get with low-cost and off-the-shelf hardware to identify and track work artifacts on a table. We then take our technology and evaluate its new potentials for fluid spatial arrangements of mobile devices versus non-spatial or spatially-agnostic cross-device interactions. We particularly address research questions **RQ2** and **RQ3** (see Chapter 1.3, page 12).

RQ2 How can technology support egocentric spatial and cross-device interaction, so it seamlessly integrates into people’s everyday practices?

RQ3 What are the benefits of spatially-aware cross-device interactions; and are they superior to non-spatial or spatially-agnostic cross-device interactions?

In *Chapter 6 – Cross-Device Interaction – Enabling Technology*, we present HuddleLamp, a low-cost spatial tracking of mobile devices placed on a table. This enabling technology is an alternative to Integrative Workplace, which allows for physical arrangements of electronic documents in space and interaction across device boundaries. To generalize for other spatial interaction and cross-device interactions, we explore the design space of HuddleLamp by presenting various multi-device and cross-device interaction techniques.

In a consecutive study and presented in *Chapter 7 – Cross-Device Interaction – Understanding Spatial Cues*, we elicit user-defined gestures for cross-device interactions. We implemented three cross-device interaction techniques using HuddleLamp technology, two spatially-aware techniques, and one spatially-agnostic technique. Then, we compare them in a controlled experiment to gain insights into users’ preference and their subjective workload (e.g., cognitive demand, effort, and frustration).

3.4 Summary

This chapter narrows down the application context of research to academic libraries. It digs into prevalent knowledge work practices of library users through empirical research such as questionnaires, interviews, and focus groups. This empirical research was conducted as a field study in the Library of the University of Konstanz and revealed two prevalent problems of knowledge workers.

First, library users want shelf browsing and value its inherent qualities such as serendipitous discoveries, which only open-stack libraries can provide. Closed-stack libraries and digital library collections, however, lack this opportunity since bookshelves are not directly accessible by library users or media do not have a tangible physical form to be represented in a shelf.

Second, working with multiple documents and fluidly arrange them in physical space is a common practice during reading & writing or sensemaking. However, working with both, analog and physical media, often disrupts this process as digital media cannot be arranged in physical space.

These two problems in bibliographic search & browsing and read & write across documents were identified as problems, which provide challenges and potentials for improvements likewise. We transformed them into potentials under consideration of established theoretical background and UbiComp.

Before conducting experimental research, we further explored the problem space through design. In a research-oriented design approach, we implemented two research prototypes Blended Shelf and Integrative Workplace. Outcomes of conducted "in the wild" user studies helped to further refine the problem space and clearly define research questions guiding experimental research of consecutive chapters.

Parts of the next Chapter 4 appear in the following publication:

Rädle, R. Jetter, H.-C. Müller, J. Reiterer, H. (2014a). "Bigger is not always better: display size, performance, and task load during peephole map navigation". In: *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*. New York, New York, USA: ACM Press, pp. 4127–4136. DOI: `10.1145/2556288.2557071`[30]

---

[30] The responsibilities for this joint publication were divided as follows: I formulated the research question, designed and conducted the study, analyzed the study data, and spearheaded the writing. Hans-Christian Jetter equally helped in formulating the research question and writing the paper. Jens Müller helped in conducting the study. Harald Reiterer supervised the work.

4

Spatial Navigation – Peephole Size & Navigation Behavior

> *Life is a peephole, a single tiny entry onto a vastness–how can I not dwell on this brief, cramped view of things? This peephole is all I've got!*
>
> — **Yann Martel**
> (Spanish-born Canadian author)

Contents

The previous chapter *Chapter 3 – Context & Analysis* revealed Blended Shelf as a complementary system when searching for literature. However, a user study also revealed limitations of the current implementation of Blended Shelf. The three limitations are (i) *privacy*, (ii) *understanding interaction*, and (iii) *commercial aspects* (see *Chapter 3.3.2 – Research Prototype 1: Blended Shelf*). We proposed ShelfHole as an alternative concept for reality-based exploration of digital library collections.

In this chapter, we investigate the tradeoff, or "sweet spot", between peephole size and both user navigation performance and user task load. To do this, we conducted a controlled lab experiment during which 16 participants completed a map navigation task on a large, vertical screen with physical navigation of simulated dynamic peepholes (see Figure 4.3, page 100).

We opted for map navigation in favor of browsing a bookshelf since most users are familiar with map navigation. It, therefore, avoided lengthy introduction of the task. Nevertheless, both map navigation and browsing a bookshelf have overlapping characteristics, which allow us to generalize from map navigation to searching and browsing a digital bookshelf. For example, both require users to navigate in a beforehand unknown information space and learn points of interest.

In the following, we formulate our hypotheses, describe the experimental design, report results, discuss our findings before we conclude by putting this research in the context of related work.

4.1 Formulation of Hypotheses

We entered our experiment with the following basic assumption about the nature of map navigation with peepholes: A typical map navigation activity can be separated into two phases, a learning phase, and a navigation phase.

The learning phase only takes place if the content of the map is unknown to the user or the spatial relations within the map are only partially remembered and must be reactivated from the users' memory. This is the case when users navigate an unknown map or a map they have not seen or used recently. During this phase, users first have to scan the entire map by physically moving the peephole to get an overview and to memorize positions, map features, and their spatial relations before they then can navigate efficiently. Naturally, a larger display size reveals more content and visual features. Therefore, it should facilitate learning and at the same time reduce the amount of physical panning in favor of more visual scanning. Our hypotheses for the learning phase were that a larger peephole size decreases (i) the traveled path lengths and (ii) the times for completing the navigation tasks.

In the navigation phase, a mental representation of the actual map is already present in the users' memory. This is either the case when a mental representation of a map remains in a user's memory after the learning phase is completed or when they are already familiar with the map. In the navigation phase users can, in principle, navigate toward destinations in the map efficiently, even if they are currently invisible. They do not have to scan vast areas of the map anymore to find their destination but can rely on their spatial memory (and proprioceptive cues and motor memory from physical navigation) to reach their targets faster and with a shorter traveled distance. In comparison to the learning phase, the navigation phase more resembles a pointing task without exhaustive scanning or searching and thus is less affected by peephole size. However, based on Fitts' law models of peephole target acquisition (Cao et al., 2008; Huber et al., 2011; Kaufmann and Ahlström, 2012), there still should be differences between the peephole sizes. Like in the learning phase, our hypotheses for the navigation phase were that a larger peephole size decreases (i) the traveled path lengths and (ii) the navigation times for completing the navigation tasks.

For the overall navigation task including both phases, we assumed that the cognitive load and the amount of physical panning increase with a smaller peephole size. Therefore, we hypothesized that the users' reported task load (based on the mental and physical demand items of NASA-TLX (Hart and Staveland, 1988)) increases for smaller peepholes.

Finally, we predicted that the smaller the peephole, the greater the likelihood that users built an unreliable or incorrect mental spatial representation of the map and thus, when exposed to similar maps, they might not be able to recognize the one they navigated in the experiment. Therefore, our final hypothesis is that the number of errors in a post-navigation map recognition task should increase for smaller peephole sizes.

4.2 Experiment

To better understand the role of peephole size during both phases of a map navigation task, we designed a controlled laboratory experiment with high internal validity. We isolated peephole size from other possible confounding variables, such as the existence of off-screen visualizations, the design of navigation gestures, and ergonomic aspects or device-specific properties (weight of the device, readability from different viewing angles, resolution, or latency).

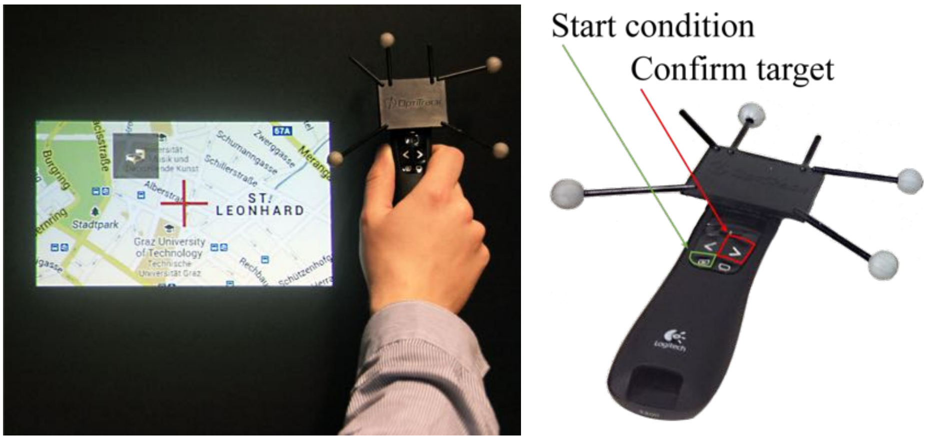


Figure 4.1.: The peephole was always displayed next to a handheld presenter device (left) with buttons and a passive IR marker for 3D tracking (right).

The study was conducted as a 4×4 within-subjects design and systematically counterbalanced using a balanced Latin Square. The independent variable, peephole size (see Figure 4.2, page 99), had four within-subjects factors: control condition (S1), projector-sized peephole (S2), tablet-sized peephole (S3), and smartphone-sized peephole (S4). We used the four different maps A, B, C, D (see Figure 4.5, page 102) to control for systemic errors and to avoid learning effects. The navigation path length, the navigation time, task load, and the post-navigation map recognition were the dependent variables.

To achieve a high degree of internal validity, we simulated the peepholes on a large display (rather than using the actual devices) so that the only variation from condition to condition was the peephole size itself. We initially discussed using different real-world devices instead of simulations, so that users would experience all device-specific properties such as different weight, resolution, or latency. However, we decided against this for following reason: Our overall goal is to understand the subtleties of peephole navigation as suggested by Pahud et al. (Pahud et al., 2013) and the different factors that contribute to navigation performance. As a first step, in this study, we wanted to focus only on the effect of peephole size which arguably is the most important property and ideally arrive at generalizable results as explained initially. Comparing devices would have led to recommending a certain device instead of a "sweet spot" peephole size without being able to isolate the role of peephole size from other device-specific properties (e.g. weight, resolution). We still would not truly understand the role of peephole size because other device properties such as weight and resolution would have come into play. Also, the recommendation would have been short-lived since such properties change with each new device generation.

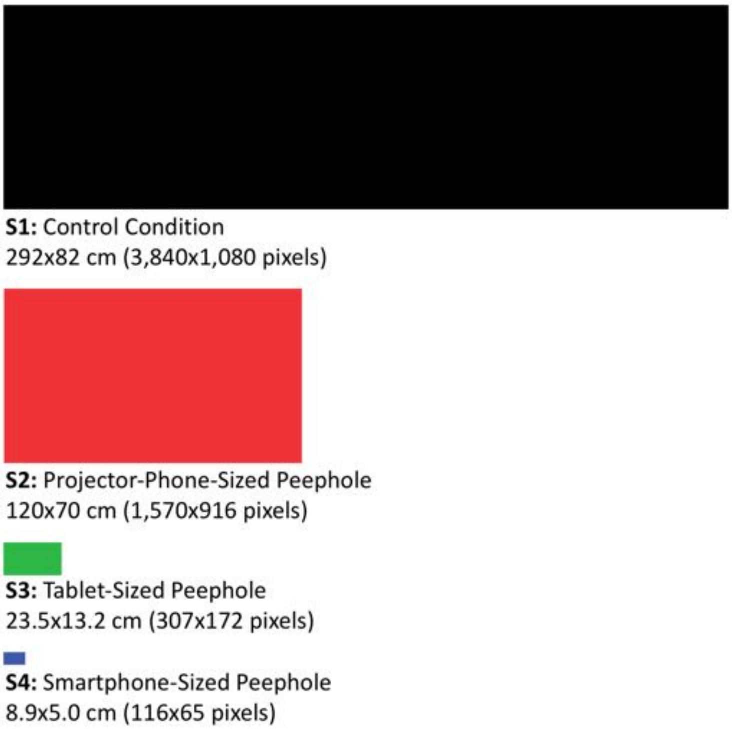


Figure 4.2.: Overview of peephole sizes S1-S4 used as independent variable in the experiment.

4.2.1 Participants

16 participants (8 female, 8 male) were recruited to take part in the experiment. The mean age was 26.6 years (SD=6.2, min=19 years, max=37 years) with a skewness of 0.64 (SE=0.56) and kurtosis of -1.15 (SE=1.10). One participant was left-handed. To get a realistic sample of participants, we excluded participants from the computer science department or with a background in computer science. 12 of the participants were students, 1 was a lecturer in linguistics, 2 were administrative staff, and 1 was a construction worker.

4.2.2 Apparatus

We used a large, vertical high-resolution screen (size 292×82 cm, 120″ diagonal, 3,840×1,080 pixels resolution) to simulate peephole sizes of typical device displays at a constant resolution of 13.1 pixel/cm (or 33.5 ppi). This resolution was lower than that of actual mobile devices, but the display quality was more than sufficient for our purposes and clearly revealed important features necessary for map navigation (see Figure 4.4, page 100). The maps used in the experiment covered the entire screen, but users were only shown a rectangular section the size of the simulated peephole while the surrounding screen was black (see Figure 4.3, page 100) (Scan QR Code #5 or enter “5” in the *MediaBrowser* application).

Video

Code #5

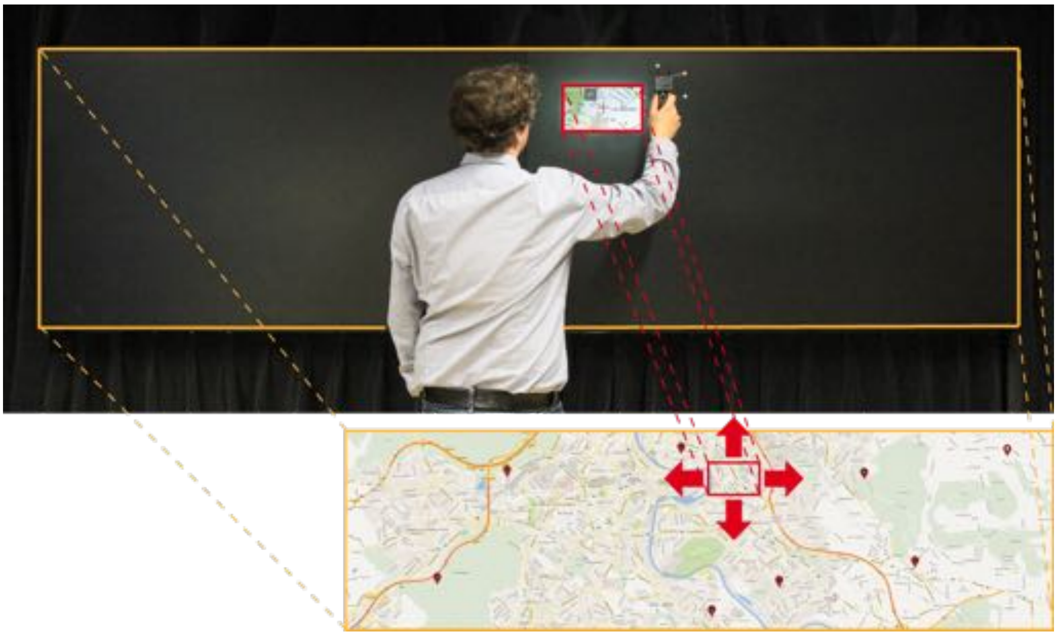


Figure 4.3.: Experimental setup simulating dynamic peephole interaction on a large vertical screen. This simulation allows to study the effect of peephole size on map navigation with high internal validity. It avoids confounding factors like weight and resolution of particular devices.

Participants used a wireless Logitech Professional Presenter R800 device (total weight 102 grams, Figure 4.1 (page 98)) to move the peephole on the screen. The presenter was equipped with passive markers and continuously tracked in space using an OptiTrack 3D motion capturing system (18 cameras) with a tracking mean error of less than .5 mm and a tracking rate of 100 Hz. Participants held the presenter in their preferred hand. A Kalman filter was used to reduce jittering caused by hand tremor and the noise or inaccuracies of the OptiTrack motion capturing system.

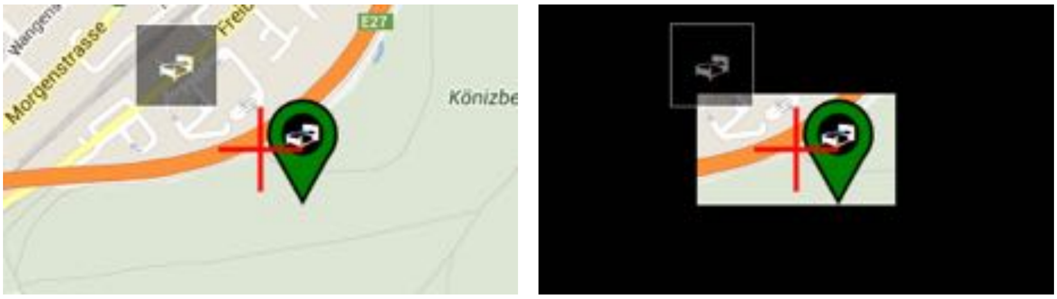


Figure 4.4.: The tablet-sized peephole S3 with 307x172 pixels (left) and the smartphone-sized peephole S4 with 116x65 pixels (right). Note that the visual features and symbols stay recognizable and have sufficient detail.

During the experiment, the peephole traveled left or right of the presenter (depending on handedness) to simulate physical navigation with a handheld dynamic peephole. By movement of their hands and lateral movement of their bodies, participants could move the rectangular peephole in the XY-plane of the display to view any location on the map, similar to the augmented maps in (Henze and Boll, 2010; Rohs et al., 2007). To minimize occlusion by hands, the anchor point was adjusted to right- and left-handed users. The ratio between physical movement of the hand in control space and the peephole's XY-movement in display and map space was always 1:1. To constrain the distance between hand and screen to realistic holding and viewing of mobile devices, the peephole only appeared on the screen when the hand was within a range of 15cm to the display. Except this, participants were free to choose their preferred head, body, and arm position during navigation and thereby set the optimal viewing distance for the display as it is the case when using an actual handheld device. However, they could not use rotation around the X-, Y-, or Z-axis as it is possible in AR see-through scenarios (Morrison et al., 2009; Rohs and Essl, 2006; Tsang et al., 2002). There was a red crosshair in the center of the peephole for selecting targets. We showed the target that the user searched for above and to the left of the crosshair.

4.2.3 Task Design

There was one condition for each peephole size. Due to the balanced Latin Square design, display sizes and maps were counterbalanced. Each map had 4 target pins that acted as navigation goals and 4 distractor pins. Maps were taken from Google Maps but were all unknown to the participants. All maps had similar visual features and complexity, such as a city with roads and a river (see Figure 4.5, page 102).

During the task, participants were asked to navigate with the peephole to a target pin in the map that shows an individual symbol, e.g., a bed (see Figure 4.4, page 100). They were asked to navigate as quickly and as precisely as possible and to select the target with the peephole's crosshair by pressing the confirm button on the presenter (see Figure 4.1, page 98). The next target was presented to the participants if the correct target was selected. Otherwise, the trial continued. The navigation path and time traveled between showing the new symbol and its selection with the crosshair were recorded. The recording of a trial started immediately after completion of the previous trial and at the last position of the peephole. All targets were systematically placed on each map to ensure comparable target distances between the maps.

In each condition, participants had to navigate to 4 targets in the same order for 8 times (blocks). This added up to 16 participants × 4 conditions × 8 blocks × 4 targets = 2048 trials with 128 trials per participant.

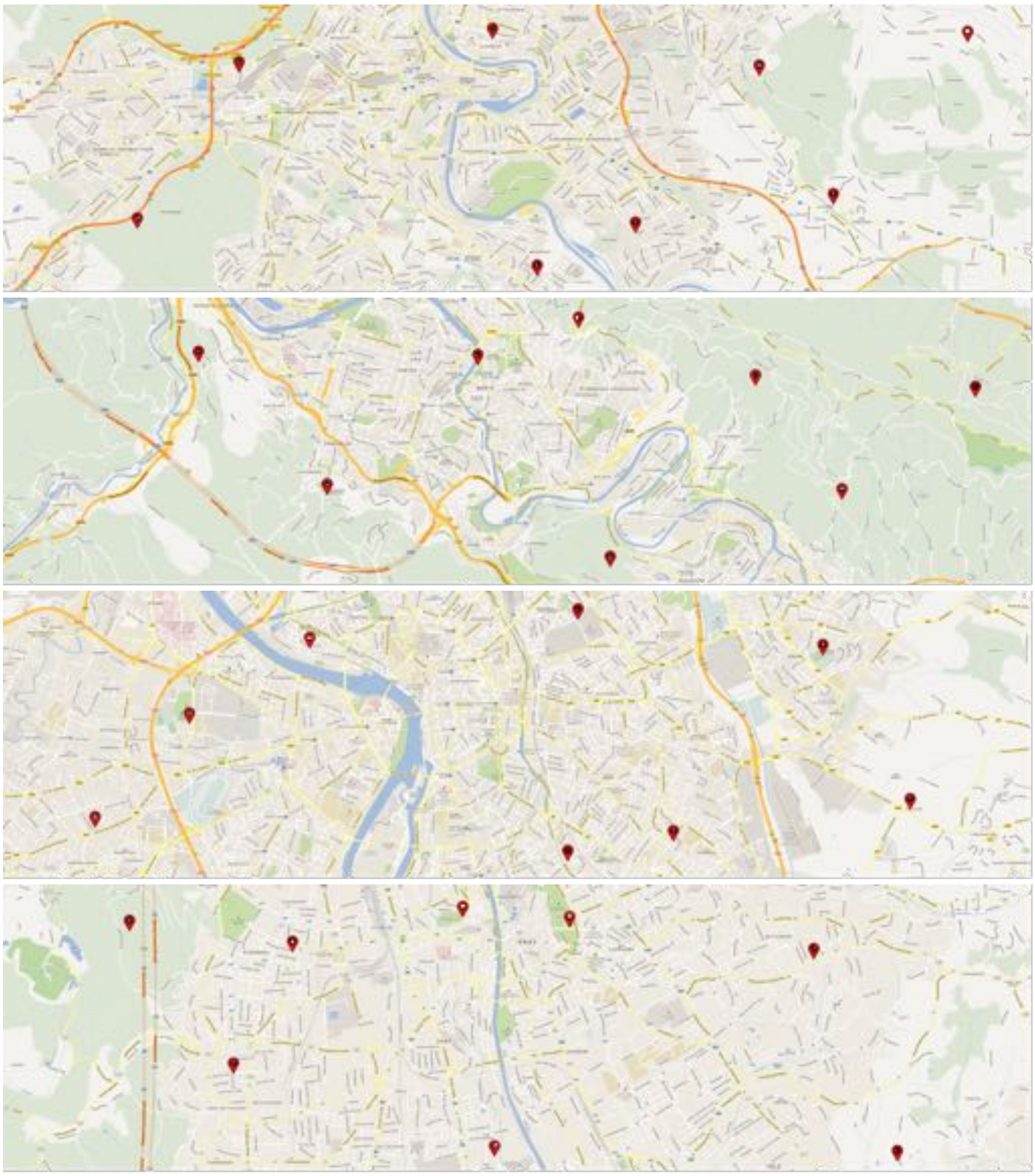


Figure 4.5.: The four maps A-D (top to bottom) used for the navigation tasks.

After each condition participants reported their subjective workload ratings using the NASA-TLX questionnaire (Hart and Staveland, 1988). After this, they chose the map they navigated from a selection of three maps. Two maps served as distractors. The purpose of this task was to test if the participants could recognize the map they had just used based on the mental representation of the map that they created during the navigation tasks.

4.2.4 Procedure

Each participant was first asked to fill out a demographic questionnaire and was asked about their dominant hand. After this, participants were introduced to holding the Logitech Presenter device with their dominant hand, its two buttons, and how to move it with their hand. To avoid learning effects of handling the peephole during the actual data collection, they then could familiarize themselves with the task, the technique for moving the peephole, and the different peephole sizes during a training phase that lasted as long as they wanted.

After this preparation phase, the actual data collection started with the four conditions. After each condition, they reported their task load using NASA-TLX and choose their map in the post-navigation map recognition task. The entire experiment lasted approximately 30 minutes per participant, and each participant was rewarded with €8 for their time.

4.3 Results

For each peephole size, Figure 4.6 (page 103) (left) shows the mean path length that the participants travelled during each block. Path lengths were normalized by dividing them by the shortest possible path length so that 1.0 is the minimum. Figure 4.6 (page 103) (right) shows the mean navigation time per block in milliseconds. Additional plots are provided for blocks 4-8 where the data points are too close together on the Y-axis to discriminate them.

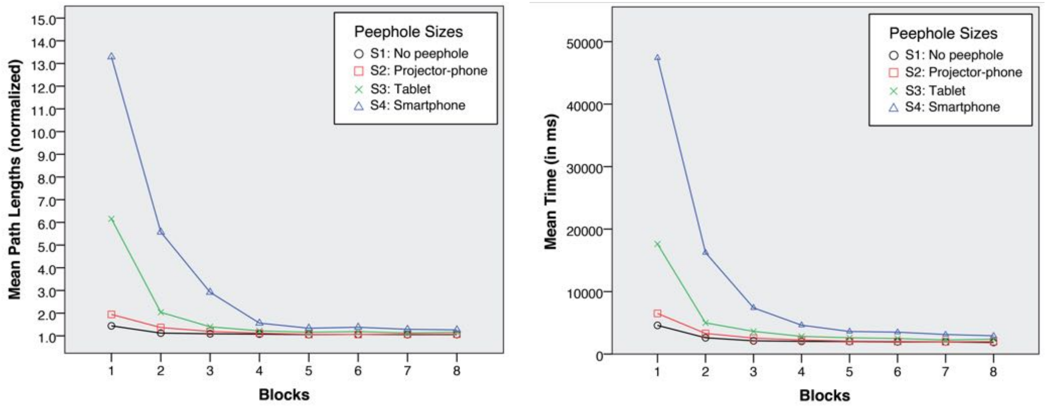


Figure 4.6.: The mean path length (left) and time (right) participants travelled during each block.

Path length and movement time analyses were done using repeated measures ANOVAs (Greenhouse-Geisser corrections are marked as GGc) with post-hoc pairwise comparisons. Table 1 shows their *p*-values for the mean of each block, the mean of blocks 1-4 (B1–B4), the mean of blocks 5-8 (B5–B8), and the mean of blocks 1-8 (B1–B8). All post-hoc tests were Bonferroni corrected.

		B1	B2	B3	B4	B5	B6	B7	B8	B1–B4	B5–B8	B1–B8
S1 vs. S2	Length	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	Time	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
S1 vs. S3	Length	**.023***	1.000	1.000	.862	**.023***	.330	**.043***	.133	**.017***	**.000***	**.008***
	Time	.061	1.000	1.000	.279	**.001***	.178	.054	**.004***	**.014***	**.000***	**.036***
S1 vs. S4	Length	**.000***	**.000***	**.008***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***
	Time	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***
S2 vs. S3	Length	.057	1.000	1.000	1.000	**.049***	.310	.291	.340	.064	.000*	**.038***
	Time	.168	1.000	1.000	.904	**.004***	.379	.095	.057	.078	**.000***	.146
S2 vs. S4	Length	**.000***	**.001***	**.015***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***
	Time	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***	**.000***
S3 vs. S4	Length	**.000***	**.008***	**.045***	**.004***	**.000***	**.020***	**.000***	.119	**.000***	**.000***	**.000***
	Time	**.000***	**.000***	**.010***	**.000***	**.000***	**.000***	**.000***	**.002***	**.000***	**.000***	**.000***

Table 4.1.: The p-values (* means significant) for pairwise comparisons of lengths and times.

Results NASA-TLX Task Load Index

The mean results of the NASA-TLX questionnaire about task load are S1: 13.07, S2: 18.49, S3: 25.51, and S4: 47.19 (scale from 0 to 100). An ANOVA with repeated measures revealed a statistically significant main effect of the subjective workload ratings on the peephole sizes, GGc: $F_{1.82,27.23} = 37.642$, $p < .001$, $partial\ \eta^2 = .715$. Pairwise comparisons (Bonferroni corrected) of subjective workload of peephole sizes revealed statistically significant differences between S1 vs. S3 ($p < .001$), S1 vs. S4 ($p < .001$), S2 vs. S4 ($p < .001$), and S3 vs. S4 ($p < .001$). All other comparisons were not significant. The individual subscales of NASA-TLX such as mental demand or physical demand are shown in Figure 4.7 (page 105).

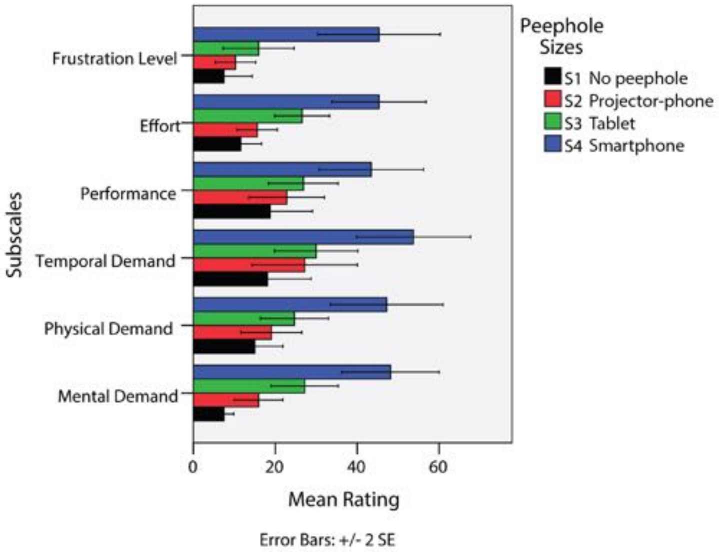


Figure 4.7.: Subscales of the NASA-TLX questionnaire.

Results Map Recognition

The results of the map recognition task revealed the following error rates for each peephole size: S1: 5 errors (31.25%), S2: 4 errors (25%), S3: 4 errors (25%), and S4: 4 errors (25%). Since there are only marginal differences in the error rates for the different peephole sizes, we have not used the error rates in the further data analysis.

4.4 Findings & Discussion

Resonating with our previous assumption about an initial learning phase followed by a navigation phase, the first blocks in Figure 4.8 (page 106) (e.g. blocks 1-4) show very long path lengths and navigation times with high standard deviations. During these first blocks, users still had to scan the information space to memorize locations and to build up a spatial mental representation of the map. This learning phase initially lead to a rapid fall of path lengths and times until the values stabilized and stayed roughly constant which indicates the beginning of the navigation phase. In the following, we discuss both phases in greater detail.

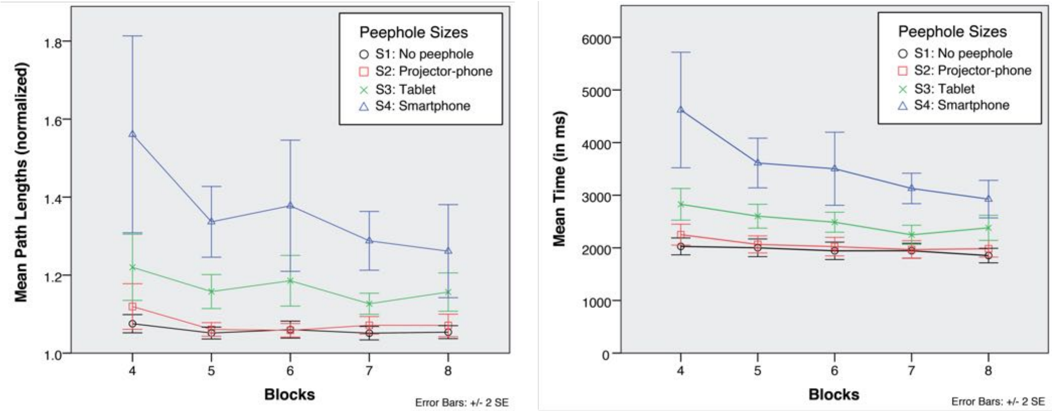


Figure 4.8.: The mean path length (left) and time (right) participants travelled during block 4 to 8.

4.4.1 Evidence of a Learning Phase

Noticeable improvements in peephole navigation occurred during blocks 1-4. This can be explained by users' improving the mental spatial representation of the map that they achieved by systematically scanning the map for targets with the peephole. The nature of this initial scanning process becomes evident when plotting peephole movements. Figure 4.9 (page 107) shows two examples of such a scanning process by participant 1 for B1 to B4 using the tablet-sized peephole S3 (top) and the smartphone-sized peephole S4 (bottom). The blue dotted lines show the movement of the peephole's anchor point on the screen. The red dots show the locations of the navigation targets. The figure illustrates characteristic scanning patterns with vertical scanning movements that are repeated horizontally or vice versa. They also visualize the potential benefit of a greater peephole size during this learning phase. Since a greater peephole reveals more visual information, it is possible to choose larger distances between the repeated movements, thus shortening the overall scanning path.

A 4×4 (peephole size × repetition) ANOVA with repeated measures revealed a statistically significant main effect of peephole size in terms of travelled path lengths, GGc: $F_{1.12,16.82} = 26.79$, $p < .001$, $partial\ \eta^2 = .641$, as well as a significant effect of repetition, $F_{1.97,29.49} = 30.69$, $p < .001$, $partial\ \eta^2 = .672$. There was also an interaction between peephole size and repetition, GGc: $F_{2.39,35.91} = 13.69$, $p < .001$,

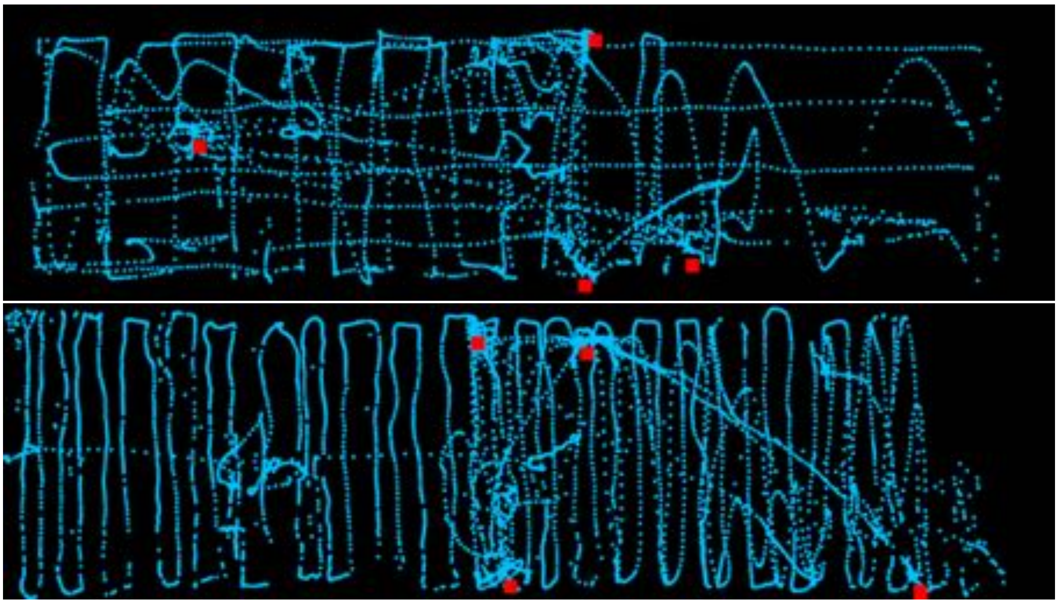


Figure 4.9.: Example of peephole movement by participant 1 during B1 to B4 with S3 (top) and S4 (bottom).

$partial\ \eta^2 = .477$. This indicates that the effects of peephole size on path lengths and navigation time is depended on the number of repetitions. We consider this as evidence of a learning process during B1 to B4. This is further supported by the fact that the same interaction effect cannot be found in the assumed navigation phase during B5 to B8 as we discuss below.

Moreover, the results show that a larger peephole facilitates this learning process and leads to better initial performance. The initial performance in B1 for path length (M_{S1}=1.44 with SE=.105, M_{S2}=1.94 with SE=.152, M_{S3}=6.16 with SE=.736, M_{S4}=13.29 with SE=2.096) shows significant differences between and the smartphone-sized peephole (S4) and all other peephole sizes (S1, S2, S3). Clearly, the smartphone-sized peephole was outperformed (see Table 4.1, page 104) (column B1). Interestingly, there are no significant differences for S1 vs. S2 and S2 vs. S3, a fact that we discuss below in a dedicated section on peephole sizes.

Looking at the entire learning phase B1–B4 reveals similar characteristics: The mean path lengths for B1–B4 (M_{S1}=1.18 with SE=.031, M_{S2}=1.41 with SE=.063, M_{S3}=2.70 with SE=.240, M_{S4}=5.84 with SE=.853) have significant differences between S1, S2, and S3 vs. the smartphone size S4, which is clearly outperformed again (see Table 4.1, page 104) (column B1-4). There are no significant differences for S1 vs. S2 and S2 vs. S3.

4.4.2 Evidence for a Navigation Phase

The different nature of the navigation phase compared to the learning phase becomes immediately visible when looking at the plots of peephole movement in Figure 4.10 (page 108) that show the same tasks as Figure 4.9 (page 107) but this time for B5 to B8. The navigation trajectories show direct navigation movements between the targets without scanning. This illustrates how participants successfully applied their mental spatial representation and proprioceptive cues of the physical peephole navigation to move efficiently between invisible but known targets without a need for scanning.

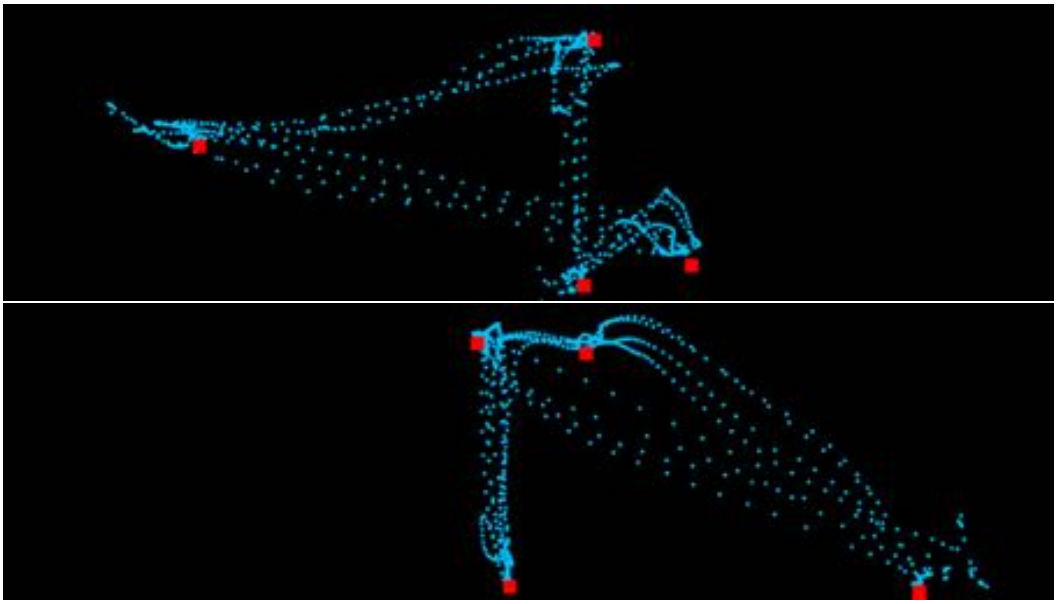


Figure 4.10.: Example of peephole movement by participant 1 during B5 to B8 with S3 (top) and S4 (bottom).

As discussed, the navigation performance regarding path lengths and navigation time substantially improved between B1 and B5. After this, as is visible in the plots of mean path lengths and mean time for B4–B8 in Figure 4.8 (page 106), the navigation performance in B5, B6, B7, and B8 stayed almost constant, however at different levels depending on the peephole size. These results indicate a gradual transition from the end of the learning phase to the beginning of the navigation phase.

A statistical indicator for the end of the learning phase and the beginning of the navigation phase is the absence of the interaction between peephole size and repetition that we witnessed for B1–B4: A 4×4 (peephole size × repetition) ANOVA with repeated measures on B5–B8 revealed a statistically significant main effect of peephole size on travelled path lengths, GGc: $F_{1.27,19.10} = 27.11, p < .001, partial\ \eta^2 = .644$ but no difference for the repetition, $F_{1.88,28.21} = 1.87, p = .175, partial\ \eta^2 = .111$. As stated above, there was no interaction between peephole size and repetition, GGc:

$F_{3.47,51.98} = 1.04$, $p = .388$, $partial\ \eta^2 = .065$. Also the small standard deviations in Figure 4.8 (page 106) allow the conclusion that participants reached almost constant performance levels for the different peephole sizes.

A post-hoc pairwise comparison with Bonferroni corrections of peephole sizes for B5–B8 (see Table 4.1, page 104) revealed statistically significant differences between peephole sizes S2 vs. S3, S2 vs. S4, and S3 vs. S4. However, it did not show a difference for the comparison of S1 vs. S2. The mean values for path lengths for B5–B8 are M_{S1}=1.05 (SE=.006) for S1, M_{S2}=1.07 (SE=.006) for S2, M_{S3}=1.16 (SE=.017) for S3, and M_{S4}=1.32 (SE=.043) for S4.

4.4.3 Peephole Sizes: Is Bigger Always Better?

Up to now, the results were mainly reflecting our initial assumptions about the existence of a learning phase, a navigation phase, and the benefits of larger peepholes that we formulated above. However, there are some unexpected observations that shed light on the question, "Is bigger always better?"

Control Condition (S1) vs. Projector-Phone-Size (S2)

Table 4.1 (page 104) shows that for all blocks in B1–B8, each individual block, the learning phase (B1–B4), and navigation phase (B5–B8), there was no significant difference between control condition S1 and the peephole S2. This is clearly a case for projector-phones since there were no significant differences in performance between S2 and a 120″ large screen without any peephole. Also, the NASA-TLX questionnaires did not report a significantly different workload with S2 compared to S1. Therefore, when comparing S1 vs. S2, bigger is not better. To expand this conclusion, peephole sizes greater than a projector-phone do not pay off in terms of navigation performance or task load when used in a map navigation scenario that is similar to our experiment.

However, in our study, S2's size of 54.7″ covers a greater field of view than might be typical in real-world uses of projector-phones. Participants stood within close range (approx. 40cm) to the screen resulting in covering approx. 127° of the users' typical field of view. In (Kaufmann and Ahlström, 2013), users stood at a distance of 200cm, so that the projection covered approx. 33.4° of the users' field of view. Interestingly, in the light of this size of S2 in our study, it is therefore even more surprising that the tablet-sized peephole S3 achieved an almost comparable performance as we discuss in the following.

Projector-Phone-Size (S2) vs. Tablet-Size (S3)

The comparison of the projector-phone condition S2 vs. the tablet condition S3 in Table 4.1 (page 104) reveals that there are no significant differences in both devices except for B5, B5-8 (navigation phase), and, as a result, also for the overall performance B1–B8. S2 outperforms S3 only during the navigation phase, but not during the learning phase. A comparison of the absolute differences in terms of path lengths and times during the navigation phase shows an 8.4% longer navigation path length and 419 ms longer navigation time per target.

While statistically significant, these differences have to be seen in relation to the aforementioned disadvantages and practicalities of mobile projections versus tablets. In our interpretation, the only moderately increased performance during navigation phase cannot outweigh the many disadvantages of mobile projection and the many advantages of using off-the-shelf tablets. Furthermore, there are no significant differences between S2 and S3 in terms of the reported subjective workload. By this, we do not imply that a tablet-sized peephole should be considered as an equivalent to a projector-phone-sized peephole in every respect. However, designers of peephole navigation systems should carefully balance the specifics of both technologies. We, therefore, suggest for use cases that are similar to our experiment that a tablet-sized peephole is more suitable than a larger one.

Smartphone-Size (S4)

Our results clearly show that a peephole with the size of a smartphone is outperformed by all other peephole sizes. This is particularly interesting with respect to tablets which are natural competitors to smartphones in peephole navigation scenarios due to their great availability, popularity, price, and mobility. The tablet-size S3 outperforms S4 in blocks B1 to B7, during the learning phase B1–B4, the navigation phase B5–B8, and the overall performance B1–B8. This is also reflected in the report of subjective workload from the NASA-TLX questionnaires, which is 82% higher for S4 than for S3 and also higher in all subscales (mental demand = 72%, physical demand = 85%, temporal demand = 76%, performance = 53%, effort = 69%, and frustration = 75%).

These findings of better navigation performance and less workload with S3 compared to S4 (13.8% in path lengths and 864ms in time) could be helpful for revisiting the study of Pahud et al. (Pahud et al., 2013). Replacing the 4.3″ device in their study with a tablet should lead to better navigation performance and reduced task load in their physical navigation condition. This could lead to different results for their comparison between virtual and physical navigation. These findings are also relevant for Kaufmann and Ahlström's study of spatial memory and map navigation performance with projector-phones vs. a smartphone (Kaufmann and Ahlström, 2013). It would be interesting to see if the reported significant differences in spatial memory still exist when replacing the smartphone with a tablet-sized peephole.

4.5 Related Work

Peephole navigation with handheld, spatially aware devices was originally conceived by Fitzmaurice in 1993 (Fitzmaurice, 1993). It was then used in 2002 and 2003 for navigation and pen interaction in 3D (Tsang et al., 2002) and 2D (Yee, 2003). In 2004, Rapp et al. transferred this concept to handheld projectors for navigating the content of a general purpose UI (e.g., calendars, emails) (Rapp et al., 2004). From then on, many more peephole designs and systems were created, including augmented reality maps (Henze and Boll, 2010; Morrison et al., 2009; Rohs et al., 2007), peepholes using handheld projectors or projector-phones (Cao and Balakrishnan, 2006; Kaufmann and Ahlström, 2012; Kaufmann and Ahlström, 2013; Löchtefeld et al., 2011; Rukzio et al., 2012), peephole navigation with smartphones, tablets, and tangible displays (Henze and Boll, 2010; Pahud et al., 2013; Rädle et al., 2013a; Spindler et al., 2009).

4.5.1 Comparative User Studies of Peephole Designs

Despite the popularity of peephole navigation, it took until 2006 for user studies to move beyond formative usability evaluations of individual systems and use controlled experiments to better understand the different design variants of peepholes more generally.

Mehra et al. (Mehra et al., 2006) simulated a handheld peephole on a 15″ screen showing a 3.3″ peephole in two conditions: 1) dynamic peephole navigation: users move the peephole with the mouse across the screen to simulate physical navigation, 2) static peephole navigation: users use the mouse to scroll/pan the information space behind the peephole that always remains in the center of the screen to simulate virtual navigation. Results showed that dynamic peepholes improved users' speed and accuracy of discriminating lengths. Mehra et al. expect a substantial increase in users' situation-awareness and better estimation of spatial relationships when using handheld peepholes.

Rohs et al. used a phone as a peephole for map navigation to compare virtual vs. physical navigation with and without visual context (Rohs et al., 2007). They found that physical navigation clearly outperforms virtual navigation with a joystick and that visual context (i.e. a map) behind the peephole does not substantially increase performance, potentially because of the costs of switching and refocusing between the two layers of information. These findings resonate with Henze and Boll who report that a simple off-screen visualization (i.e. arrow) can decrease the task completion time and that its effect is stronger than that of having visual context (Henze and Boll, 2010). Rohs and Essl compared different peephole designs such as panning, zooming, and halo (Rohs and Essl, 2006). They report that the halo off-screen visualization is significantly faster and that only in complex situations zoom and halo show comparable performance, while the combination of halo and zooming is detrimental. In our study of peephole size, we, therefore, used only panning without zoom, no off-screen visualizations, and no visual context around the peephole to avoid confounding variables and to achieve better internal validity.

In 2013, three similar studies that compared physical vs. virtual touch-based peephole navigation were published. Kaufmann and Ahlström conducted a study to find out if navigation performance and spatial memory performance during map navigation can be improved by using a projector-phone with a peephole interface (54.7″) instead of a smartphone (4″) with a touchscreen. They report that users performed navigation in the zoomable map equally well, but that spatial memory performance was 41% better for projector-phone users (Kaufmann and Ahlström, 2013). In *Chapter 5 – Spatial Navigation – Spatial Memory*, we compare physical vs. virtual navigation with a tablet (10.1″) in a zoomable landscape. Opposed to Kaufmann and Ahlström, we report a significantly better navigation performance (47% decrease in path lengths and a 34% decrease in task time), but no significant effect on users' spatial memory. Finally, Pahud et al. show for a map navigation task with a smartphone as peephole (4.3″) that physical navigation is significantly slower than virtual navigation unless navigation happens between a few known targets (Pahud et al., 2013). In the light of these contradicting results and 20 years after Fitzmaurice (Fitzmaurice, 1993), Pahud et al.'s concluding remark that the understanding of the subtleties of peepholes is still not sufficient appears particularly true. Therefore, we designed our research to explore these subtleties by isolating the peephole sizes from above studies (54.7″/4″, 10.1″and 4.3″) in a controlled map navigation experiment to understand their effect on navigation behavior, navigation performance, and task load.

4.5.2 Fitts' Law Peephole Target Acquisition Models

Another stream of related research concerns formulating models of peephole target acquisition based on Fitts' law and validating them with one-directional (Cao et al., 2008; Huber et al., 2011; Kaufmann and Ahlström, 2012) or multi-directional pointing (Rohs and Oulasvirta, 2008) or AR tasks (Rohs and Essl, 2006). While this work is of fundamental importance, we believe that for understanding the subtleties of real-world map navigation with dynamic peepholes these models are only a first step. They accurately model a subtask of navigation, namely the time and precision of pointing at a distant target. However, real map navigation is far more complex than only one-directional pointing between two targets, since it is a multi-directional task that involves recalling multiple different (off-screen) locations from a mental representation of a 2D map and navigating between them. Such map navigation also includes initial phases of learning the yet unknown locations and spatial features or, at least, reactivating them from memory. All these aspects of map navigation are not part of Fitts' law models because Fitts' law does not consider them. Fitts' law models cannot help with finding design tradeoffs for peephole size since they propose that pointing performance always gets better with growing peephole size and thus assume that "bigger is always better...". They do not take limiting factors or boundaries into account. For instance, upper boundaries like the users' maximum field of view or the practicalities of using large displays or projections or lower boundaries like the higher mental and physical demand when using small or very small peepholes. This is why we chose an experimental approach to measure the "sweet spot" for map navigation instead of attempting to approximate it using existing predictive models.

4.6 Summary & Contributions

We found that previous studies (Kaufmann and Ahlström, 2013; Pahud et al., 2013; Rohs et al., 2007) have not sufficiently explored the peephole's size as an independent variable and how it affects navigation behavior, path lengths, navigation times, and user task load. This is surprising since it seems plausible that these aspects are all strongly dependent upon peephole size. A larger peephole reduces the need for slow physical panning and search in favor of a faster visual scanning of the display's content. It also allows for recognition rather than recall from spatial memory because it reveals more visual features that support user orientation all at once.

However, a study of simulated tunnel-vision in front of large displays that included a task comparable to peephole navigation showed that the effect of a reduced peripheral vision and field of view is surprisingly small (Ball and North, 2008). If this is also true for peepholes, it will open important design opportunities. In real-world systems, larger peepholes and displays increase cost, energy consumption, and weight, and the devices become more cumbersome. Alternatives are small and lightweight handheld projectors which can produce a relatively large peephole. However, some practical problems (hand jittering, finding surfaces in the right size and lighting conditions for projection, privacy concerns when using projections in public spaces) come into play. Designers must make concessions due to these constraints. They want users to experience the benefits of larger peepholes while avoiding the many disadvantages that result from using and handling larger devices or mobile projections. Therefore answering the question of how small peepholes can become without overburdening their users during map navigation is of great practical relevance.

4.6.1 Empirical Contributions

With this study of peephole map navigation, we found a "sweet spot" between peephole size and both user navigation performance and user task load. By simulating different peephole sizes from 4″ (smartphone size) up to 120″ (control condition), we found that a smartphone-sized peephole is outperformed by all other sizes and that larger peepholes significantly improve learning speed, navigation speed, and reduce task load. However, this added benefit diminishes with growing sizes, and peephole sizes greater than a projector-phone do not pay off in terms of navigation performance or task load anymore. Our data shows that a relatively small, tablet-sized peephole can serve as a "sweet spot" in terms of both user navigation performance and user task load (see Figure 4.11, page 114).

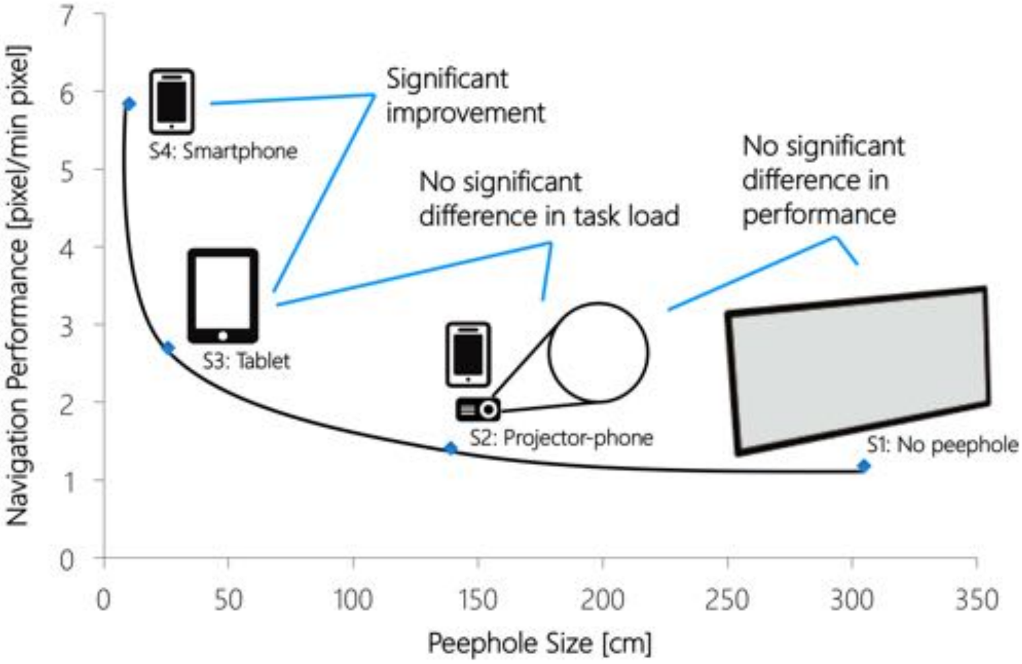


Figure 4.11.: Plot shows diminishing benefit of navigation performance with growing peephole sizes.

We have also shown that for understanding the subtleties of real-world map navigation with dynamic peepholes, existing models of peephole target acquisition based on Fitts' law (Cao et al., 2008; Huber et al., 2011; Kaufmann and Ahlström, 2012; Rohs and Oulasvirta, 2008) are only a first step. They were not intended to model different phases of map navigation such as a learning phase and a navigation phase whose existence we have shown using a statistical and visual analysis of the users' navigation paths in our study. By this, we have contributed to the better understanding of the subtleties of peephole navigation as motivated by Pahud et al. in (Pahud et al., 2013).

Parts of the next Chapter 5 appear in the following publication:

Rädle, R. Jetter, H.-C. Butscher, S. Reiterer, H. (2013a). "The effect of egocentric body movements on users' navigation performance and spatial memory in zoomable user interfaces". In: *Proceedings of the 2013 ACM international conference on Interactive tabletops and surfaces - ITS '13*. New York, New York, USA: ACM Press, pp. 23–32. DOI: `10.1145/2512349.2512811`[1]

---

[1] The responsibilities for this joint publication were divided as follows: I formulated the research question, designed and conducted the study, analyzed the study data, and spearheaded the writing. Hans-Christian Jetter helped in formulating the research question and writing the paper. Simon Butscher helped in analyzing the study data. Harald Reiterer supervised the work.

5

Spatial Navigation – Spatial Memory

> *No man has a good enough memory to be a successful liar.*
>
> — **Abraham Lincoln**
> (16th President of the United States)

Contents

RO1

This chapter builds on results of *Chapter 4 – Spatial Navigation – Peephole Size & Navigation Behavior*. We use a tablet-sized peephole as "sweet-spot" to further explore the effect on users' navigation performance and spatial memory when navigating in virtual information spaces such digital library collections.

In addition to the experiment presented in the previous chapter, we now introduce zooming to provide a more realistic navigation. We present two experimental studies comparing traditional multi-touch navigation to peephole interaction. Each experiment is explained in detail, and the results are reported and then discussed. The chapter concludes with a distinction to related work and a short section summarizing this chapter.

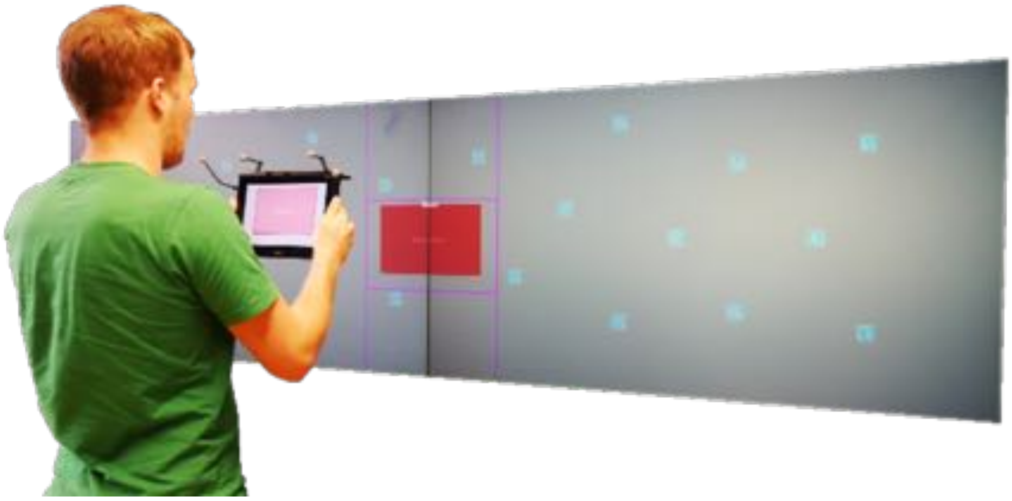


Figure 5.1.: Egocentric navigation to navigate in large information spaces projected on high-resolution displays.

5.1 Zoomable User Interfaces

Recent work in HCI indicates potential advantages of physical navigation over virtual navigation in large information spaces both for single users (Andrews et al., 2010; Ball et al., 2007; Ball and North, 2008) and for multi-user collaborative work (Jakobsen and Hornbæk, 2012). However, as discussed in *Chapter 1.2.1 – Physical Navigation in Front of Wall-Sized Screens*, even a large, high-resolution display can only provide a limited amount of information, for example because of the users' visual acuity, and it also restricts the ways in which users can interact and manipulate the information, e.g., using touch or stylus input.

ZUIs (or "multi-scale UIs") on a secondary personal device can provide the user with a second, mobile window to the large primary display. The large display's entire content becomes a virtual "canvas" that can be navigated through the mobile display using zooming and panning techniques and manipulated by touch or stylus input. This is particularly useful when navigating or collaborating in large virtual workspaces with very detailed data, e.g., high-resolution maps, satellite imagery, large documents, image collections, or large virtual whiteboards. In collaborative scenarios, ZUIs with personal devices can also enable different coupling styles of collaboration, e.g., parallel work vs. tightly-coupled collaboration (Tang et al., 2006).

Unfortunately, these advantages come at a cost. ZUIs introduce a greater cognitive load due to the frequent switches between application level (e.g., reading, annotating, search) and system level tasks (e.g., view management with panning and zooming) (Andrews et al., 2010). To reduce this load, Bederson formulates design principles for

ZUIs (Bederson, 2011). It must be possible for users to navigate within information spaces without any training and it must be difficult or impossible for them to get lost in Desert Fog (Jul and Furnas, 1998). He also states that spatial layout and movement by panning and zooming alone are not sufficient cues for memorizing content and locations. Therefore, we want to test whether an egocentric navigation can increase the efficiency and reduce the cognitive load of ZUI navigation and if its greater kinesthetic and proprioceptive feedback offer additional cues that facilitate the encoding of the location and identity of objects in the users' memory. In our study, we compared egocentric navigation with multi-touch navigation. Latter consists of traditional drag-to-pan and pinch-to-zoom touch gestures.

We isolated two main tasks that are relevant when interacting with large, high-resolution displays. First, in the navigation task, users search for specific objects on a large spatial canvas. Second, in a recall task users need to recall precisely the locations of objects based on their spatial memory. In consequence of the open research about the advantages of egocentric body movements and based on **RQ1**, we generated the following four hypotheses H1–4:

- **H1 Navigation Performance:** Users perform better in navigation tasks within a zoomable user interface when using egocentric navigation compared to multi-touch navigation.
- **H2 Spatial Memory:** The object location and identity recall capabilities of users increase with egocentric body movement compared to multi-touch navigation.
- **H3 Subjective Workload:** The workload assessment of users significantly differs between egocentric body movement and multi-touch navigation.
- **H4 Long-term Memory:** The interaction with egocentric body movement has a positive effect on users' long-term memory compared to multi-touch navigation.

5.2 Experiment 1 (E1)

The first experiment compares two navigation techniques and allows us to confirm or reject Hypotheses H1–H3. Both techniques allow panning and zooming of a large virtual canvas. However, the baseline technique uses traditional touch input and the second technique uses egocentric body movements. According to our H1 and H2, we assume that egocentric body movements will result in better spatial memory performance and navigation performance. Our experiment compares the impact of multi-touch versus egocentric navigation on a zoomable UI. For the experiment, we decided to use more abstract and generic spatial tasks instead of focusing on a specific real-world application, thus increasing the internal validity. Still, it could resemble a shelf-browsing task or another knowledge work activity where design artifacts were created on a tablet and then collected and arranged in space on a wall-sized display to project and share one's internal thoughts.

In the experiment, participants had to perform two tasks: (1) a navigation task and (2) a recall task. In the navigation task, they had to search for eight presented objects similar to a memory game (see Figure 5.2, page 122). This task is analogous to many real-world applications in which users have to search for a particular information object in a large information space (e.g., a book). The objects were repeated, which solidifies them in users' spatial memory. The second task then exemplified a task in which users had to recall the spatial location and identity of previously searched objects.

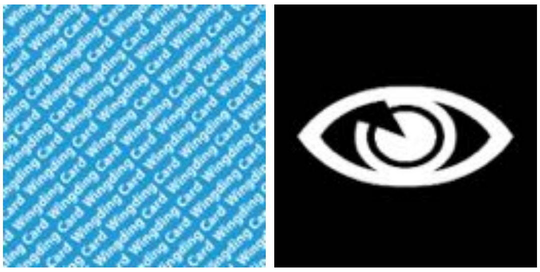


Figure 5.2.: Both sides of a memory card. The memory card hid the symbol (left) until a certain zooming level was reached and then uncovered the underlying symbol (right).

Figure 5.1 (page 120) provides an overview of the environmental setup. Both the high-resolution display and the mobile device are connected and share a large virtual canvas. Objects (memory cards) are spatially distributed on the large canvas. The mobile device provides a dynamic view of the canvas whereas the view on the large display visualizes the whole canvas statically. Two navigation techniques were provided to manipulate the dynamic view on the tablet. They are explained in the following paragraphs. Also, to raise the awareness of users, a purple rectangle was presented on the large display that represents the view currently visible on the tablet (see Figure 5.1, page 120). With this, users have the chance to orient themselves spatially and keep their global orientation without getting lost in Desert Fog.

Multi-touch condition (see Figure 5.3, page 123): The participants are provided with a PC desk and an office chair. Both were arranged by the participants at the beginning of the multi-touch condition so that they could comfortably sit and see all of the objects on the large canvas. The tablet is laid out on the desk, and users can navigate the large canvas on the tablet using conventional multi-touch drag-to-pan and pinch-to-zoom gestures. While there are alternative interaction techniques for pan and zoom, we chose this technique because it reflects the dominant design in ZUIs for multi-touch devices. To not penalize the multi-touch condition, we let participants decide whether they would like to hold the tablet in an upright position or leave it lying on the desk in front of them. All participants chose to place the tablet in the horizontal position on the desk (Scan QR Code #6 or enter “6” in the *MediaBrowser* application).

Video

Code #6

Figure 5.3.: A participant solving the navigation task in the multi-touch interaction condition.

Egocentric condition (see Figure 5.4, page 124): In this condition, the users pan and zoom in the large information canvas using peephole interaction. The technique uses the position and the distance of the tablet in relation to the large display for navigation. This technique is similar to PaperLens (Spindler et al., 2009). However, in this condition, we used a vertical display instead of a horizontal display and enlarged the interaction space to force participants to move their entire body instead of relying only on arm movements. As in peephole interaction, lateral and perpendicular movements of the device resulted in panning. Also, orthogonal movements led to zooming (move towards large display = zoom in, move away from large display = zoom out). By doing so, we could observe the effect of body movements on the users’ navigation performance and spatial memory and compare it with multi-touch navigation without body movement in the other condition (Scan QR Code #7 or enter “7” in the *MediaBrowser* application).

5.2.1 Method

We conducted a controlled lab experiment and recruited 24 subjects. The navigation technique (m = multi-touch, e = egocentric) and the object pools (A = PoolA, B = PoolB) were independent variables. The navigation performance (cost factor), the recall error rate, the navigation task completion time, and the subjective reports were dependent variables. The study was conducted as a within-subjects design with

Video

Code #7

Figure 5.4.: A participant solving the navigation task in the egocentric interaction condition.

repeated measures. It was systematically counterbalanced, and each participant was randomly assigned to one of the four groups (mAeB, mBeA, eAmB, eBmA).

Apparatus

The system consisted of two displays: a large, high-resolution, non-interactive wall-sized display and a tablet providing multi-touch input. The large display consisted of 2 eyevisCube EC-67-SXT+ (67″ each) and a total resolution of 3.840×1.080 pixels. The physical dimensions of the horizontally concatenated cubes were 292×82cm. The tablet was an Acer Iconia Tab W500 (10.1″ TFT LCD Display LED Backlight) with a resolution of 1.280×800 pixels. It provided a capacitive multi-touch input and weighs[1] approximately 940g. At this time, available Apple iPad (4th Generation) weighs 660g. We opted for the Acer Iconia Tab W500 because it natively operates with Windows 7, which allowed us to run the same software stack on the large wall-sized display as well as on the tablet. Of course, the 280g difference in weight could potentially penalize the egocentric navigation condition. We will discuss this aspect later in the NASA TLX section.

A rigid body marker was attached to the tablet to track the egocentric body movements (see Figure 5.5, page 126). The rigid body was present in both conditions to avoid any bias. For the egocentric navigation, the total interaction volume in front

---

[1] "The weight of the [Acer Iconia Tab W500] device is 940g, whereas a further 610g for the Keydock and 190g for the power adapter. – `http://www.notebookcheck.net/Review-Acer-Iconia-Tab-W500-Keydock-Notebook.53964.0.html` (last accessed: May 28th, 2016)"

of the wall-sized display was approximately 292×160×82cm. In the multi-touch navigation, participants were asked to position the PC desk at a distance so that they could comfortably see the entire large canvas. The mean distance of the PC desk to the wall-sized display was approximately 160cm.

The software was implemented in Microsoft .NET 4.0/WPF with C#. The ZUI, semantic zoom, and view synchronization were implemented with the ZOIL API (Jetter et al., 2012b). An OptiTrack motion capture tracking system from NaturalPoint with 18 cameras (10 default lenses, 8 wide-angle lenses) allowed for precise tracking of the tablet for the egocentric body movement condition. It tracks motion within a tolerance less than 0.5 mm. We implemented an OptiTrackInputModule for the ProximityToolkit (Marquardt et al., 2011b) to measure proxemic dimensions, such as the position, distance, and orientation of the tablet in relation to the large display. The mobile device was calibrated in the interaction volume so that the minimum scale factor of .75 on the mobile device was reached at a distance of 190cm and the maximum scale factor of 8.0 was reached at a distance of 30cm in front of the wall-sized display.

Participants

24 subjects (12 female) participated in the experiment. The age of participants ranged from 19 to 57 years ($\bar{x}$=25.42, SD=8.87). Age was non-normally distributed, with skewness of 2.97 (SE=0.47) and kurtosis of 8.51 (SE=0.92). Two participants were left-handed, and one participant had red-green color blindness. However, this did not bias our study since all tasks could be accomplished simply by perceiving shapes and monochromatic colors. All participants were members of the university, but none of them had a background in computer science or was affiliated with the computer science department (18 students, 2 Ph.D. students, 2 apprentices, 2 administrative staff).

Task Design

We used a different object pool with numbers to introduce the tasks and two training pools with letters that were different from the object pools that were used in the actual tasks.

During the actual tasks, 22 objects were equally distributed across the virtual canvas in each pool. Each object had a size of 50×50 pixels. Since a within-subjects design was applied, we created two different object pools with different symbols and varying object locations to prevent learning effects across conditions (see Figure 5.2, page 122). The object pools were also systematically counterbalanced throughout the two navigation conditions. Both object pools were tested in a pre-test to guarantee that the pools would be easily distinguished from each other (to avoid confusion with a previous pool) and their objects (e.g., avoid having an airplane and a helicopter in the same pool since both are aircrafts).

During the navigation task, participants had to search for 8 objects in 8 blocks (= 8 objects × 8 blocks = 64 trials). The other 14 objects served as distractors to increase the difficulty of the task. In total there were 3072 trials (24 participants × 2 conditions × 8 objects × 8 blocks).

Task 1: Navigation Task

The first task focused on the navigation performance in a ZUI. It compared navigation cost and navigation time of the two navigation techniques: multi-touch and egocentric. In the navigation task, the participants had to pan and zoom to the center of the overview display, which is referred to as "home position" in the following. The home position was visualized as a red rectangle. Both, Figure 5.3 (page 123) and Figure 5.4 (page 124) illustrate the red rectangle in the picture-in-picture at the top right.

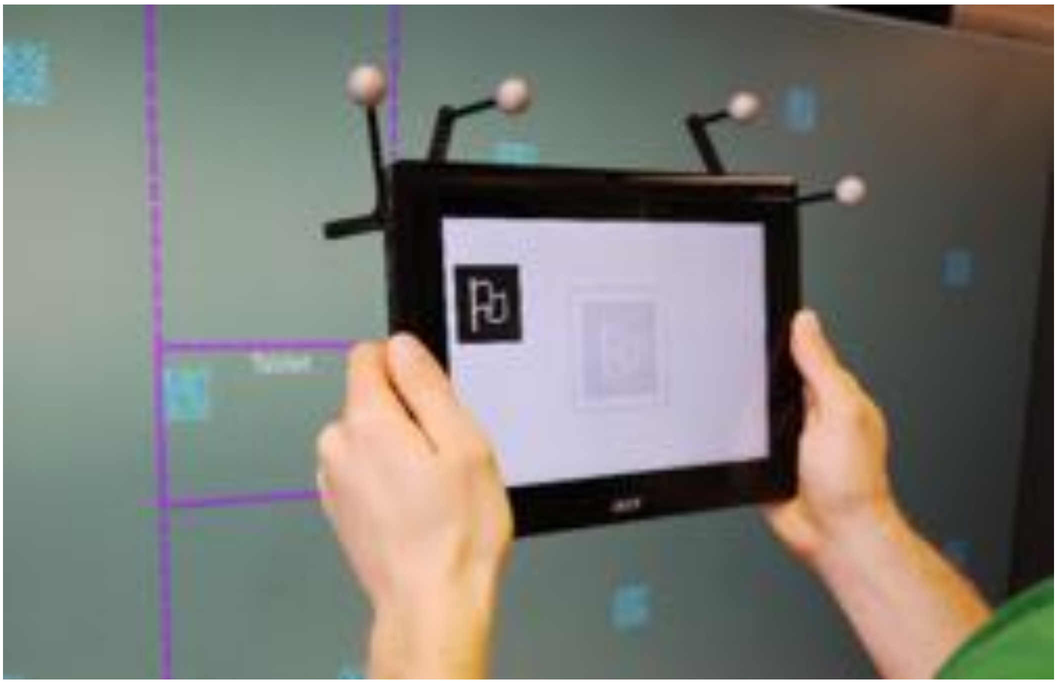


Figure 5.5.: A user navigating in the large canvas using the egocentric navigation. The purple rectangle on the overview display indicates the current position of the viewport on the large canvas.

After the tablet's view matched with the home position, it disappeared and a translucent object (see Figure 5.5, page 126) was presented in the center of the tablet's screen. Then, the participant had to find the corresponding on the large virtual canvas. Figure 5.5 (page 126), for instance, illustrates the view on the tablet displaying the uncovered flag object on the top left and the requested and translucent object in the center of the tablet's screen. A match was accepted as a sufficient if it matched with the presented translucent object on the tablet within an offset of 10 pixels in position and 30 pixels in size. The tablet revealed memory cards content only if the scale factor of the view is > 4.0 or above while the objects on the large display permanently show the back of the memory card. Also, the distance between the objects assured that only one uncovered symbol was visible on the tablet at a time, to avoid building a spatial memory based on relative relationships (e.g., object X is to the right of object Y). After matching the requested object, the home position appeared again, and the participants had to match it again in order to continue with the next object. Each match was signaled by a compound audio feedback for both the homing action and the correctly identified object.

Task 2: Recall Task

The second and consecutive recall task requested that participants recall the positions of the objects in the previous navigation task. Therefore, the participants were seated at the mobile desk and in front of the large display (see Figure 5.6, page 128). The large display showed an empty canvas. Then, the first object of the sequence from the previous navigation task was presented in the center of the large display. Participants had to place the object at their recalled position based on their spatial memory. The object could be moved with the help of the arrow keys on a keyboard. The keyboard was used to avoid any effects originating from the use of muscle memory. If they were satisfied with the current position, they confirmed it with the Enter key and the object disappeared. After that, the next object appeared in the center of the screen, and they had to continue the task until all eight objects were positioned. Each time a participant pressed an arrow key, the object moves by 25 pixels in the given direction. This helped to speed up the reproduction process instead of positioning objects pixel-wise. The original positions of the objects, however, can be matched exactly.

Each x- and y-offset to the original object was considered an error. Figure 5.7 (page 129) illustrates the grid and the calculation of the overall error that was measured as the Euclidian distance between the original location and the participants' recalled location.

Figure 5.6.: A participant performing the recall task. He is seated at the mobile desk and provided with a keyboard. Participants used the arrow keys of the keyboard to move objects to recalled positions and place them with the enter key.

Procedure

In the beginning, the participants were welcomed by the experimenter and asked to fill out a pre-test questionnaire. The questionnaire included questions regarding demographic data, their experiences of touch technology, and their handedness, which could be important for later video analysis. Then, all participants were introduced to the first navigation technique according to their group. They were allowed to practice before the actual tasks began. Through this, we wanted to avoid any learning effects based on the interaction technique and enable participants to familiarize themselves with the input method. For the practice training, each participant had to complete 4 objects with 4 repetitions (= 4 objects × 4 blocks = 16 trials). Afterward, participants were asked if they felt comfortable with the navigation technique; if they concurred, the study began. Each session lasted approximately 90 minutes, and participants were compensated with €12.

5.2.2 Results

During each trial, we logged the navigation time and path including the x- and y-position as well as the zoom factor s. Although navigation time was measured, we did not apply a Fitts' law (Balakrishnan, 2004) test because our research did not focus on the input device and its index of performance for navigation. Both navigation techniques had different sampling rates. Therefore, we had to reduce the

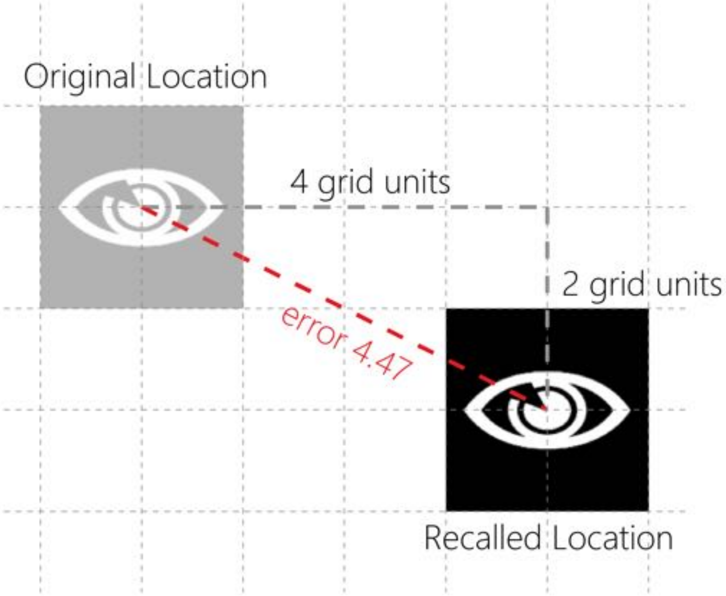


Figure 5.7.: Calculation of the recall error between the recalled location and the original location. The objects are 50×50 pixels in size and the grid cells are 25×25 pixels.

input data of both interaction techniques to a 10Hz sampling rate. On the sampled data, we applied Leifert's ZUI cost metric (Leifert, 2013) (see Figure 5.8, page 130) that is similar to the metric proposed by Jetter et al. (Jetter et al., 2012c). The variables W (width) and H (height) define the viewport size (e.g., 1280 × 800 pixels), Δx and Δy define the x- and y-movement in pixels between $t-1$ and t, Δs the scale factor between $t-1$ and t, and c is the cost of the navigation step:

$$c = H \cdot \Delta x + W \cdot \Delta y - \Delta x \Delta y + W \cdot H \cdot |log \Delta s| \tag{5.1}$$

Equation 5.1.: The original ZUI cost metric (Jetter et al., 2012c).

$$c = H \cdot \Delta x + W \cdot \Delta y + W \cdot H \cdot |log \Delta s| \tag{5.2}$$

Equation 5.2.: The modified ZUI cost metric.

The original ZUI cost metric (see Equation 5.1, page 129) calculates the updated pixels between a view t_{n-1} and the view t_n discretely. Since horizontal and vertical panning happens in parallel (see Figure 5.8, page 130) (red rectangle), Jetter et al. included the term $-\Delta x \Delta y$. Leifert removed the term $-\Delta x \Delta y$ in her cost metric (see Equation 5.2, page 129) to achieve a better approximation of the continuous movement and shifting during user interaction (see Figure 5.8, page 130).

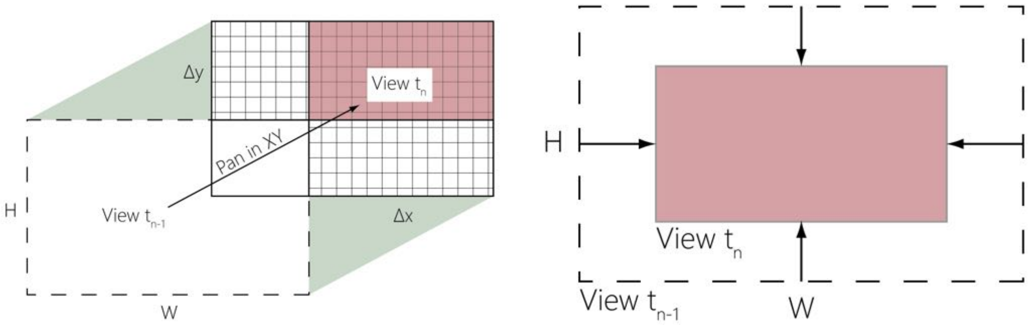


Figure 5.8.: Illustration of modified ZUI cost metric, which approximates the continuously updated pixels when the participants shifted the view (left). Illustration of cost factor for zooming (right).

Results for Navigation Performance

The resulting navigation performance per trial using the ZUI cost metric was 5,509,852.07 pixels^2 (SD = 1,513,327.43) for the multi-touch condition and 2,910,584.94 pixels^2 (SD = 671,967.12) for the egocentric condition. Block 1 was removed in both conditions since participants were navigating the objects for the first time and therefore their chosen navigation paths were entirely random. A 2×7×8 (navigation technique × block × object) analysis of variance (ANOVA) with repeated measures was performed on the traveled navigation paths. The analysis revealed a statistically better navigation performance for the egocentric navigation, $F_{1,23} = 117.03$, $p < .001$, $partial\ \eta^2 = .84$. The results show that the navigation with egocentric navigation is 47% more efficient than with multi-touch navigation; therefore, we can confirm our Hypothesis H1.

Tan et al. reported in their study that females may benefit more from kinesthetic cues provided by touch screen devices than males (Tan et al., 2002). Our results, however, did not show a significant difference for the navigation technique between the genders ($F_{1,22} = .342$, $p = .565$, $partial\ \eta^2 = .015$).

The design of the experiment can be regarded as successful since the analysis did not show a significant interaction between navigation technique and group order ($F_{3,20} = .683$, $p = .573$, $partial\ \eta^2 = .093$). The mean navigation performance for Blocks 2 to 8 are plotted in Figure 5.9 (page 131). As mentioned above, Block 1 was omitted in both conditions since participants were unfamiliar with the virtual canvas and object locations at the beginning. It illustrates a learning effect for the virtual canvas in both conditions. However, the navigation paths for egocentric navigation are shorter in an earlier stage compared to the multi-touch navigation.

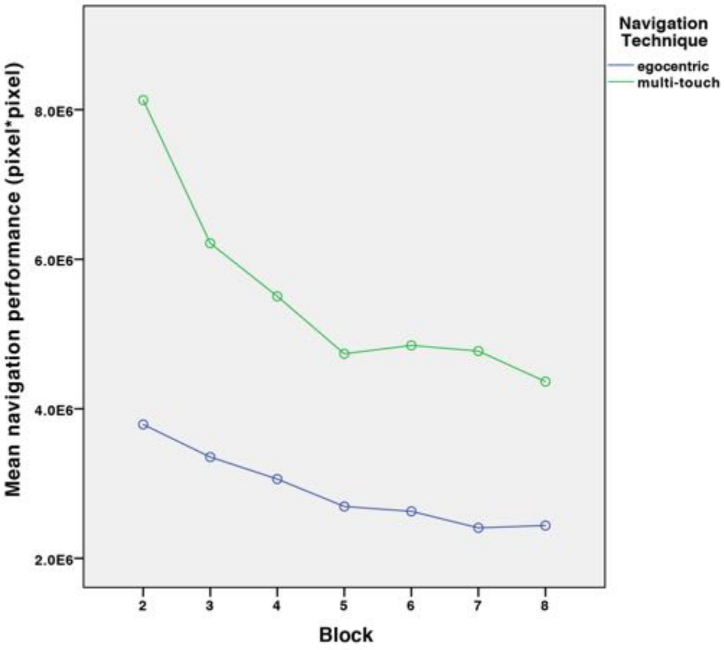


Figure 5.9.: The navigation cost per trial based on the ZUI cost factor for Blocks 2 to 8.

Results for Navigation Time

The navigation time was recorded starting when the participants matched the red rectangle and ending when they matched the target object (see Figure 5.1, page 120). The mean navigation time per trial for the multi-touch condition was 10,460.82 ms (SD = 4,102.97) and the mean for the egocentric condition was 6,868.94 ms (SD = 2,260.19). A 2×7×8 ANOVA (navigation technique × block × object) with repeated measures was performed on the collected data. The analysis revealed that the technique had a statistically significant effect on the navigation time, $F_{1,23} = 13.96$, $p < .05$, $partial\ \eta^2 = .38$. Again, there was no interaction between the navigation technique and gender ($F_{1,22} = .102$, $p = .752$, $partial\ \eta^2 = .005$). Also, the analysis of variance did not show a significant interaction between navigation technique and group order ($F_{3,20} = .1.853$, $p = .170$, $partial\ \eta^2 = .218$).

As shown in Figure 5.10 (page 132), the mean navigation time in the multi-touch condition improved quickly with every block. To ensure that the significant effect of navigation time was also present in the final block, we also performed a 2×8 ANOVA on Block 8 only. This still revealed that the navigation technique had a statistically significant effect on the navigation time, $F_{1,23} = 5.72$, $p < .05$, $partial\ \eta^2 = .199$.

The mean navigation time of both conditions is illustrated in Figure 5.10 (page 132). Again, Block 1 was omitted in both conditions since participants were unfamiliar with the virtual canvas. Remarkably, all participants learned object locations and

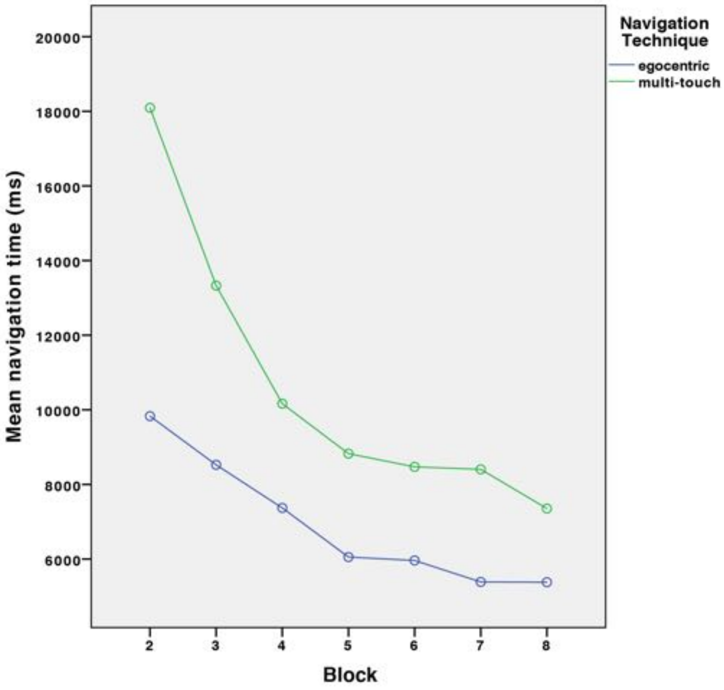


Figure 5.10.: The mean navigation time per trial for Block 2 to Block 8.

identities as the task progressed. The results show that egocentric navigation is 34% faster than in the multi-touch condition.

Results for Spatial Memory and Recall

The comparison of the results of the recall task with multi-touch and egocentric navigation did not show significant differences ($F_{1,23} = .120$, $p = .732$, $partial\ \eta^2 = .005$). Participants performed equally well in multi-touch and egocentric conditions when they had to recall the locations and identities of objects they had to search for in a previous task. The mean error in grid units was 5.69 (SD = 2.63) for the multi-touch condition and 5.39 (SD = 2.97) for the egocentric condition. Figure 5.11 (page 133) illustrates the distribution of recalled object locations (top = PoolA; bottom = PoolB). Black squares indicate the original location of objects. Colored squares show finally recalled locations (red = multi-touch; blue = egocentric). In our analysis, we also tested the serial position effect (Murdock Jr., 1962) (primacy and recency effect and found no significant differences between the first and the last objects and middle objects 2 to 7.

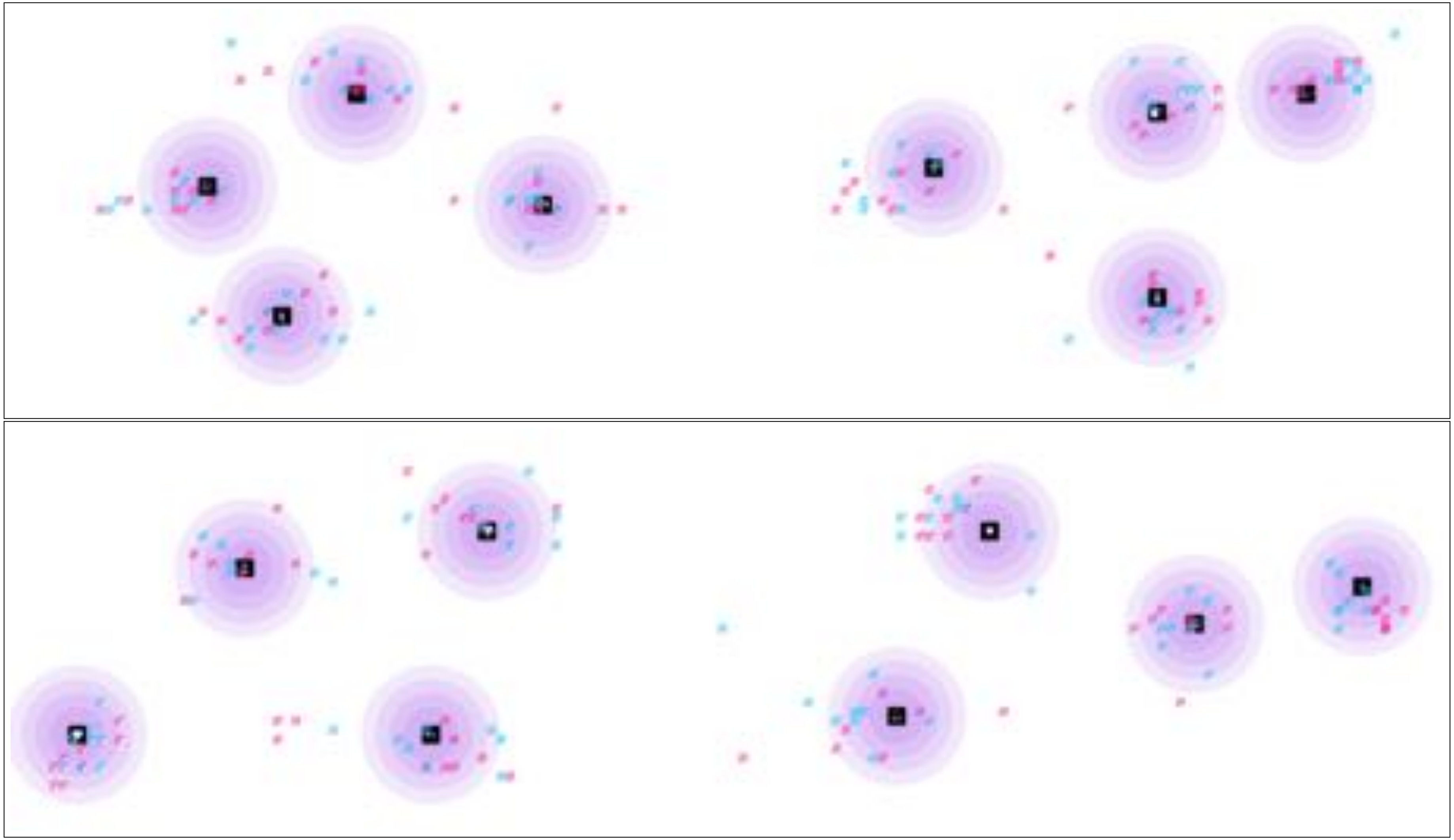


Figure 5.11.: Distribution of recalled object locations (top = PoolA; bottom = PoolB). Black squares indicate the original location of objects. Colored squares show finally recalled locations (red = multi-touch; blue = egocentric).

NASA TLX

Participants rated their subjective workload at the end of each condition. For this, we used the standardized NASA TLX. We did not find any significant difference in the overall subjective workload when comparing the navigation techniques egocentric navigation ($\bar{x}$=39.65, SE=13.43) and multi-touch navigation ($\bar{x}$=42.40, SE=17.73) (t = .753, p = .459, r = .009) (see Figure 5.12, page 134). However, if we look closely into the subscales and compare them pairwise, the two subscales mental demand and physical demand denote significant differences (see Figure 5.13, page 135). All calculated statistical significance values were based on a two-tailed paired-samples t-test.

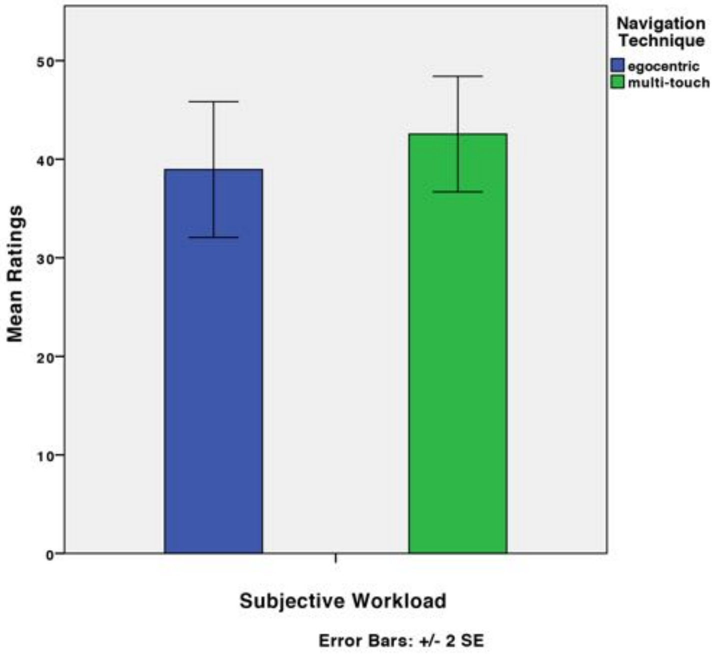


Figure 5.12.: The subjective workload assessment was measured using the NASA Task Load Index. Overall subjective workload did not reveal any statistical significant differences.

On average, participants reported a significantly greater mental demand for multi-touch navigation ($\bar{x}$=53.75, SE=4.67) than for egocentric navigation ($\bar{x}$=40.83, SE=4.17), $t(23) = 2.262$, $p < .05$, $r = .43$. Probably, this is reasonable because users needed to transform recalled spatial information into 2D touch gestures for the multi-touch navigation whereas in the egocentric navigation users applied knowledge used in everyday life.

Unsurprisingly, on average, participants reported significantly less physical demand for multi-touch navigation ($\bar{x}$=36.25, SE=5.267) than for egocentric navigation ($\bar{x}$=55.42, SE=4.94), $t(23) = -3.960$, $p < .05$, $r = .64$. Of course, participants had to move around and hold the tablet (940g) in laborious positions in the egocentric

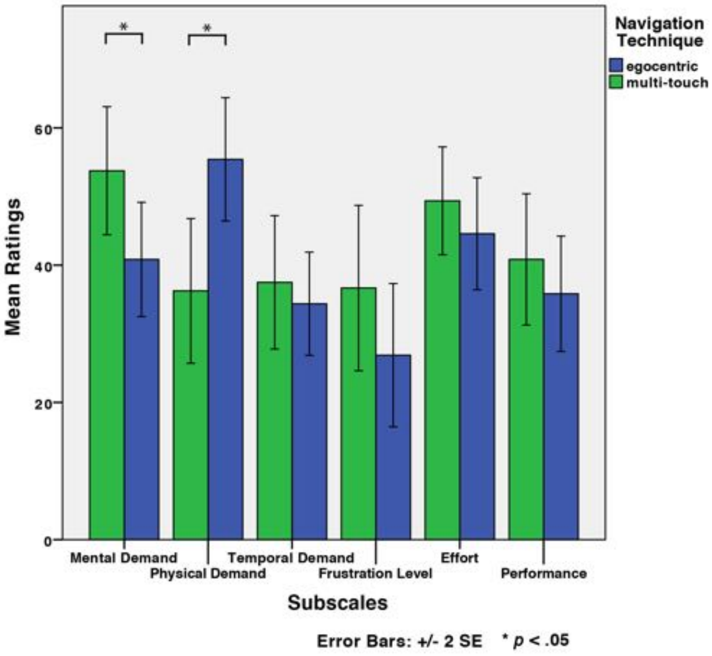


Figure 5.13.: The subjective workload assessment was measured using the NASA Task Load Index. Pairwise comparison showed significant differences for mental demand and physical demand.

navigation, while in the multi-touch condition, the participants were seated and could rest their arms on the table. This difference, however, might diminish with future mobile and lightweight device technologies and only further strengthen benefits of egocentric navigation. For example, the latest iPad Air (2nd Generation) and 9.7″ iPad Pro both weigh 437g, which is more than half weight of the Acer Iconia Tab W500.

Even though the pairwise comparison of individual NASA TLX subscales did reveal significant differences, the calculated overall subjective workload of users does not significantly differ between egocentric body movement and multi-touch. Therefore, we can reject Hypothesis H3.

Results of Subjective Preference

Nineteen out of 24 participants preferred the egocentric navigation over the multi-touch navigation. The participants that preferred multi-touch reported very frequent usage of touch devices in the pre-test questionnaire (e.g., smartphones). Also, two participants who do not own multi-touch devices stated that egocentric navigation "*feels more natural*" to them. This reinforces the applicability of Reality-Based Interaction framework and Blended Interaction framework for the design of interactive UbiComp systems.

5.2.3 Discussion

The key finding of our experiment was that egocentric navigation performs significantly better in terms of traveled path length and time. Apparently, in the egocentric condition, participants naturally applied a navigation style with combined zooming and panning operations that resulted in more efficient navigation paths. We observed how users navigated between two targets by moving the device back or forth and simultaneously moving laterally in front of the large display. This resulted in navigation paths that are close to the "optimal" paths in a ZUI with exponential savings (Jetter et al., 2012c). In the multi-touch condition, however, users made use of this combined movement less often and more often alternated between pure zooming or panning operations. This resonates with the findings of Jetter et al. (Jetter et al., 2012c) that observed the same for the multi-touch condition in their study of navigation performance in a ZUI on a tabletop.

Since the results showed significantly better navigation performance with egocentric navigation, but no significant differences in spatial memory performance, we cannot attribute the improved navigation to better spatial memory. This still could mean that there was an additional memory source applied during the spatial interaction tasks, e.g., motor memory. However, this assumption has not been verified by our experiment and remains an open research question.

In addition, a review of the recorded video exposed an interesting observation. Participants, especially those with less experience with multi-touch devices, approached incorrect targets repeatedly within the same trial in the multi-touch condition. This happened even though participants were provided with the current view on the large display (see Figure 5.1, page 120) (purple rectangle). We assume that this was the case because users in the multi-touch condition tended to lose their global orientation, i.e., the awareness of where their current view on the tablet is located within the canvas. Users in the egocentric condition, however, used their physical position in front of the large display as a spatial cue to maintain this global orientation.

This resonates with the results of the NASA TLX questionnaire, in which egocentric navigation was rated as significantly less mentally demanding than multi-touch. Thus, it leaves more cognitive resources for system-level tasks because fewer resources were used for application-level tasks (e.g., zooming and panning). This benefit also shows the willingness of the users to accept the significantly greater physical demand (53%), since 19 out of 24 participants preferred the egocentric navigation technique over multi-touch for the given task. Besides, the heavy weight of the tablet eventually affected the other five participants' choice and future studies with lightweight tablets would even lead to clearer results in favor of egocentric navigation technique.

5.3 Experiment 2 (E2)

In a second experiment E2, we wanted to study if the greater proprioceptive and kinesthetic feedback of body interaction would lead to stronger encoding as well as fixation of object location and identity in users' spatial memory as formulated in our Hypothesis H4.

The experiment measures the effect of the two navigation techniques egocentric versus multi-touch on users' long-term memory. We randomly selected eight participants from the first experiment E1; two participants from each of the four groups: mAeB, mBeA, eAmB, and eBmA. Eight is the minimum number of participants needed to counterbalance conditions for experiment E2 using a Latin square study design. Since experiment E1 already lasted approximately 90 minutes, we did not want to push all participants' patience. Therefore and at the end of experiment E1, we asked a few to take part in an additional but shorter experiment until all slots were filled. Of course, the number of participants limits the statistical power and thus does not yet allow for generalization of results. However, it can be seen as a pre-study that allows judging whether further investigation in this research direction is promising.

5.3.1 Participants & Procedure

The eight participants (3 female) were aged between 19 to 30 years ($\bar{x}$=22.75, SD=3.20). Participants were distracted for 15 minutes showing them a game on a tabletop. The distraction was deliberately planned to ensure that participants did not think or talk about the previous experiment. Participants were then asked to recall the objects from the endmost navigation technique after the 15 minutes of distraction. The recall task was the same as in Experiment E1. Again, the error was calculated based on the original location of the object and the recalled location (see Figure 5.7, page 129).

5.3.2 Results & Discussion

The mean error for the multi-touch condition was 6.26 (SD=1.09, SE=.54) and the mean error for the egocentric condition was 4.56 (SD=.32, SE=.16). A one-way ANOVA revealed significant differences for the navigation technique, $F_{1,6} = 8.979$, $p < 0.05$, $\omega = .71$, indicating that egocentric body movements increase long-term spatial memory (see (Tang et al., 2006)). The spatial memory indicates an improvement of 27% in favor of the egocentric navigation.

Although the results may show a significant improvement for long-term spatial memory in egocentric navigation, it is too early to generalize this for the entire user population because of the small sample size (N=8). However, this second experiment is a good starting point for further investigation of the effect of egocentric body movements on users' long-term spatial memory and to examine whether the underlying mechanisms differ from those for spatial memory from E1.

5.4 Related Work

Our work here draws upon the existing studies that measured the influence of interaction and navigation techniques or visualizations on users' spatial memory and navigation performance.

5.4.1 Spatial Interaction Techniques and Bodily Movements

Fitzmaurice demonstrated spatial interaction with a 3D virtual environment in his Chameleon prototype (Fitzmaurice, 1993). Although he did not study the effect on users' spatial memory when using spatial interaction, the Chameleon prototype enabled an early experimental evaluation of the advantages of spatial interaction, and he observed a positive user experience.

Later, Peephole (Yee, 2003) interaction used a small mobile device and no additional overview projection to navigate in a large virtual canvas. The researchers implemented different usage scenarios for interaction with 2D and 3D environments. The focus of their work is on the study of different interaction techniques, including selection tasks, navigation, and manipulation such as drawing. Although they mention that spatially-aware displays enable the use of spatial memory for navigation, they have not conducted an experiment that supports this statement.

Spindler et al. combined a spatially-aware display with a magic lens (Bier et al., 1993). Their tangible PaperLens (Spindler et al., 2009) enables physical navigation in volumetric, layered, or zoomable information spaces. In a following work, they also studied the minimum thickness of layers, users' holding accuracy, and the physical boundaries of the interaction volume (Spindler et al., 2012). However, they have not studied the effect of tangible magic lenses on users' spatial memory.

Ball et al. conducted a user study that measured the effect of physical vs. virtual navigation in a 2D virtual space when interacting in front of a large, high-resolution display (Ball et al., 2007). Participants were allowed to move freely in front of the display and all interaction with it was performed with a Gyration GyroMouse. Ball et al. found that physical navigation outperforms virtual navigation for tasks such as navigation, search, or pattern matching, but the effects on users' spatial memory were not tested.

5.5 Summary and Contributions

In this chapter, we presented two consecutive experiments that research the effect of egocentric body movements on users' navigation performance, spatial memory, and long-term spatial memory. In contrast to the related work mentioned above, our research specifically investigates the effect of egocentric body movements versus multi-touch on users' navigation performance and spatial memory when interacting with a zoomable user interface.

The results of Experiment E1 shows a significantly better navigation performance in a ZUI for egocentric navigation in terms of path length (47%) and task time (34%). We provide possible explanations based on a more frequent use of combined zooming and panning operations that resulted in more efficient navigation paths. Participants also reported a significantly lower mental demand which resonates with our observation that in the egocentric condition, they used their physical position in front of the large display as a spatial cue to maintain their global orientation.

Parts of the next Chapter 6 appear in the following publications:

Rädle, R. Jetter, H.-C. Marquardt, N. Reiterer, H. Rogers, Y. (2014c). “HuddleLamp: Spatially-Aware Mobile Displays for Ad-hoc Around-the-Table Collaboration”. In: *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces - ITS ’14*. New York, New York, USA: ACM Press, pp. 45–54. DOI: `10.1145/2669485.2669500`[2]

Rädle, R. Jetter, H.-C. Marquardt, N. Reiterer, H. Rogers, Y. (2014b). “Demonstrating HuddleLamp: Spatially- Aware Mobile Displays for Ad-hoc Around-the-Table Collaboration”. In: *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces - ITS ’14*. New York, New York, USA: ACM Press, pp. 435–438. DOI: `10.1145/2669485.2676584`[3]

---

[2] The responsibilities for this joint publication were divided as follows: I spearheaded the technical implementation of HuddleLamp and did the comparative technical evaluation of the tracking. The writing of the paper was divided among myself, Hans-Christian Jetter, and Nicolai Marquardt. Hans-Christian Jetter also contributed to the technical implementation of HuddleLamp. Nicolai Marquardt did all sketches. Harald Reiterer helped designing HuddleLamp interaction techniques and gave continuously advise on this work. Yvonne Rogers and Harald Reiterer supervised the work.

[3] This publication is accompanying a demo session. The technical demonstration of HuddleLamp received a **People’s Choice Best Demo Award** at the *International Conference on Interactive Tabletops and Surfaces 2014* in Dresden. Responsibilities for this joint publication were divided same as (Rädle et al., 2014c)

6 Cross-Device Interaction – Enabling Technology

> *Complain about the way other people make software by making software.*
>
> — **Andre Torrez**
> (Former CTO of Federated Media Publishing)

Contents

RO2

Empirical research and the research-oriented design approach in *Chapter 3 – Context & Analysis* best shows the potential of Integrative Workplace for cross-document and cross-media interaction. However, a user study also revealed two technical issues of the system. Its (i) *tracking quality & accuracy* is often unreliable and inaccurate and requires (ii) heavy *instrumentation of environment*.

In this chapter, we present a different and alternative technical solution to cross-document interaction using multiple mobile devices. It only requires a single low-cost and off-the-shelf RGB-D camera mounted above a table (e.g., integrated into a desk lamp). This technology enables a new kind of computer-supported around-the-table collaboration without interactive tabletops. Users can still sit around ordinary tables that can remain cluttered with non-digital objects (e.g., printouts, maps, notebooks) while their digital collaborations happen using spatially-aware mobile screens that blend into existing spatial and social practices. Unlike traditional interactive tabletops, HuddleLamp needs only a few low-cost and off-the-shelf hardware components, so it can be used in improvised settings or where the costs for hardware and administration typically prohibit the use of large interactive tabletops, e.g., public libraries, schools, community centers. In the remainder of this chapter, we focus on our following contributions:

As our primary contribution, we introduce our novel hybrid sensing approach for HuddleLamp. This approach combines RGB and depth input for detecting and tracking movements of multiple mobile screens with sub-centimeter precision by exploiting their optical characteristics in both the RGB and IR range. We evaluate the tracking quality of our approach regarding accuracy, precision, and reliability with a controlled experiment and discuss capabilities and limitations.

As secondary contributions, we first introduce our web-based architecture and JavaScript API for enabling truly ad-hoc, walk-up-and-use applications with no need for installing any native mobile apps or instrumenting devices with any markers or hardware before collaborating. Second, we validate our architecture and API with five example interaction techniques that used to be possible only in instrumented rooms or on and above large tabletops. They demonstrate the future design space for spatially-aware multi-device around-the-table collaboration with HuddleLamp.

6.1 HuddleLamp's Technical Setup

Software

Code #8

In the following, we describe HuddleLamp's technical setup, and its architecture and algorithms. To facilitate replication outside our lab, we used only off-the-shelf hardware and free or open source software components. We provide the source code as open source (Scan QR Code #8 or enter “8” in the *MediaBrowser* application) .

6.1.1 Hardware Components

HuddleLamp uses a low-cost short range time-of-flight (TOF) depth camera which delivers a 1280×720 RGB color image and a 320×240 depth image at 25 or 30 fps. It shares its technical specifications with a Creative Senz3D or SoftKinetic DepthSense 325. The camera is fixed to an Artemide Tolomeo Tavolo desk lamp in which it replaces the light bulb (top right corner of Figure 6.1 (page 146)) (Scan QR Code #9 or enter “9” in the *MediaBrowser* application). Using the lamp, users can conveniently move the camera into its downward-facing operating position that lies 0.8m above the horizontal surface to track. This results in a rectangular tracking region of approximately 1.0×0.6m from which the camera receives sufficient RGB and depth information to track mobile devices, their spatial configurations, and users’ hands and to distinguish them from non-interactive objects.

The camera was chosen for its small size (110×30×25mm) and low-noise depth data. A further advantage of this particular camera and its Perceptual SDK by Intel is that its RGB and depth images can be aligned without calibration. Therefore, it is easy to retrieve RGB values for depth pixels and vice versa. This facilitates the processing of data in our hybrid sensing approach. However, in principle, hybrid sensing should also work with other TOF cameras (e.g. Kinect v2) and a higher resolution and larger field of view.

For vision processing and for communicating with the mobile devices via a web socket server, we use a Windows PC or laptop. For better portability, we have considered integrating a single-board PC (e.g. Raspberry Pi) directly into the lamp. However, at this stage, the vision processing is still too computationally expensive to achieve our targeted tracking rate of 25-30 fps with ARM CPUs.

6.1.2 Software Components & System Architecture

The computer vision application for processing the RGB and depth data was implemented in C# for WPF/.NET 4.5.1 and Emgu CV (OpenCV bridge for C#). For finding and decoding fiducial markers in the RGB stream we use the glyph decoder of the AForge.NET library[1]. While the vision application is active, the incoming camera data is processed, and all the identified device and hand positions, device

---

[1] AForge.NET is C# framework for computer vision and artificial intelligence – `http://www.aforgenet.com/` (last accessed: March 30th, 2016)

Video

Code #9

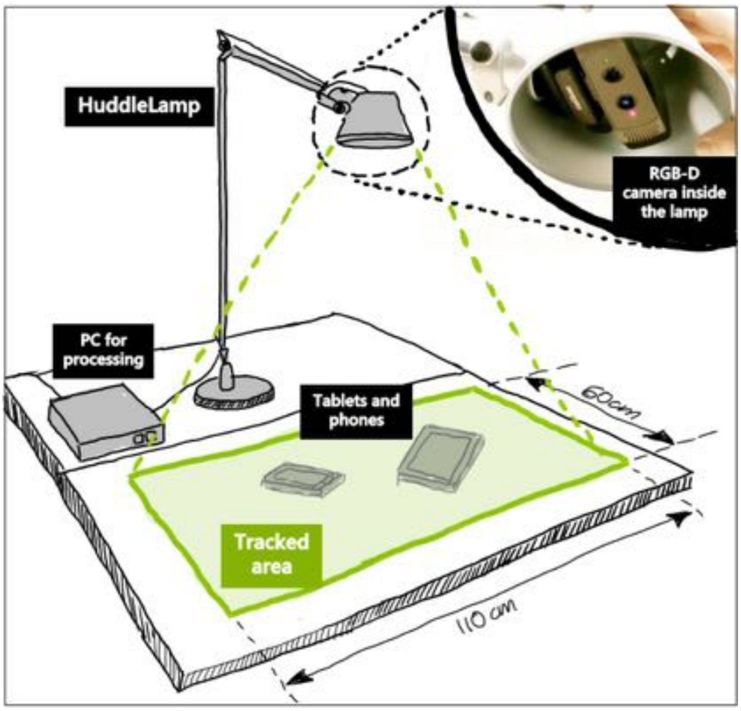


Figure 6.1.: Technical setup of HuddleLamp with an integrated RGB-D camera (Tracking region: 1.0×0.6m).

orientations, and device occlusion states are streamed as JSON to connected mobile devices via a web socket server.

Client applications on the mobile devices access this stream using HuddleLamp's JavaScript API for browsers such as Safari (Mobile), Chrome, and Internet Explorer. If desired, the API also provides a shared virtual workspace that is spatially-situated in the physical tracking region and can be accessed by the devices. It uses a shared object storage implemented on top of the Meteor web platform[2] to distribute rendering and interaction across devices.

6.2 HuddleLamp's Hybrid-sensing

HuddleLamp's main purpose is to identify and track mobile screens of different sizes with sub-centimeter precision and to distinguish them from other objects or hands. To this end, we combine RGB and depth information and verify our results with virtual fiducial markers. This "hybrid sensing" compensates for the limitations of sensing RGB, depth, IR, or fiducials alone and, unlike (Li and Kobbelt, 2012; Lucero et al., 2010; Marquardt et al., 2011b; Marquardt et al., 2012a; Merrill et al., 2007;

---

[2] Meteor is an open source platform for the web, mobile, and desktop – `https://www.meteor.com/` (last accessed: May 2nd, 2016)

Rekimoto and Saitoh, 1999), works without instrumenting devices or rooms with custom radio-hardware, markers, or tags.

6.2.1 Detection of Mobile Screens Using Low IR Reflectance

The first step of hybrid sensing is detecting the regions that possibly contain mobile screens. For this, we use two optical characteristics that all mobile screens we worked with have in common. First, they are obviously rectangular and thus can be easily recognized as rectangles in a camera image. Our second characteristic, however, is not visible to the eye, and we only learned about it during own experimentation.

We found that mobile screens generally have a very low reflectance for the modulated IR signal that is emitted by TOF cameras. Therefore, whenever a screen enters the view, the screen and its bezel absorb rather than reflect the IR signal. The reflected signal becomes so weak that the camera cannot reliably measure depth and returns "low confidence" for most pixels inside of screens. However, depth values are available for pixels outside of screens, e.g., from the table's surface. Figure 6.2 (page 148) shows this in side-by-side comparisons of RGB images (left) and depth confidence images (right). The right images are created simply by drawing all pixels that have a depth value in black and all "low confidence" pixels in red ("low confidence" means signal intensity is < 87; intensity ranges from 1 to 32,767). The top row of Figure 6.2 (page 148) shows low IR reflectance for devices of very different size, generation, screen type, screen brightness, and screen content (1. Samsung/Google Nexus 10; 2. LG/Google Nexus 5; 3. Nokia Lumia 620; 4. Apple iPad Air; 5. Lenovo ThinkPad Yoga; 6. Microsoft Surface 2 Pro; 7. Apple iPad 3; 8. Apple iPod Touch; 9. Samsung Galaxy S2; 10. Apple iPhone 4; 11. Nokia 106).

The second row illustrates that low IR reflectance (here that of an Apple iPad Air) cannot only be observed when devices are perpendicular to the camera but also for other angles (from bottom to top: 45°, 31°, 22.1°, 18.6°).

The third row shows that many everyday objects (from left to right: sheet of paper, notepad, watch, cup, back of Apple iPad, book) have normal to high IR reflectance and their depth is measured correctly. They, therefore, disappear among the other black pixels in the depth confidence image. The other objects (from left to right: Microsoft Surface tablet under notepad, Microsoft Surface keyboard, Apple logo on the back of Apple iPad, Apple iPhone on a book) have low IR reflectance and thus show in the right image as red shapes. Generally speaking, apart from screens, black or dark gray objects with a matte non-glossy finish are likely to have a low

Figure 6.2.: Low IR reflectance of mobile screens and objects.

IR reflectance. However, as we explain below, such false positives can be easily identified as non-screen objects at a later stage.

The novelty of our screen tracking approach lies in not discarding but using these regions of low IR reflectance with missing depth values. The only other approach that we are aware of that uses IR reflectance (in that case structured light from a Kinect v1) to distinguish materials (e.g. skin from paper) is (Steimle et al., 2013). We can achieve a similar segmentation of screens from non-screens "for free," by using a TOF camera and a threshold for signal strength without additional computation. Unlike Dippon et al. (Dippon et al., 2012), we can, therefore, avoid the need for object detection and segmentation algorithms that use actual depth data.

In Figure 6.4 (page 151), the RGB image R1 and the depth confidence image D1 show the example of eight mobile screens on a table under realistic lighting conditions. Thereby D1 is almost entirely unaffected by ambient light, screen content, screen backlighting, or even the bright reflections from overhead light sources in R1. Since IR tracking works with its own light sources, D1 would also look very similar in a room that appears completely dark to the human eye. In general, IR reflectance tracking is very reliable concerning all kinds of other light sources, and it is very easy to extract the positions and orientations of screens from an image like D1 just by using low-pass filtering, Canny edge detection, and finding contours and convex hulls. The result of this process is shown in image D2 and contains 7 out of 8 screens.

6.2.2 Sauron's Eye

Only using the depth confidence image would however have limitations. One limitation lies in D1's centre region, which we informally refer to as "Sauron's eye"[3] (see Figure 6.4, page 151). Here the IR signal from the camera hits the screen surface with almost 90°. The IR signal is directly reflected back into the depth sensor with great intensity and saturates a small region of pixels in the center. This results in a small red ellipse in the centre of D1 that is the "pupil" of Sauron's eye (like low confidence pixels, saturated pixels in D1 are red). Around this inner ellipse, the reflectance remains high, but not high enough to saturate. Therefore there is another elliptical region around the saturated pixels, for which the depth values can be read. They show as a black "iris" region around the pupil that only gradually turns into red pixels when moving further away from the center. Finally, Sauron's eye is contained by the device bezels that typically have a lower IR reflectance than screens.

6.2.3 Hybrid Sensing: Fusing RGB and IR Detection

While the depth confidence image enables reliable tracking without strong interference from other light sources, Sauron's eye negatively affects the tracking quality in its center. The closer screens get to it, the less visible they are in D1. For example, one smartphone from R1 becomes almost invisible in D1 with only parts of its bezels still visible. Another disadvantage of D1 is its low resolution of 320×240 pixels or less compared to R1 with 1280×720 pixels. This reduces the accuracy of tracking positions and orientations. To compensate for these disadvantages, we employ a hybrid sensing approach that complements the results from depth confidence tracking with those of RGB tracking in R1. This RGB tracking method uses standard processing steps such as binary thresholding, low-pass filtering, Canny edge detection, and finding contours and convex hulls (see Figure 6.3, page 150). R2 in Figure 6.4 (page 151) shows the result with five of eight possible screens from the example RGB image R1.

---

[3] Sauron's primary appearance in Peter Jackson's movie trilogy Lord of the Rings is as the Eye of Sauron. This visual appearance looks very similar to the resulting effect of direct IR illumination as seen in Figure 6.4D1 (page 151).

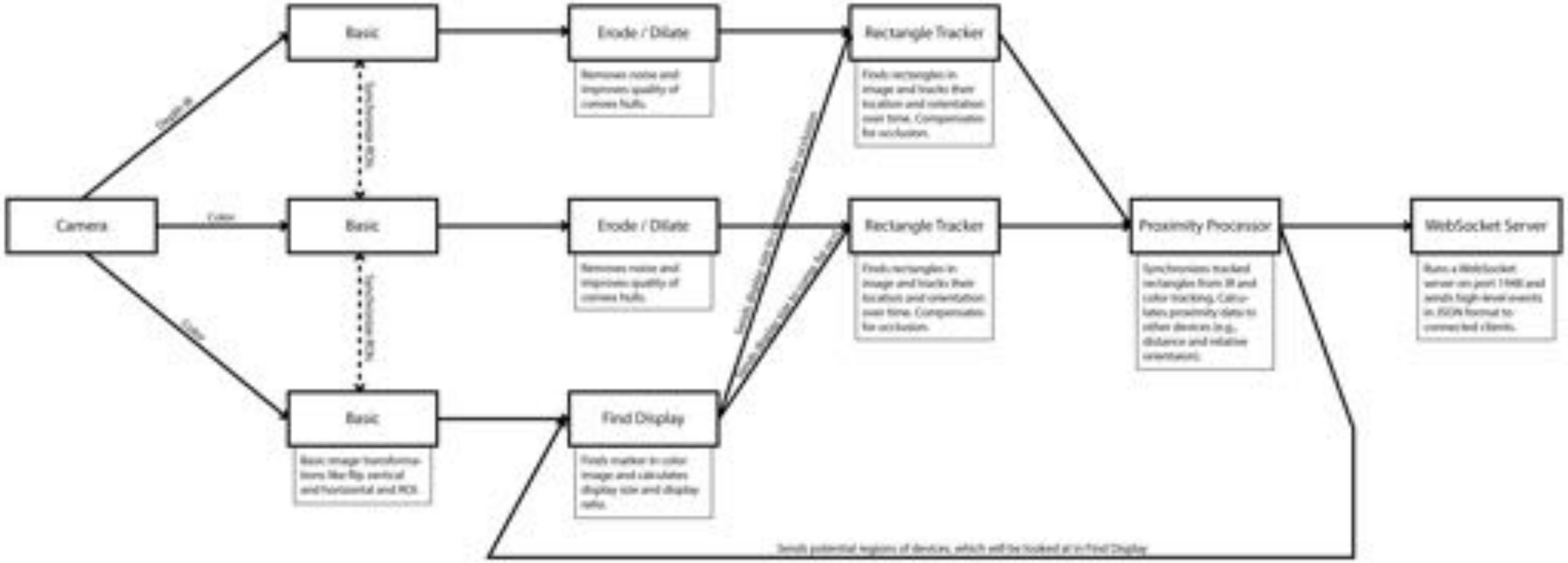


Figure 6.3.: Schematic illustration of HuddleLamp tracking pipeline. It shows each processing node with a brief description of its function and further illustrates the image and data flow in the pipeline.

It is important to notice the limitations of this simple RGB tracking. While it does not have a blind spot in the center and has higher accuracy because of the higher resolution, the reflections from overhead light sources can easily deform device contours or cut through them, so that they are not recognized as screens anymore. Also, the devices' bezels must always have a color that is clearly distinguishable from the table's surface. In summary, the RGB tracking is less reliable but more accurate and thus can serve to improve tracking whenever it can provide more accurate positions and orientations or the depth confidence tracking fails. Therefore, RGB and depth tracking mutually complement each other. Consequentially, the merged result in H1 (see Figure 6.4, page 151) contains all eight screens.

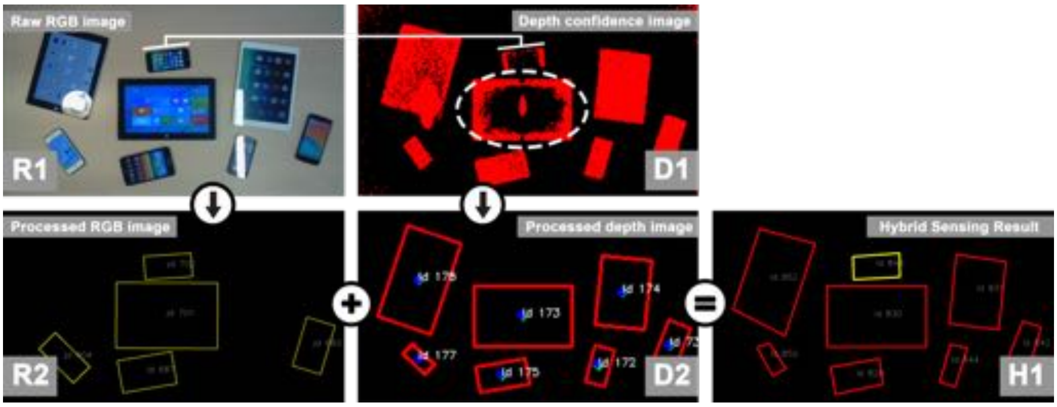


Figure 6.4.: Hybrid Sensing. R1: RGB camera input. D1: Depth confidence image containing "Sauron's eye" (dashed ellipse). R2: Mobile screens detected in R1. D2: Mobile screens detected in D1. H1: Merged result of hybrid sensing from R2 and D2.

6.2.4 Identification of Displays, Display Size, and Orientation

After having determined the positions of rectangles that potentially are screens in the current frame, the vision processing associates them with the history of tracked rectangles from previous frames. This association is done by smoothing all previous positions of each rectangle using a Kalman filter and pairing these predicted points with the closest rectangle in the current frame. This enables the vision system to track the rectangles' movements over time, assign them with an internal ID, and look into their past.

However, the detected rectangles in Figure 6.4 (page 151) are not necessarily actual screens. At this stage of the processing, the result contains regions that are only likely candidates for being a screen, but can also be false positives, e.g., the aforementioned dark non-glossy objects with low IR reflectance and a rectangular shape. For example, tablet covers or protective pouches made from felt. To identify

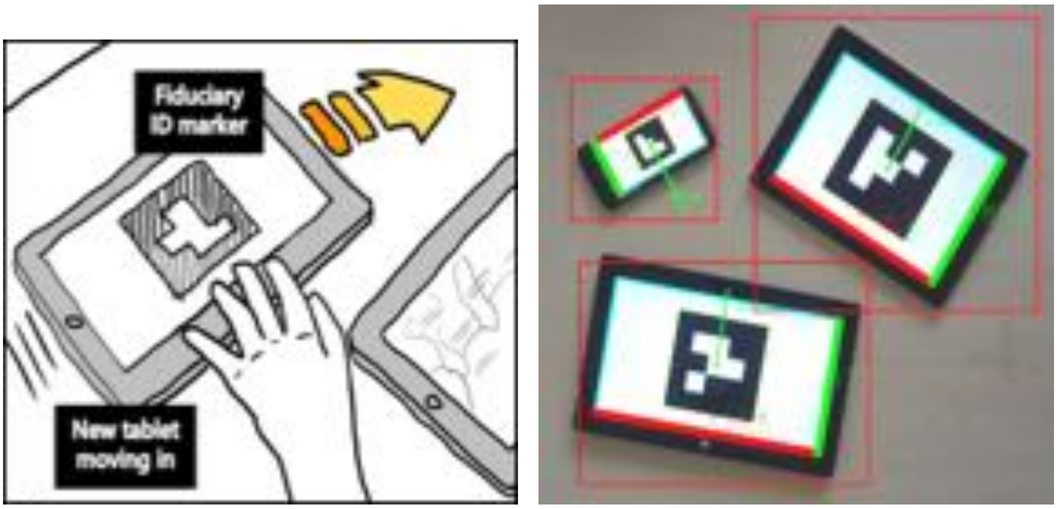


Figure 6.5.: The vision system determines ID, orientation, and screen size using fiducial markers on a white background.

such false positives, we ask the screens to display fiducial markers and by this verify that they are truly interactive displays. To this end, the vision system broadcasts an identification request to all connected, but so far unidentified, devices. They react by displaying a fiducial marker containing a unique ID on a white background that covers their screen (see Figure 6.5, page 152). The vision system then looks for the presence of such a marker in the RGB image and decodes their value to verify that a new device has joined the huddle (and not a false positive) and to associate the corresponding region with the device ID. This spontaneous device association procedure, namely pattern matching (Chong et al., 2014), usually takes only 1-2 seconds, and the fiducial markers disappear immediately once a device has been identified.

This identification process in the RGB image also enables us to determine the orientation and size of the screen in camera coordinates. The initial orientation is that of the found marker and the screen size is determined by the size of the white background surrounding it. This data is later used by the vision system to track the current orientation and also by our JavaScript API to compensate for different screen sizes, orientations, and resolutions (see below).

6.2.5 Hand Detection and Tracking

HuddleLamp's hand detection and tracking happens entirely in the depth image. Using background subtraction, depth thresholding, and flood-fill segmentation, we identify contours in the depth image that could contain an arm and hand, based on the assumption that users always reach into the camera from outside (see Figure 6.6, page 153). The hand position can then be approximated by the centroid of the hand's depth minima. Like for the screens, these positions are tracked over time by using a Kalman filter and a distance threshold.

Figure 6.6.: Segmentation of hands and arms (white overlays) and estimated location and height of hands (red circles).

Our hand tracking is not intended for providing a detailed representation of finger positions or detecting hand gestures, but to provide light-weight low-precision information about hands' locations and depths for cross-device or above-the-table interaction techniques. For more precise manipulations of on-screen objects, users can continue to use the low-latency and high-precision multi-touch detection of their capacitive touch screens. For example, in the interaction techniques described below, we only used our hand tracking for low-precision tasks, such as using the hand position to determine the destination device for a cross-device "pick, drag, and drop" gesture. The exact on-screen position for dropping the object is determined using the device's touchscreen and not the hand tracking.

6.2.6 Technical Evaluation of Hybrid Sensing

We conducted a controlled experiment to evaluate precision, accuracy, and reliability of hybrid sensing and RGB-only tracking. They are defined as follows:

- Precision is the standard deviation of the tracked position and orientation of a fixed tablet over time and thus measures noise and jittering (units: mm or degree).
- Accuracy is the spatial accuracy of a tablet's position compared to a ground truth (unit: mm).
- Reliability is the percentage of frames in which a present tablet was tracked (its tracking state is "true") during measurement (unit: %).

Apparatus

As apparatus we used the setup in Figure 6.7 (page 154) with a tracking region of 1020×570mm and a camera height of 780mm. The effective usable resolution for the RGB image was 1280×720 pixels and for the depth image 283×159 pixels. We used a black Apple iPad 2 as mobile device.

Figure 6.7.: Setup of technical evaluation.

For the experiment, we studied seven different conditions divided into three categories: Lighting, Occlusion, and RGB-only. The Lighting and Occlusion conditions used hybrid sensing. In RGB-only hybrid sensing was turned off.

The Lighting conditions had different levels of illumination: 20 lux, 1600 lux, 2200 lux. For comparison: 400 lux is the recommended illumination for offices and classrooms and 1000 lux for hospital examination and treatment, or for difficult industrial assembly. In the 2200 lux condition we simulated strong ambient light using two R7 halogen lamps (400W) with a peak emission in IR (approx. 800nm) (see Figure 6.7, page 154).

The illumination for the three Occlusion conditions was held constant at 1600 lux. The conditions simulated occlusion with 1 finger, 1 hand, and 2 hands (see

Figure 6.8.: Picture taken from camera input stream shows the three levels of occlusion: 1 finger, 2 hands, and 1 hand.

Figure 6.8, page 155). The RGB-only condition was also at 1600 lux but without simulating any occlusion.

Procedure

In each condition, a tablet was systematically moved to 21 different positions on a table where it was fixed for a frame-by-frame measurement of tracking state, position, and orientation for 10 seconds (see Figure 6.9, page 156). Positions were defined by a 3×7 grid on the table. Two angle brackets with scales were used for tablet movement to guarantee exact positioning before measurement. The grid size was empirically determined in a pre-test. It provides sufficient detail for data gathering and compensates for eventually distorted camera images. Additional rows and columns did not add to the technical evaluation quality. In the grid, adjacent positions were exactly 100mm distant to each other. The camera image was centered on the center point of the grid. The grid distance served as ground truth for accuracy measurement.

Results

The results show sub-centimeter precision and accuracy for all conditions (see Table 6.1, page 156). As expected, the best precision and accuracy was achieved for RGB-only due to the higher resolution of the RGB image but with a low reliability of only 89.5% due to the reflections of ceiling and ambient light sources. Under the same conditions, switching to hybrid sensing increased the reliability to 100% with only a small decline in precision and accuracy due to increased dependence from the lower resolution depth image when RGB tracking failed. Still, accuracy remained below 2.3mm and thus well inside the sub-centimeter range.

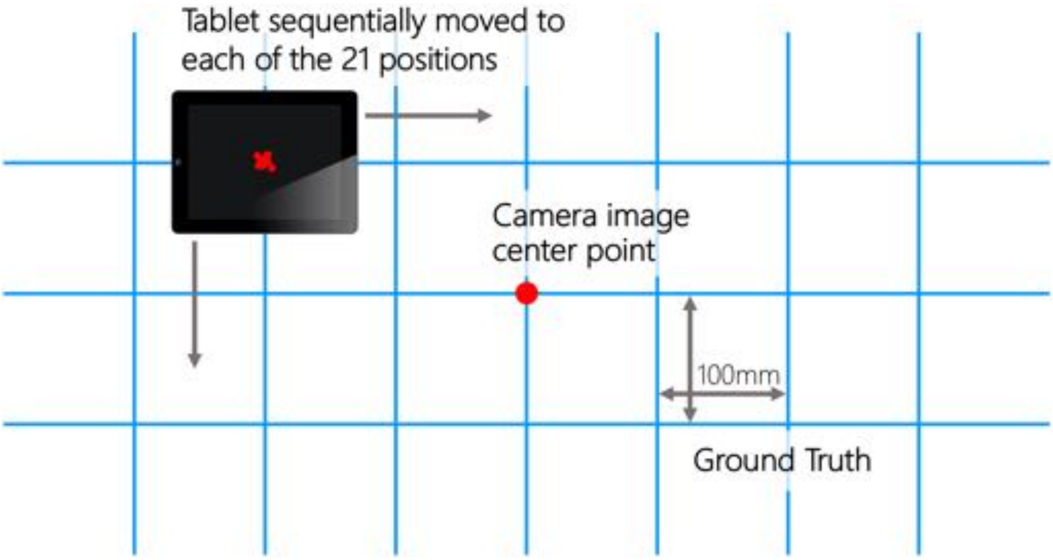


Figure 6.9.: Grid with 21 target positions denoted as grid intersection points. The tablet was systematically moved to each position and held steady for 10 seconds during measurement. Adjacent target positions where 100mm distant to each other.

For the 1 hand and 2 hands occlusion, the reliability of hybrid sensing decreased to 89.8% since it sometimes failed to track occluded tablets close to Sauron's eye. In contrast, 20 lux or 2200 lux did not negatively affect the 100% reliability of hybrid sensing, but its accuracy and precision. Nonetheless, the worst accuracy was still below 3.5mm and well inside the sub-centimeter range. As we discuss below, there is, however, an upper limit for ambient light.

Condition	**Precision** [in mm or degree]			**Reliab.** [in %]	**Accuracy** [in mm]			
	Mean (of all SD points)			**SD**	**Mean**		**SD**	
Lighting	**X**	**Y**	**Angle**		**X**	**Y**	**X**	**Y**
20 lux	1.22	1.26	.32	100.0	3.34	3.40	2.16	1.48
1600 lux	.78	1.05	.29	100.0	2.10	2.24	1.47	1.67
2200 lux	1.25	1.24	.37	100.0	3.29	2.80	2.25	1.41
Occlusion								
1 finger	.95	.74	.17	100.0	2.33	2.68	1.65	1.92
1 hand	1.08	1.02	.32	99.9	2.27	2.43	1.97	1.71
2 hands	1.36	1.81	.44	89.8	3.61	3.49	2.88	2.25
RGB-only								
1600 lux	.60	.63	.14	89.5	1.78	1.13	1.24	.70

Table 6.1.: Precision and accuracy of HuddleLamp's hybrid sensing and RGB-only tracking.

6.2.7 Challenges and Limitations of Hybrid Sensing

The natural segmentation of mobile devices using their low IR reflectance worked robustly for different lighting conditions and we achieved to process the hybrid RGB and depth input at the maximum frame rate of the camera (30 fps) with an Intel Core i7 laptop. As we discuss below, there is still a perceptible latency on the UI, but it is not originating from hybrid sensing, but from web sockets and rendering on mobile devices which are slower than on PCs.

Since we use an overhead camera, users' arms and hands can sometimes occlude devices and deform device contours during touching or moving mobile screens. As the experiment shows, this can decrease reliability to 89.8%, so that devices and their ID are sometimes lost and a fiducial marker must be flashed again for device identification. This can interrupt the users' flow of interaction, especially when reflections of overhead light sources inhibit marker recognition and make it necessary to move the device in a reflection-free area. We could potentially improve this by using alternative means of optical device identification, e.g., flashing full-screen color sequences as in (Schwarz et al., 2012).

Finally, there is an unavoidable upper limit of ambient light for all consumer TOF cameras with an integrating CMOS detector. They stop working as soon as the IR light that is reflected into the detector from the sun or other light sources becomes many times stronger than the modulated camera signal. Therefore, HuddleLamp's depth confidence tracking does not work outdoors on bright days or if sunlight shines directly on an indoor table. However, this upper limit was not reached even in our 2200 lux condition.

6.3 HuddleLamp's JavaScript API

For enabling truly ad-hoc multi-device collaboration, it is necessary to provide an API that lets devices of all operating systems easily join the devices on the table without prior installation of native apps or applications. HuddleLamp achieves this by providing an API for writing collaborative applications with HTML5, CSS3, and JavaScript, so that the application becomes simply a web page that can be opened in every device's browser. As soon as this web page is loaded, users can put their device on the table into the camera's view and by doing so add that device to the huddle. For the users' convenience, a QR code with that URL can be attached to the lamp or desk. After joining a huddle, the web-based application can access the JavaScript API to make optional use of three key features.

```
var huddle = Huddle.client()
  .on("devicefound", function() {
    // tracking found device
  })
  .on("devicelost", function() {
    // tracking lost device
  })
  .on("proximity", function(data) {
    var location = data.Location;
    var x = location[0];          // global x-position between 0.0 and 1.0
    var y = location[1];          // global y-position between 0.0 and 1.0
    var angle = data.Orientation; // global device orientation
    var pres = data.Presences;    // array of hands or other devices
  })
  .connect(host, port);
```

Listing 6.1: HuddleLamp's JavaScript API to access the data stream from the vision server.

First, it provides a web socket connection from the application to the vision server that returns a JSON data stream with device and hand positions, device orientations, and device occlusion states. This data is also provided as events using the JavaScript observer pattern (see Listing 6.1, page 158).

Second, the API provides a shared virtual workspace that can be accessed from all mobile devices. This workspace contains objects that can be arranged in space using the well-known multi-touch user manipulations from tabletops, e.g., drag, pinch-to-zoom-and-rotate, flick. However, in contrast to tabletop SDKs (e.g. the ScatterView control of the Microsoft Surface SDK), it is synchronized for all connected devices via a server. We implemented a shared object storage on top of the Meteor web framework, so that all manipulations on one device become instantly visible on all other devices to enable collaborative cross-device work.

Third, the API enables a homogeneous rendering of the workspace across different devices with respect to their different locations, screen sizes, and resolutions. Each device can become part of a multi-device display that renders the virtual workspace on the individual screens as if the workspace was physically situated in the tracked region on the table. Only that part of a workspace (e.g. a map) and its objects (e.g. images, videos) is rendered that lies underneath the device, correctly preserving absolute positions and orientations. The API achieves this by translating, rotating, and scaling the local rendering of the workspace based on the devices' individual screen size, position, orientation, and aspect ratio that were determined by hybrid sensing. There is no need to create a database of devices with their screen sizes. The API thus enables a fully interactive and reconfigurable Junkyard Jumbotron[4].

---

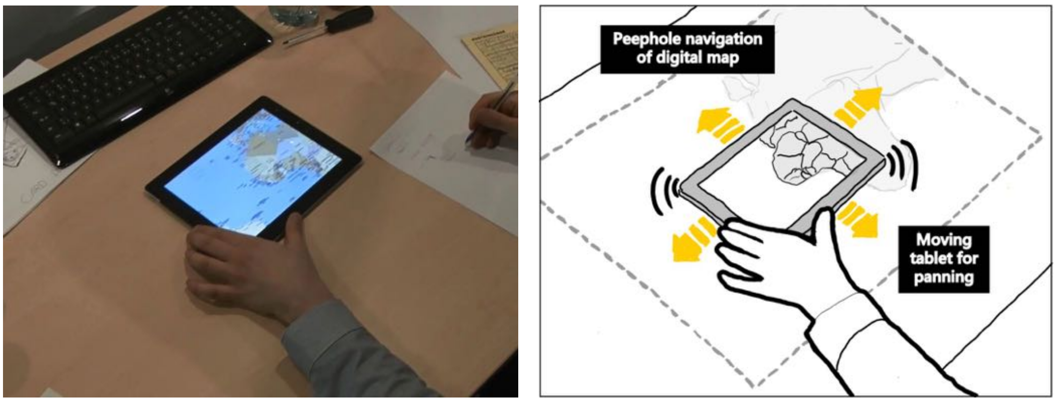
[4] Junkyard Jumbotron, Rick Borovoy and Brian Knep – `http://jumbotron.media.mit.edu/` (last accessed: March 31, 2016)

6.3.1 Validation of API with Example Interaction Techniques

To validate the design of our API, we implemented five examples of existing and novel cross-device gestures and interaction techniques that used to be possible only in instrumented rooms, as part of fixed installations or exhibits, or above instrumented surfaces. These examples serve as validation, but they also demonstrate the possible design space for future HuddleLamp applications.

Peephole Navigation

As presented in *Chapter 4 – Spatial Navigation – Peephole Size & Navigation Behavior*, peephole navigation is a promising technique for navigating large information spaces. HuddleLamp's API with a spatially-situated workspace enables peephole navigation with one or more mobile displays of different sizes. For example, we built a demonstrator in which users can move one or more tablets to navigate physically in a virtual world map (see Figure 6.10, page 159) (Scan QR Code #10 or enter "10" in the *MediaBrowser* application).



Video

Code #10

Figure 6.10.: Peephole navigation with HuddleLamp: physically navigating a large virtual map with a tablet.

Huddle Navigation

We extended peephole navigation to the novel technique of "huddle navigation." It enables users to create large multi-device displays that are similar to interactive tabletops simply by moving multiple tablets or smartphones side-by-side (see Figure 6.11, page 160). By this, users let devices join into a "huddle." Users can zoom, rotate, and pan all screens in the huddle synchronously with multi-touch. They can also create two or more huddles based on device proximity. Each huddle then zooms, rotates, and pans independently to support collaboration of multiple users or user groups (Scan QR Code #11 or enter “11” in the *MediaBrowser* application).

Video

Code #11

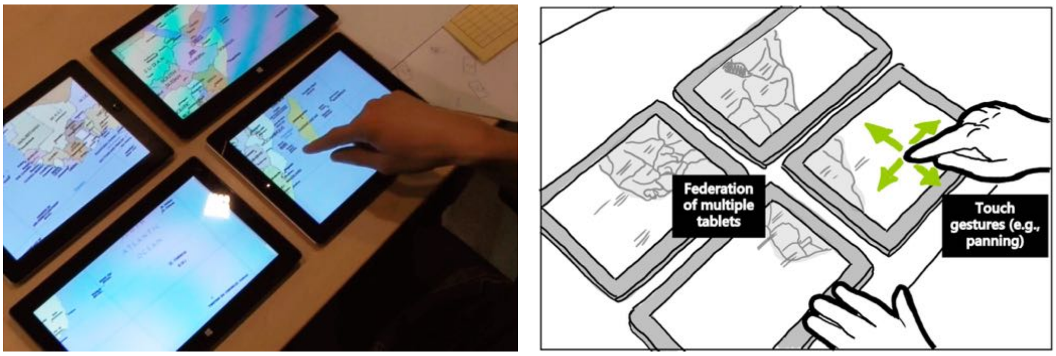


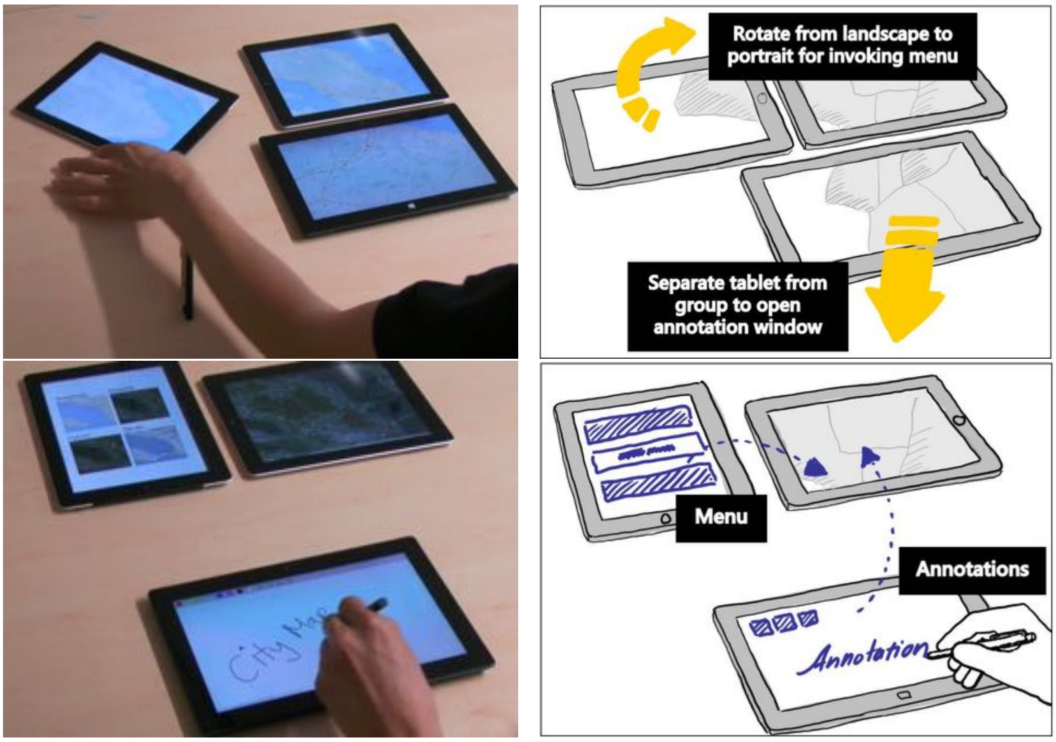
Figure 6.11.: An illustration of synchronous huddle navigation.

Spatially-Aware Menus and Modes

In the previous example, a tablet that was moved close to a map huddle automatically turned into an extension of the map. But we also used proximity and orientations for more complex behavior such as "spatially-aware menus and modes". For example, when a user rotates a tablet in a huddle from landscape to portrait orientation, this reveals a tool palette or menu where display parameters for the entire huddle can be altered, e.g., visible layers or data points of a map. When moving the tablet away from the huddle and closer to the user, the tablet automatically switches into note-taking mode and users can use a stylus to take personal notes or for annotating content (see Figure 6.12, page 161) (Scan QR Code #12 or enter “12” in the *MediaBrowser* application). With a good mapping of spatial relations to switching menus or modes, a HuddleLamp application can achieve a seemingly intelligent spatially-aware behavior that proactively supports the users during collaboration. This also enables smooth transitions between tightly-coupled collaboration (tablet is shared in the huddle) and loosely-coupled parallel work (tablet is picked up and used as personal display).

Cross-Device Flicking and Touch & Flick Browsing

Cross-device flicking is an interaction technique to enable users to flick objects with their fingers (see Figure 6.13, page 161) (left). A simple physics simulation makes objects accelerate and stop at new locations in the workspace. Objects can stop on the same screen, but also fly to the screen of another device.

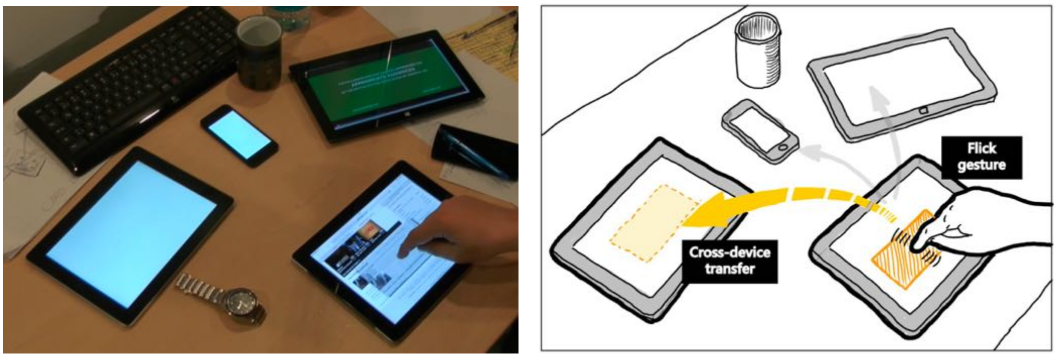


Video

Code #12

Figure 6.12.: Spatially-aware menus and modes change the role of devices based on their orientation or distance.

A variation of cross-device flicking is "touch-and-flick browsing". We implemented this technique for web browsing with multiple tablets or smartphones. Content items from a web page such as a link, video, or image can be opened on another screen simply by touching and flicking them towards a neighboring device. The content is then shown on the destination device using the entire available screen space, similar to how links can be opened in new tabs or new windows in desktop browsers (Scan QR Code #13 or enter “13” in the *MediaBrowser* application).



Video

Code #13

Figure 6.13.: Cross-device flicking and Touch & Flick Browsing.

Pick, Drag, and Drop

A different way of moving objects between devices is the "pick, drag, and drop" hand gesture. An example of such an above-the-table object manipulation technique that uses 3D graphics and physics simulations has been discussed in (Hilliges et al., 2009). It has inspired us to implement our own "above-the-tablets" technique, but with using a simpler 2D implementation. The gesture is initiated by tapping an object on a source mobile screen to pick it up and then lifting the hand and moving it above the desk to drag it to another destination device (see Figure 6.13, page 161) (Scan QR Code #14 or enter “14” in the *MediaBrowser* application). During "pick, drag, and drop," HuddleLamp continuously tracks the hand position and the picked object travels with the hand over the surface and across the screens lying beneath it. To drop the object, users move their hand above the destination device and then tap its touchscreen at the destination location. Inspired by Hilliges et al. (Hilliges et al., 2009), the current height of the hand above the surface is also used to increase the rendered size of the object to create a simple illusion of 3D movement in space.

Video

Code #14

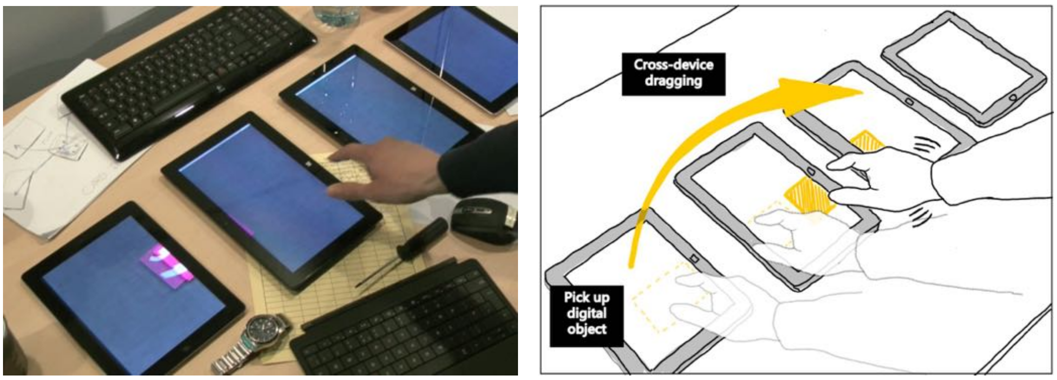


Figure 6.14.: Cross-device pick, drag, and drop.

6.4 Related Work

HuddleLamp relates to four different strains of previous research that we sample and discuss below.

6.4.1 Spatial Interaction in Instrumented Environments

The use of space and spatial configurations plays a key role in human cognition (Hollan et al., 2000) and natural social interactions (Greenberg et al., 2011). To leverage our spatial skills for interaction, one approach is to create instrumented rooms that capture the spatial interactions and configurations of users and objects as system input. For example, LightSpace uses multiple depth cameras and projectors to provide interactivity on and between physical surfaces (Wilson and Benko, 2010). The steerable Beamatron (Wilson et al., 2012) projects directly into the environment, and the LuminAR actuated desktop lamp design allows touch detection and projection onto desks (Linder and Maes, 2012). As a mobile solution, SurfacePhone (Winkler et al., 2014) uses a similar approach to facilitate single or multi-user around-the-phone interactions through projected interfaces. The OmniTouch (Harrison et al., 2011) wearable device introduces a high flexibility of where and when such interfaces are displayed. We focus on enabling ad-hoc device assemblages and interaction around tables with the mobile devices people already carry, such as phones and tablets. Conductor (Hamilton and Wigdor, 2014) recently introduced techniques for orchestrating such cross-device interactions, which HuddleLamp facilitates by automatically tracking spatial relationships between devices, allowing for fluent transfer and sharing.

Greenberg argued that we can enhance the interaction with ubicomp technology by designing systems that consider fine-grained inter-entity relationships, such as the distance or orientation between people and devices (Greenberg et al., 2011). As proxemic interactions this has been applied in various contexts, such as the interaction with a large surface media player (Ballendat et al., 2010), games (Greenberg et al., 2011), or interactive advertisements (Wang et al., 2012). Later, the GroupTogether system considers people's spatial F-formations (Kendon, 2010) during small group collaboration to initiate cross-device interactions (Marquardt et al., 2012a). The majority of these systems rely on high-end motion capturing systems (Marquardt et al., 2011b) that are ideal for prototyping interaction techniques but are difficult to deploy in environments out of the lab. To enable lightweight, ad-hoc scenarios of use, our system keeps the effort for instrumentation minimal, only uses a single RGB-D camera for tracking, and does not require any additional augmentations of devices.

6.4.2 Presence, Pairing, and Position of Co-located Mobile Devices

To enable cross-device interactions (Lucero et al., 2011), the system first needs knowledge about the presence and position of other devices around it. Various techniques establish such connections: synchronous gestures (Ramos et al., 2009) that pair devices when stitching stroke across the screens (Hinckley et al., 2004); other approaches include shaking devices simultaneously (Holmquist et al., 2001), bumping (Ramos et al., 2009), or performing pinching gestures (Ohta and Tanaka, 2012). Another approach is to use custom sensing hardware, for example infrared-, hall-, or radio-based position sensing. Siftables use infrared emitters and transceivers for detecting nearby cubes (Merrill et al., 2007). Likewise, mobile devices with magnets and hall sensors can sense presence (Huang et al., 2012), or instrumented phones can utilize custom-built radio tracking for positioning (Lucero et al., 2011) — though often with relatively coarse-grained spatial resolution (~1m). Similarly, GroupTogether requires custom-built radio-based position trilateration for device positions (Marquardt et al., 2012a). As explained shortly, our novel sensor fusion approach brings this tracking to a new level by not requiring any radio-based trilateration hardware and allowing precise sub-centimeter tracking.

Alternatively, built-in cameras can infer device location information. Back facing phone cameras can infer relative positions from extracted features in the downward-facing camera images (e.g., legs, feet) (Dearman et al., 2012). Similarly, Schmitz et al. (Schmitz et al., 2010) and Li and Kobbelt (Li and Kobbelt, 2012) use the front facing camera of devices to detect fiducial markers on the ceiling. Improving this tracking for reliably handling different environmental conditions remains an ongoing research challenge. Alternatively, cameras positioned above a table can track the spatial layout of devices placed on the table below. Rekimoto and Saitoh's tracked devices' location by attaching fiduciary markers (Rekimoto and Saitoh, 1999), an approach later also applied to track multiple devices for DynamicDuo (Piazza et al., 2013a; Piazza et al., 2013b). Kray et al. later used a related approach to determine the position of phones in one of three discrete spatial zones that trigger sharing actions between the phones (Kray et al., 2008). Taking these approaches further, we consider both depth- and RGB-tracking data to improve tracking quality for devices placed on a table.

6.4.3 Reconfigurable Tiled Displays

A special case of co-located mobile devices is (re)configurable tiled displays. ConnecTables (Tandler et al., 2001) allow dynamically reconfigurable display assemblages when devices are in proximity. At a larger scale, Phone as a Pixel (Schwarz et al., 2012) creates large-scale ad-hoc displays composed of smaller devices, each serving as a pixel of a large virtual display. The position of each pixel/device is calculated by decoding a sequence of colour transitions that encodes the ID of that device with a camera (Schwarz et al., 2012). Similarly, the web-based Junkyard Jumbotron[4] system stitches together devices' screens to a single large virtual display. Even though it is not built for real-time tracking, it allows a manual process where users take pictures of fiduciary markers shown on the screens and send it to a server for calibration processing. One approach towards allowing the display of dynamic digital content are Schmitz et al.' (Schmitz et al., 2010) and Li and Kobbelt's (Li and Kobbelt, 2012) ad-hoc multi-displays for mobile interactive applications, though their tracking performance or accuracy is not yet tested.

6.4.4 Above and around the surface interactions

For more expressive interactions, the interaction space with surfaces can be extended to include the space above and around the surface. Recent work explored this continuous interaction space with lenses above (Spindler et al., 2009) or around tabletops (Marquardt et al., 2011a), with techniques for picking up and manipulating content on an interactive touch surface (Hilliges et al., 2009). TangibleLenses (Spindler et al., 2013) and FlexPad (Steimle et al., 2013) enable handheld interactions with rigid or flexible surfaces, and LightSpace allows bi-manual gestures to transfer content between multiple surfaces (Wilson and Benko, 2010). One important design goal of the HuddleLamp system was to enable similar cross-device interactions (including tracking people's gestures), around the ad-hoc assemblages of phones and tablets on a table.

6.5 Summary and Contributions

We have described HuddleLamp, a sensing system in the form of a desk lamp with an integrated RGB-D camera that tracks the movements of multiple mobile displays on a table for around-the-table collaboration. We described our implementation and our approach of hybrid sensing, i.e., detecting mobile screens by exploiting their optical properties in the IR range and additionally using RGB images and fiducial markers to track screens better and to distinguish them from other objects or users' hands.

After a technical evaluation of our hybrid sensing approach, we also introduced HuddleLamp's web-based architecture and JavaScript API for ad-hoc collaboration that enables users to add or remove displays and reconfigure them in space at any time without installing any software.

We discussed how this can be used to create large multi-device tiled displays for multi-user and multi-touch interaction. Beyond our five demonstrated examples of different interaction techniques, we believe that HuddleLamp's setup and hybrid sensing tracking will allow the rapid exploration of future cross-device interactions supporting group collaborations.

We will do so, in a first step in the next chapter. Thereby, we explore the usefulness of spatial cues for cross-device sensemaking.

Parts of the next Chapter 7 appear in the following publication:

Rädle, R. Jetter, H.-C. Schreiner, M. Lu, Z. Reiterer, H. Rogers, Y. (2015). "Spatially-aware or Spatially-agnostic?: Elicitation and Evaluation of User-Defined Cross-Device Interactions". In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*. New York, New York, USA: ACM Press, pp. 3913–3922. DOI: `10.1145/2702123.2702287`[5]

---

[5] The responsibilities for this joint publication were divided as follows: I formulated the research question, designed studies (phase 1 and phase 2), conducted study (phase 2), analyzed the study data (phase 1 and phase 2), and spearheaded the writing. Hans-Christian Jetter equally helped in formulating the research question, writing the paper, planning study (phase 2), and analyzing the study data (phase 1). Mario Schreiner and I implemented the research prototype (phase 2). Zhihao Lu co-planned and conducted the study (phase 1). However, he analyzed the study data independently of this publication and reported his results in his master thesis (Lu, 2014). Harald Reiterer and Yvonne Rogers supervised the work.

7

Cross-Device Interaction – Understanding Spatial Cues

> *A point of view can be a dangerous luxury when substituted for insight and understanding.*
>
> — **Marshall McLuhan**
> (Canadian Communications Professor)

Contents

With HuddleLamp and our hybrid sensing approach presented in the previous chapter, it became possible to track the positions and movements of off-the-shelf mobile devices on tables and also above-the-table hand movements. The availability of these new technologies calls for more research on spatially- aware versus spatially-agnostic interactions, especially as the necessary sensing was still considered futuristic and far beyond current sensing technology just two years ago (Hamilton and Wigdor, 2014).

We have therefore conducted a user study to better explore the design space for future cross-device interaction with multiple mobile devices in two phases. In Phase 1, we conducted four elicitation sessions, each with a group of 4-5 participants. In each session, we elicited user-defined gestures for 19 typical cross-device tasks ignoring current technological restrictions and leaving the question of spatially-aware versus spatially-agnostic interaction to users. In phase 2, we used a low-cost tracking system to implement one spatially-agnostic and two spatially-aware techniques that were suggested in phase 1 and evaluated them in a controlled experiment with 12 participants.

In the following, we introduce previous and related work, report about our two-phased user study and its results and we conclude by summarizing our findings in the form of design recommendations for future cross-device systems.

7.1 Classification of Cross-Device Interactions

Today, performing tasks among multiple devices is often tedious (Greenberg et al., 2011) but new UIs promise to achieve a more natural use of multiple devices (Chen et al., 2012; Hamilton and Wigdor, 2014) with (1) new interaction techniques, (2) better sensors, and (3) user-defined gestures.

We differentiate between three main categories of cross-device interactions: synchronous gestures, spatially-agnostic interactions, and spatially-aware interactions.

7.1.1 Synchronous Gestures

Chen et al. propose a multi-tablet system for single-user reading activities that uses synchronous "conduit" interactions to move information between devices (Chen

Figure 7.1.: Three identified cross-device gesture categories: Synchronous Gestures, Spatially-Aware Gestures, and Spatially-Agnostic Gestures.

et al., 2012; Chen et al., 2013). These interactions use temporal simultaneity and sequence to express directed cross-device actions. For example, first, users designate a target by touching a device with their non-dominant hand and then they use the dominant hand, which offers more precision, to tap the item to be transferred. This approach is loosely based on Hinckley's and Rekimoto and Saitoh's pioneering work that suggested synchronous gestures such as device bumping (Hinckley, 2003), pen-based stitching (Hinckley et al., 2004), or synchronous tapping (Rekimoto and Saitoh, 1999) for exchanging content or creating multi-device tiled displays. Similarly, Lucero et al. used synchronous touch-based "pinching" across phones to create multi-device huddles in (Lucero et al., 2010; Lucero et al., 2011).

7.1.2 Spatially-Agnostic Interactions

In contrast, Hamilton and Wigdor's multi-tablet system Conductor uses traditional menus with color-coded device names to select the tablets to share information with or to chain tasks across them (Hamilton and Wigdor, 2014). Therefore, in Nacenta et al.'s terminology (Nacenta et al., 2009), Conductor's referential domain for selecting devices is "non-spatial". However, Hamilton and Wigdor's own experimentation and user study revealed that keeping track of multiple, often very similar devices is a surprisingly significant challenge and also that users extensively used spatial configurations of tablets for categorical organization (Hamilton and Wigdor, 2014). Therefore, it is possible that a non-spatial (or spatially-agnostic) interaction diminishes the benefits of cross-device interaction. It could create a mismatch between the spatial referential domain in which the users' intention is expressed and the non-spatial way in which the interaction technique requires the user to select a destination device (Nacenta et al., 2009). In other words, while menus are a robust and familiar way to select items that have no clear spatial relation, they might feel cumbersome for selecting one out of many devices from a spatial configuration.

7.1.3 Spatially-Aware Interactions

Alternatives are spatially-aware interaction techniques that use real-world spatial configurations as the referential domain, for example, hyperdragging or pick-and-drop of objects between laptops and table surfaces (Rekimoto and Saitoh, 1999) and flicking/throwing objects within AR settings (Voida et al., 2005), tabletops (Reetz et al., 2006), or from phones towards large displays (Dachselt and Buchholz, 2008; Dachselt and Buchholz, 2009). Throwing and flicking techniques are also frequently named by participants in gesture elicitation studies for multi-display interactions (e.g. (Kray et al., 2010; Seyed et al., 2012)) including our own elicitation study in this paper. This popularity of a spatial referential domain resonates with user studies that observed how important space and spatial configurations are as meaningful cognitive resources during knowledge work in offices (Kidd, 1994) or sensemaking on large screens (Andrews et al., 2010). However, in our elicitation study, participants also raised concerns about the accuracy of throwing and flicking and the danger of inadvertently sending content to wrong devices. This is particularly relevant in mobile cross-device interaction where the target screens are often too small and/or far away. More reliable spatially-aware techniques are world-in-miniature or radar views in which target regions or devices are selected in a top-down map-like representation of the environment (Biehl and Bailey, 2004; Reetz et al., 2006; Wigdor et al., 2006). User studies on tabletops revealed that these approaches can be more accurate, but also slower than flicking (Reetz et al., 2006).

7.2 Phase 1: Gesture Elicitation Study

Phase 1 of our study was aimed at eliciting user-defined cross-device gestures during four focus groups (see Figure 7.2, page 173). We explain why we decided for this rationale and why our methodology differs from traditional gesture elicitation.

7.2.1 Elicitation of User-defined Gestures

Nielsen et al. propose gesture elicitation studies for the design of intuitive and ergonomic gestural interfaces and to avoid arbitrary gesture sets that are rather designed for reliable recognition by technology than for easy learning and use by humans (Nielsen et al., 2004). Similarly, Nacenta et al. found that user-defined gestures are preferred by users and are also easier to remember (Nacenta et al., 2013). Wobbrock et al. successfully elicited multi-touch gestures for typical tasks from non-technical users (Wobbrock et al., 2009) and, since then, many similar studies have been used for connecting phones, public displays, and tabletops (Kray

Figure 7.2.: Gesture elicitation study showing participants in action and used props to delineate their suggestions.

et al., 2010), for diagram editing with multi-touch and pen (Heydekorn et al., 2011), for multi-display environments (Seyed et al., 2012), for active tokens querying big data (Valdes et al., 2014), or for skin input (Weigel et al., 2014). As we describe below, our work substantially differs from this work in two respects: First, we are not primarily interested in a single, ideally "optimal" gesture set but rather in a great breadth of user suggestions and deep insights into users' underlying thinking and metaphors. Second, we do not stop at eliciting gestures, but evaluate them in a controlled experiment to learn about their cognitive and ergonomic properties during repeated use. To our knowledge, only (Heydekorn et al., 2011; Nielsen et al., 2004; Voida et al., 2005) have pursued this approach but only for domains other than cross-device interaction.

To learn more about users' ideas, preferences, expectations, and mental models for cross-device gestures, we decided to prompt users with typical cross-device tasks and then elicit from them what their suggestions for the corresponding cross-device interactions are. For reducing bias and increasing their creativity, we primed them with a video showing latest cross-device techniques and asked them to be imaginative and to ignore any technological restrictions they know about. We also avoided commenting on the feasibility of their suggestions during their discussions.

The three main topics guiding this first phase of our study were as follows:

1. A great advantage of synchronous gestures such as SyncTap (Rekimoto, 2004), conduit (Chen et al., 2012), bumping (Hinckley, 2003), stitching (Hinckley et al., 2004), or pinching (Lucero et al., 2011) is that they only need built-in sensors. However, they must be learned and executed by the users across screens in the right sequence and with the right timing. How "intuitive" are such synchronous gestures and would users — not only designers — suggest them themselves?

2. Spatially-agnostic interactions such as device selection menus are familiar from GUIs and therefore are likely to be suggested by users (see "legacy bias" discussed in (Morris et al., 2014)). However, as discussed, a non-spatial selection is also potentially cumbersome (Nacenta et al., 2013). Would users be aware of this and what would their suggestions and opinions be?

3. Spatially-aware interactions such as throwing or flicking are popular suggestions in gesture elicitation, e.g. in (Kray et al., 2010; Seyed et al., 2012). They are fast and efficient (Reetz et al., 2006), because they have an open-loop control paradigm (Nacenta et al., 2009) and they make use of the users' natural understanding of space and physical movement. However, they are also less precise than other spatially-aware interactions, e.g., the top-down world-in-miniature or radar view representations in (Biehl and Bailey, 2004; Reetz et al., 2006; Wigdor et al., 2006) which offer closed-loop control (Nacenta et al., 2009) with higher accuracy but slower interaction (Reetz et al., 2006). We wanted to learn about users' preferences and opinions on this and if they would discuss speed and precision as important criteria.

7.2.2 Reducing Bias with Partners, Production, and Priming

To reduce bias, we did not introduce our three guiding topics or categories of interactions to the users. We only applied them afterward to categorize users' suggestions and to analyze their verbal comments, opinions, and their feedback from questionnaires. We also decided against a traditional gesture elicitation study like in (Wobbrock et al., 2009) and opted for an approach similar to Morris et al.'s proposal of reducing legacy bias with partners, priming, and production (Morris et al., 2014).

We used partners, i.e., focus groups of up to 5 partners in each session, to collect more and different suggestions, comments, and explanations. By enabling partners to fruitfully build upon one another's ideas and asking them to decide for a single preferred interaction, we hoped to create more reflection, discussion, and different opinions about the designs. Similar to (Marquardt et al., 2012a; Voida et al., 2005) our sessions, therefore, contained an element of co-creation instead of pure elicitation. As a result, we received many novel and elaborate suggestions including details for physical input and visual output.

We also employed production and priming (Morris et al., 2014). Production happened by requiring groups to produce at least three proposals for each task before deciding for their preferred one to move beyond few simple, legacy-inspired techniques. Priming happened by showing each group an introductory video with a variety of latest cross-device interactions to reduce the group's bias towards legacy-inspired GUI interaction. Also, users were encouraged to perform their suggestions with physical props such as tablets, pens, and paper to think more about the capabilities and affordances of mobile form factors instead of technological restrictions (see Figure 7.2, page 173). Nevertheless, we still followed a strict formal procedure during elicitation with carefully selected materials, questionnaires, and tasks.

7.2.3 Task Set

An initial set of tasks was extracted from cross-device systems or elicitation studies in literature (Chen et al., 2012; Hamilton and Wigdor, 2014; Kray et al., 2010; Seyed et al., 2012; Valdes et al., 2014). This initial set of 22 tasks was intended to represent the most typical and relevant cross-device tasks. In a pilot study, we then identified redundant tasks and those tasks that were too complex to understand for non-technical users. After removing these tasks the final set contained 19 tasks Table 7.1 (page 177), also see Figure 7.3 (page 176).

Tasks 1-9 represent typical cross-device object movements. Tasks 10-14 deal with stitching and duplicating screens. Tasks 15-19 are miscellaneous tasks such as pairing a wireless keyboard with a tablet or copying all files from all other devices to the personal device. In the table, "source" and "destination" define the involved devices and direction. Parentheses mean that a remote device, e.g. "(Tablet)," is not held in the hands of the users and is not lying directly in front of them. A distance of "in reach" means it is within an arm's length and "far" means users have to stand up and walk to reach it.

During the study, we prompted each group with one task after another. Each prompt was an animation that first showed the starting point and then the outcome of each task. To avoid bias, the animation did not show any user interactions. For example, it first showed two tablets lying on a table and that an object is on the screen of the

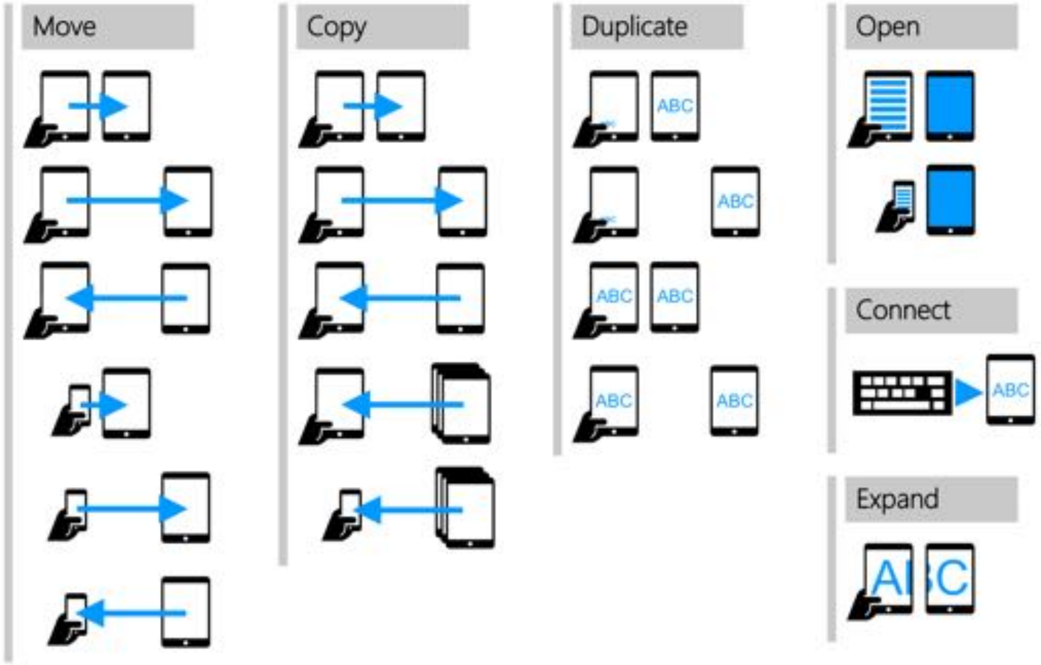


Figure 7.3.: 19 typical cross-device tasks extracted from cross-device systems or elicitation studies in literature.

first tablet. Then it showed that this object disappears and appears on the screen of the second tablet but without hinting at a possible user interaction.

7.2.4 Participants and Groups

We recruited 17 participants (7 female) through mailing lists and posters on a university campus. The age of participants ranged from 18 to 43 years ($\bar{x}$=26.4, SD=7.2). All participants had several years of experience with using a smartphone ($\bar{x}$=3.9, SD = 2.7). 13 participants also had experience with using a tablet ($\bar{x}$=2.2, SD=1.1). After reviewing the pool of participants, we manually assigned them to four groups to avoid too heterogeneous focus groups. Group A (N=4) were undergraduate students of computer science. Group B (N=4) were researchers in computer science or professional software developers. Group C (N=5) and Group D (N=4) were students from non-technical subjects such as comparative literature, anatomy, neuroscience, linguistics, architecture, and financial risk management.

7.2.5 Procedure

In each session, the group sat around a table with different switched off tablets, sheets of paper, and marker pens, in case users wanted to demonstrate or sketch their suggestions. After showing the introductory video, the recording of the elicitation

T#	Function	Object	Source	Destination	Distance
1	Move	File	Tablet	(Tablet)	In reach
2	Move	File	Tablet	(Tablet)	Far
3	Move	File	(Tablet)	Tablet	Far
4	Move	File	Phone	(Tablet)	In reach
5	Move	File	Phone	(Tablet)	Far
6	Move	File	(Tablet)	Phone	Far
7	Copy	File	Tablet	(Tablet)	In reach
8	Copy	File	Tablet	(Tablet)	Far
9	Copy	File	(Tablet)	Tablet	Far
10	Expand	View	Tablet	(Tablet)	In reach
11	Duplicate	Screen	Tablet	(Tablet)	In reach
12	Duplicate	Part of Screen	Tablet	(Tablet)	In reach
13	Duplicate	Screen	Tablet	(Tablet)	Far
14	Duplicate	Part of Screen	Tablet	(Tablet)	Far
15	Open	File	Tablet	(Tablet)	In reach
16	Open	File	Phone	(Tablet)	In reach
17	Connect	Keyboard	-	Tablet	In reach
18	Copy	All files	(All tablets)	Tablet	Far
19	Copy	All files	(All tablets)	Phone	Far

Table 7.1.: Set of 19 tasks used for the elicitation study.

began, and the following procedure was repeated for each of the 19 tasks. First, the animation for the task was shown on a projector or large screen. Second, the group was asked to think of interactions and discuss them with their group members. Groups were asked to produce at least three different interactions. Third, the group chose one favorite interaction for the given task and, after this, each group member filled out a questionnaire with 7-point Likert-scales about the understandability of the task, their personal agreement with the group's favorite, and how difficult it was to propose an interaction. After the 19 tasks, there was a debriefing and closing discussion. Participants were handed post-test questionnaires with 7-point Likert scales to ask whether they always understood what they were asked to do and if they felt that they could express their ideas during the session. Each session lasted between 1.5 to 2 hours, and participants were compensated for their time with £20.

7.2.6 Results & Discussion

The questionnaires revealed that participants had no problems with understanding what they needed to do ($\bar{x}$=6.76, SD=0.75) and expressing their ideas ($\bar{x}$=6.06, SD=1.25). The agreement and difficulty ratings after each task revealed high overall agreement ($\bar{x}$=5.99, SD=1.15) and low to neutral difficulty ($\bar{x}$=3.57, SD=1.67).

Not surprisingly, the tasks with the lowest agreement and highest difficulty were tasks 3, 9, 18, and 19 which all involved retrieving objects from one or multiple sources outside one's reach. However, there were no indications of more general problems with the study's design, its social setting, or the difficulty of the task set.

Synchronous, Spatially-Aware, and Spatially-Agnostic

After analysis and thematic coding of the video recordings, we took the favorite gestures from each group and all tasks (19 tasks × 4 groups = 76 favorites) and categorized them: 12 favorites (15.8%) were synchronous gestures (e.g. bumping devices), 10 favorites (13.2%) were spatially- agnostic interactions (e.g. select a target device from a menu of devices by name), and 54 favorites (71.1%) were spatially-aware interactions (e.g. flicking an item to a remote device).

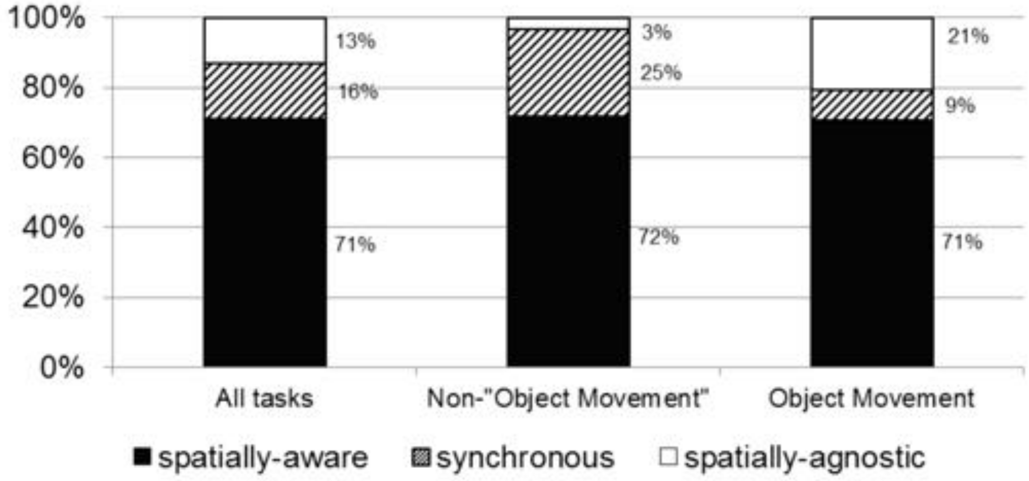


Figure 7.4.: Categorization of favorites.

We checked if the predominance of spatially-aware interactions was due to the many cross-display object movement tasks in the task set (tasks 1-9, 18, 19) by excluding them. For these "non-object movement" tasks 10–17, the spatially-aware interactions were again most popular at 71.9%, synchronous gestures were second at 25% and spatially-agnostic interactions third at 3.1% (Figure 2). This clearly indicates how strongly participants associated cross-device tasks with interactions in space and how much their thinking and suggestions were of a spatial nature. Although they were already familiar with some of the spatially-agnostic or synchronous cross-device techniques that are popular today, spatially-aware interactions still seemed to be most appealing, even for non-object movement tasks.

Synchronous gestures were most popular for task 10 "expanding view to other device" (4 of 4 groups) and task 17 "pairing keyboard and tablet" (2 of 4 groups). In general, participants seemed to use synchronous gestures mostly when they wanted to refer to the device itself or the entire screen, but not for cross-device object movement tasks. For such object movement tasks, only four favorites (9.1%) were synchronous gestures (mostly similar to the "conduit" gesture (Chen et al., 2012)) while 20.5% were spatially-agnostic (Figure 2). In summary, synchronous gestures seemed to be most important in the context of expanding screens or pairing of devices but were not popular for object transfer tasks.

We were surprised that spatially-agnostic interactions were almost as popular as synchronous gestures for all tasks and twice as popular for object movement tasks (20.5%). They were especially popular in Group C, which had a mixed, non-technical background. While this could be seen merely as a case of a strong legacy bias (Morris et al., 2014) it also hints at the ongoing importance and high practical relevance of more traditional menu-based interactions for cross-device interaction (e.g. demonstrated in (Hamilton and Wigdor, 2014)). This also resonates with the surprisingly good performance of menu-based techniques that we observed later in phase 2.

Suggestions for Spatially-Aware Interactions

We further analyzed the 54 favorites with spatially-aware interactions. 25 favorites (46.3%) were open-loop flicking/throwing gestures between devices. Participants discussed their potential limitations about precision and control. For example, Group D discussed the problem of inadvertently sending content to a wrong person or device in a room with 4 or 5 other tablets. Group C also discussed that imprecise flicking might result in content ending up on the wrong tablet, and Group A discussed the problem of how to flick content between two devices when there is a third device lying between them.

Groups B and D, therefore, suggested a slingshot metaphor instead of plain flicking/throwing. Inspired by games like Angry Birds, they suggested that direction and force/distance of flicking could be better controlled when an item is first pulled back from its current position with a finger on the touch screen and is then launched in the opposite direction after lifting the finger. They also considered additional visual output during aiming such as highlighting the prospective target device to have more control, thus turning the slingshot from an open-loop into an intermittent or closed-loop control paradigm (Nacenta et al., 2009).

Another suggestion was using visual proxies that represent remote devices on the local screen. This was suggested for 18 (33.3%) of the 54 spatially-aware favorites. For example, Group A suggested representing all remote devices as bubbles on the edges of the local screen. They appear where the imaginary line between the center of the local device and the center of the remote device intersects with the local screen's boundaries. These bubbles can be used as proxy targets for drag and drop or flicking to remote devices. Alternatively, Group C suggested an overhead map or radar view that contained live representations of all devices at their current locations as proxy targets (similar to (Biehl and Bailey, 2004; Reetz et al., 2006; Wigdor et al., 2006)).

Conclusions and Input for Phase 2

Given the great role that both spatially-aware and spatially-agnostic interactions played in participants' suggestions, we decided to explore such techniques further in phase 2 of our study. We decided for implementing the two spatially-aware techniques (edge bubbles and radar view) that users suggested to address the problem of insufficient control of open-loop flicking/throwing. Furthermore, given that spatially-agnostic interactions played a greater role than synchronous gestures in object movement, we also decided to implement a non-spatial menu-based technique to compare both approaches.

7.3 Phase 2: Interaction Techniques & Prototype

Software

Code #15

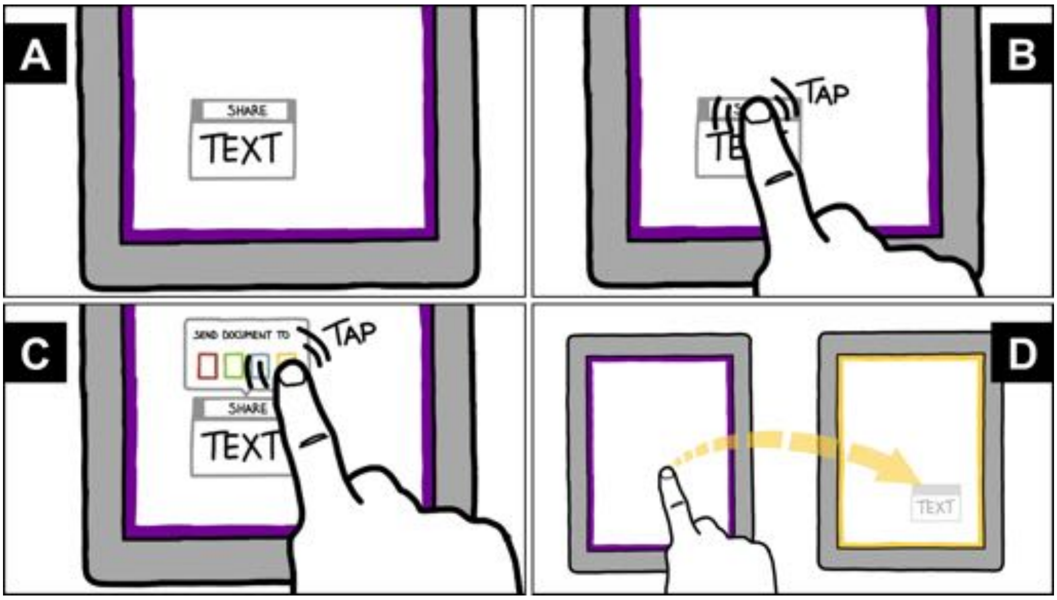
For phase 2 of our study, we integrated the edge bubbles, radar view, and menu techniques from phase 1 in a prototype application for multi-tablet sensemaking. The application prototype enabled users to use tablets to search a database with a few hundred text documents for keywords and to read the found documents (Scan QR Code #15 or enter "15" in the *MediaBrowser* application) . Users could highlight parts of the document in different colors, annotate documents, and copy relevant parts of a document into a summary document. For the experiment, we focused on three cross-device operations between a local tablet and a remote tablet.

1. Duplicating the current view of the document on the local tablet on a remote tablet.
2. Selecting a piece of text from the current document on the local tablet and copying it to a remote tablet.
3. Selecting an object on the local tablet and moving it to a remote tablet.

To enable a fair comparison, we ensured that they have equivalent functionality so that all three operations were possible with all three interaction techniques. Based on the observations in (Hamilton and Wigdor, 2014), we also assigned a unique color to each tablet. This color is always visible on the edges of the screen to facilitate identification and selection (see Figure 7.5 for example).

7.3.1 Interaction Technique 1: Menu

In our experiment, the spatially-agnostic menu technique represented the many suggestions of traditional GUI techniques from phase 1. They were particularly popular in Group C and are also used in recent publications (Hamilton and Wigdor, 2014). First, the object to move or copy must be identified on the local tablet, and its "Share" button must be pressed (see Figure 7.5A+B, page 181) (Scan QR Code #16 or enter “16” in the *MediaBrowser* application). This opens a context menu for selecting the destination tablet from a horizontal list of rectangles representing the remote tablets 1–4 by their color (see Figure 7.5C, page 181). Please note that they are ordered by an internal ID number and not by their spatial location since their locations are unknown to a spatially-agnostic technique. By tapping one of the rectangles, a remote tablet is selected as the target and the object is moved or copied there (see Figure 7.5D, page 181). Figure 7.5 shows the example of moving an object. For copying text or duplicating a view, the necessary interactions are almost the same. In these cases, a "Copy" button appears next to the currently selected text and for duplicating the view there is a "Share" button that is permanently shown in the bottom right corner of the screen (Scan QR Code #16 or enter “16” in the *MediaBrowser* application).



Video

Code #16

Figure 7.5.: Menu cross-device interaction.

7.3.2 Interaction Technique 2: Radar View

Following the suggestions from phase 1 (particularly from Group C for task 2), the radar view is a spatially-aware technique similar to (Biehl and Bailey, 2004; Reetz et al., 2006; Wigdor et al., 2006) that shows a top-down map instead of just a list. The map shows color-coded rectangles as visual proxies for all devices at their current real-world locations from an overhead perspective and is updated in real-time when devices are moved (see Figure 7.6C, page 182). To open the map and to select a destination device, the text or object to move or copy is dragged and dropped on the "Open Radar" button in the bottom right corner of the screen. Tapping on one of the colored rectangles in the map then selects the corresponding remote tablet as the destination device and closes the radar view. For duplicating the current view, users only press the "Open Radar" button and select a tablet without dragging text or an object to it (Scan QR Code #17 or enter “17” in the *MediaBrowser* application).

Video
Code #17

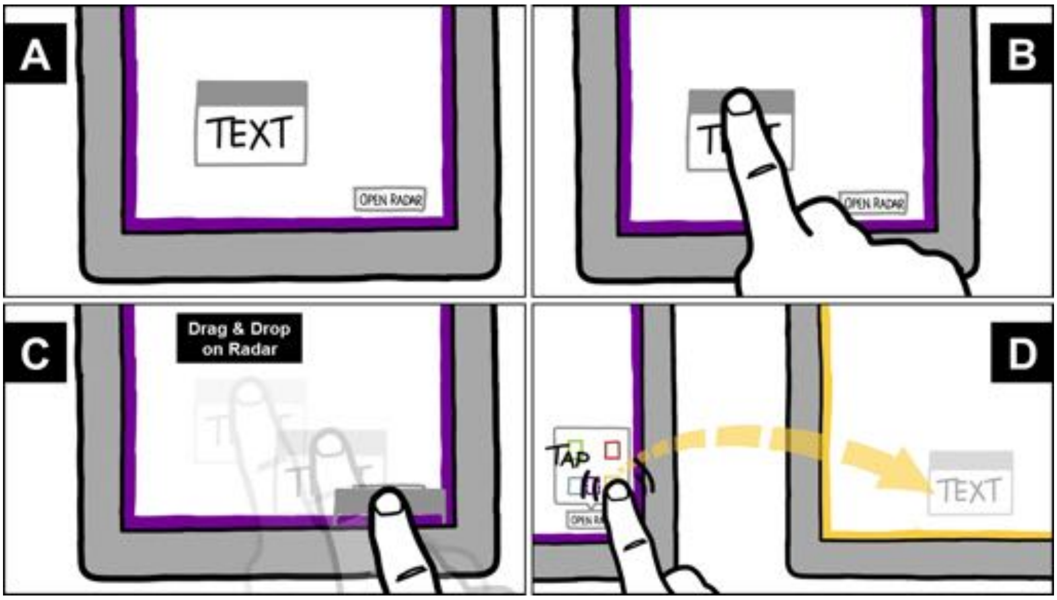


Figure 7.6.: Radar view cross-device interaction.

7.3.3 Interaction Technique 3: Edge Bubbles

Edge bubbles is another spatially-aware technique mainly based on suggestions made by Group A for tasks 1, 2, 5, 6, 8, and 9. Colored semi-circles around the edges of the screen serve as visual proxies for remote devices and, similar to off-screen-visualization techniques (Baudisch and Rosenholtz, 2003; Frisch and Dachselt, 2013), they indicate in which direction remote devices lie (see Figure 7.7A, page 183). The distance to a device is mapped to the radius of its bubble. The locations of the bubbles are defined by imaginary lines between the center point of the local device and the center points of the target devices in the real world. Each bubble is located where this imaginary line intersects with the edges of the local screen. The positions of the bubbles are updated in real-time and thus always reflect changes in the physical configuration of devices. Dragging and dropping an object on one of the edge bubbles moves the object to the corresponding target device. Tapping an edge bubble duplicates the current view of the local device on the remote device or copies selected text to it (Scan QR Code #18 or enter "18" in the *MediaBrowser* application).

Video

Code #18

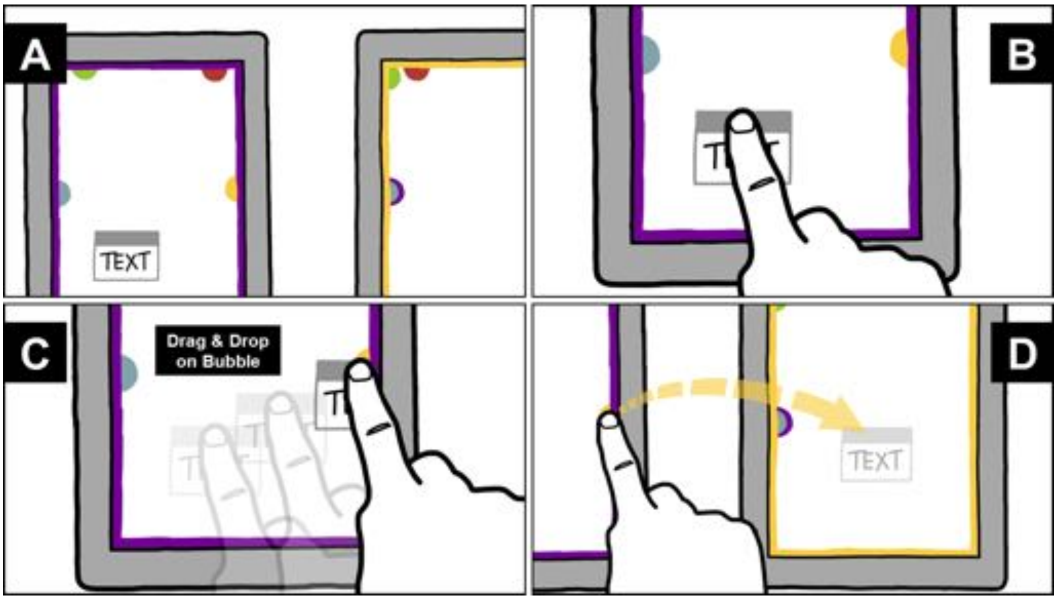


Figure 7.7.: Edge bubbles cross-device interaction.

7.3.4 Comparative Evaluation

Our comparative evaluation of menu, radar view, and edge bubbles was inspired by Nielsen et al., who suggest evaluating user-defined gestures in experiments to test them for their cognitive and ergonomic quality (Nielsen et al., 2004). Therefore, we tested the interaction techniques to (i) learn about ergonomic aspects, e.g., memory or stress and (ii) to better understand the benefit of spatially-aware visual proxies of devices compared to spatially-agnostic menus. To this end, we designed the study as a controlled laboratory experiment with a within-subjects design and an independent variable interaction technique with three within-subjects factors: menu, radar view, and edge bubbles. The order of the three tasks Duplicate View, Copy Text, and Move Object was kept constant, but the interaction techniques were systematically counterbalanced for each task using a balanced Latin Square. Dependent variables were the task completion time, and subjective measures (ranking of techniques by preference, how much users liked a technique, mental demand, effort, and frustration level). Additionally, a questionnaire with two open-ended questions asked for improvements of the interaction techniques and suggestions for other cross-device interaction techniques. To achieve a high degree of external validity, we chose realistic tasks during a typical sensemaking task and, as described above, the study prototype was a fully functional sensemaking application. To achieve higher internal validity, each task was repeated for 48 times per condition.

Participants

12 participants (7 female, 5 male) were recruited to take part in the experiment. The mean age was 24.3 years (SD=2.5, min=20 years, max=28 years). 11 participants were right-handed, and 1 participant was left-handed. None of the participants had color vision deficiency and thus no problems with the employed color coding. We only chose participants without a background in a computer science related field. 8 participants were students from non-technical subjects such as economics or law, 2 were research assistants in politics and physics, 1 was a kindergarten teacher, and 1 an occupational therapist.

Apparatus

Figure 7.8 (page 185) shows the physical setup of the experiment. As working surface we used a conventional office desk (1.2×0.8cm). Five Apple iPads (9.7″ diagonal) were provided as tablet devices in a U-shaped start configuration as illustrated in Figure 7.8 (page 185). To achieve a higher internal validity, we provided the tablet configuration in advance instead of providing participants a stack of devices and let them arrange devices as they like. Of course, this would allow us to study users' device configuration strategies; however it would not allow comparing interaction performance validly across participants and conditions. Nevertheless, investigating device configurations and users' cross-device interaction strategies is a likewise interesting research direction and could be explored in future studies.

Resonating with the task set (see Table 7.1, page 177), the remote tablets to the left and right of the local tablet could be reached comfortably within an arm's length and were therefore considered as "in reach". The other two tablets could only be reached by leaning forward and reaching out to them. The symmetric layout was chosen to account for different handedness of participants. A number was assigned to each tablet and placed next to it (see Figure 7.8, page 185).

To track the positions of tablets for the spatially-aware interaction techniques radar view and edge bubbles, we used the HuddleLamp vision tracking. In our setup, both techniques could also be simulated without any device tracking. However, to take

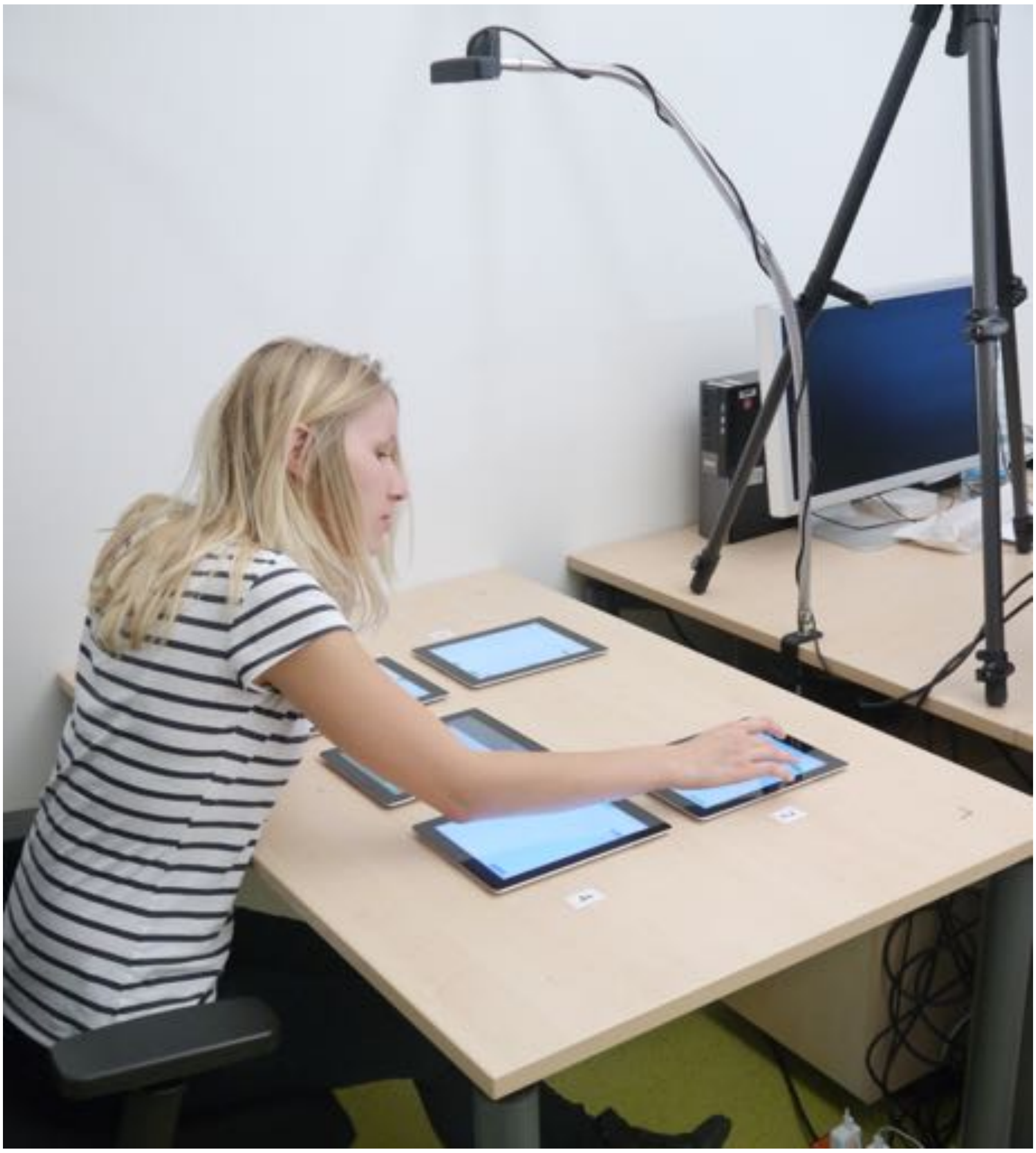


Figure 7.8.: Our experimental setup for multi-tablet cross-device interaction.

RQ2 one step further, we opted for HuddleLamp to evaluate whether this technology seamlessly integrates into people's everyday cross-device practices or it hampers them. For this reason, we used actual tracking data to update the radar view and edge bubbles in real-time to expose users to the limited accuracy, reliability, and noise in real-world tracking settings. For HuddleLamp, a Creative Senz3D RGB-D camera was set up at a height of 78cm, which provided a tracking region of 102cm×57cm.

The application running on each tablet was implemented in HTML5/JavaScript for Safari Mobile. Tablets were wirelessly connected to the tracking system to continuously receive location and orientation data for all tablets from a Web socket connection.

Task Design

The study consisted of the three tasks (i) Duplicate View, (ii) Copy Text, and (iii) Move Object and each task consisted of three conditions: menu (M), radar view (RV), and edge bubbles (EB). For each trial, users were prompted a number between 1 and 4 to indicate the destination device. The trial was noted as successful if the target device was correct; otherwise, an error was noted.

In each condition, participants repeated the cross-device interaction 48 times (each remote tablet was 12 times a target device). The sequence of numbers was randomized to avoid learning effects. Participants were asked to perform the cross-device interaction quickly and without errors. In total there were 12 participants × 3 tasks × 3 interaction techniques × 48 repetitions = 5184 trials with 432 trials per participant. Duplicate View and Copy Text always used the center tablet as the source device and the different tablets 1–4 as destination devices. It was also necessary to confirm the end of each trial by closing the duplicated view or deleting the copied text on the remote device. The Move Object task began with the center tablet as a source device and with one of the tablets 1–4 as the destination device. The destination tablet was then used as source device in the next trial and so forth.

Procedure

After signing a consent form and filling out a demographic questionnaire, the participants were introduced to the first task in their assigned first condition. We did not include a training phase due to the simplicity of the task and the many repetitions. After participants had completed the task for the first condition, they were introduced to the following condition until all three conditions of this task were completed. After the task, participants were asked to rank the three interaction techniques by bringing them into an order from most favored to least favored. They also rated each interaction technique in a questionnaire with four subscales: Liked (scale 0 to 100, 0: did not like, 100: liked it), Mental demand, Effort, and Frustration (all subscales from NASA TLX (Hart and Staveland, 1988), 0: low, 100: high). Two open-ended questions at the end of the questionnaire asked for the reason of the ranking and for possible improvements to any of the interaction techniques. This procedure was repeated for each of the three tasks. Each session lasted about 1.5 hours, and participants were compensated for their time with €12.

7.4 Overall Results and Discussion

For data analysis, Kendall's W coefficient (exact method) was used for the ranking of interaction techniques. The analysis of task completion time was done using ANOVAs with repeated measures, with post-hoc pairwise comparisons (all Greenhouse-Geisser corrected). The subjective ratings were analyzed with Friedman's ANOVA and Wilcoxon Signed Rank Test was used for post-hoc comparisons. All post-hoc tests were Bonferroni corrected. Figure 7.9 (page 187) show the sub-scales Liked, Mental demand, Effort, and Frustration for tasks 1–3.

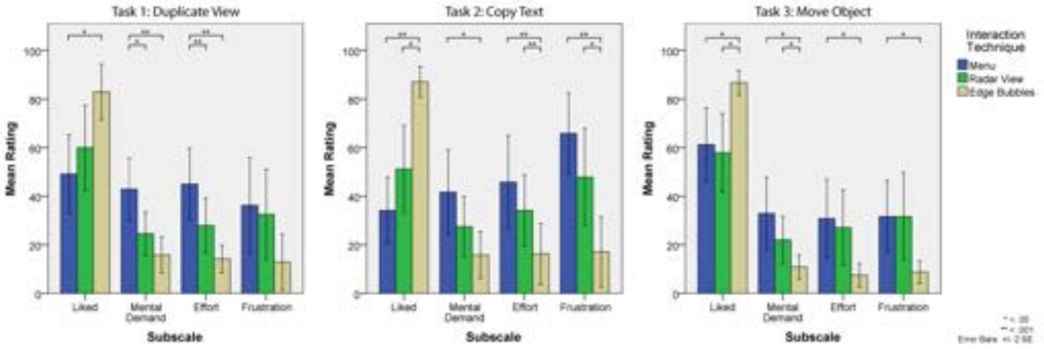


Figure 7.9.: Subjective ratings Liked, Mental demand, Effort, Frustration for each task.

The ranking for tasks T_{1-3} was significant with a Kendall's W_{T1} = .361 ($\chi^2(2)$ = 8.67, $p < .05$), W_{T2} = .72 ($\chi^2(2)$ = 17.17, $p < .001$), and W_{T3} = .65 ($\chi^2(2)$ = 15.50, $p < .001$). For all tasks, cross-device interactions were ranked in the following order: M_{EB} (T_1: 1.3, T_2: 1.1, T_3: 1.1), M_{RV} (T_1: 2.2, T_2: 2.2, T_3: 2.3), and M_M (T_1: 2.5, T_2: 2.8, T_3: 2.6) (values from 1 most favored to 3 least favored). The statistical analysis revealed that the order of the mean rankings from most favored to least favored for each task was consistently edge bubbles, radar view, and menu. The spatially-aware techniques proved to be favored by users even after many repetitions and the popularity of spatially-aware techniques during the elicitation study in phase 1 also clearly showed in the results of phase 2. However, as we show in the following section, it is not possible to generalize this to all spatially-aware techniques.

7.4.1 Spatially-aware interaction is not always better

For all tasks T_{1-3}, an ANOVA revealed a statistically significant difference between interaction techniques for task completion time (T_1: $F_{1.88,20.69} = 22.69$, $p < .001$, $partial\ \eta^2 = .67$, T_2: $F_{1.45,15.98} = 56.04$, $p < .001$, $partial\ \eta^2 = .84$, T_3: $F_{1.82,20.04}$ $= 24.95$, $p < .001$, $partial\ \eta^2 = .69$). All Friedman's Tests revealed statistically significant differences between interaction techniques for subjective ratings Liked, Mental demand, Effort, and Frustration for all tasks 1–3 (see Figure 7.9, page 187).

The spatially-aware edge bubbles outperformed menu in terms of task times for tasks 1 and 2 and consistently scored higher than the non-spatial menu on the Liked subscale for tasks 1–3. However, surprisingly the also spatially-aware radar view was outperformed by the menu in terms of task time in task 3 and never scored significantly higher than the menu in the Liked subscale. The differences between the two spatially-aware techniques are also visible in the higher Effort and Frustration in tasks 2 and the higher Mental Demand for task 3 for radar view than for edge bubbles. Moreover, for all tasks, the task times for radar view are significantly higher than for edge bubbles. It seems that, despite its popularity among users, spatial awareness alone does not lead to greater user performance and better usability.

A potential explanation why edge bubbles is clearly superior to radar view in our experiment seems to be the higher mental demand (see the significant difference in mental demand between edge bubbles and radar view) of mapping the virtual proxy objects on the screen to their real-world counterparts. To use the radar view, users eventually must locate the destination tablet on the map instead of directly swiping objects towards it in egocentric fashion like for edge bubbles. As, one participant commented that edge bubbles "[. . .] is very intuitive because proximity and direction are clear and natural [. . .]"[1]. The identification of the destination tablet in the radar view condition, however, requires mentally switching from the natural egocentric view of the environment to an exocentric top-down view. One participant stated that radar view "requires more effort" and is "is not intuitive". The difference in the mental demand is also reflected other participants' comments. For example, one participant commented that edge bubbles "worked well"[2] and "relationships [were] clearly recognizable"[3] and also radar view "worked well" but was a "little more exhausting".

This higher mental demand also seems to diminish the benefits of the radar view compared to the menu. This becomes visible in the absence of significant differences between them for Frustration in all tasks. Also, while the radar view helps to identify devices faster when the spatial configuration is unknown or very dynamic, the performance of menu improves after users have internalized the mapping of colors to tablets over time. Using the menu then only requires a sequential scanning of a one-dimensional list of colored objects. This explains why the menu is faster than

---

[1] Translated from German Edge Bubbles "ist sehr intuitiv da räumliche Nähe und Richtung eindeutig und natürlich"

[2] Translated from German "hat gut funktioniert"

[3] Translated from German "Zusammenhänge klar erkennbar"

radar view in the last task 3 and that there are no differences anymore in Mental Demand and Effort for tasks 2 and 3.

Of course and as mentioned earlier, the mentally demanding switch from an egocentric view to a top down view in the radar view is just one possible explanation for coming off poorly. However, further studies are needed to investigate if the higher mental demand inevitably relegates to a mental switch.

7.5 Summary

In the following, we summarize our results and discussions from phase 1 and 2 in four findings that can inform the design of future cross-device interaction.

First, Phase 1 has clearly revealed that users expect cross-device interactions to be spatially-aware (71.1% of all suggestions). In phase 2 the spatially-aware edge bubbles technique outperformed other techniques in a controlled experiment and was the most favored technique, even after many repetitions.

Second, as shown in the experiment, spatially-aware techniques have to be designed with care. The edge bubbles technique succeeded because of its directness of interaction and its spatial representation that did **not** require mentally switching between an egocentric and a top-down view. Top-down views such as radar views seem to introduce a cognitive load that can entirely diminish their advantages over simple spatially-agnostic menus.

Third, spatially-agnostic interactions such as menus were popular for cross-device object movements (20.5% of suggestions), particularly for tasks involving one or multiple remote devices as sources. In comparison to maps or radar views, they can show a good or even equivalent performance when the number of devices is small, and their spatial configurations are not changing rapidly.

Fourth, synchronous gestures were popular (25% of suggestions) for tasks concerned with expanding views or pairing devices or whenever users wanted to refer to the device itself or its entire screen. They seem to naturally fit pairing tasks but were only suggested in very few cases for cross-device object movement (9.1%).

8

Conclusion

> *If you can't explain it simply, you don't understand it well enough.*
>
> — **Albert Einstein**
> (Physicist & Nobel prize winner)

Contents

This chapter concludes with a brief thesis summary. It then integrates findings and limitations from presented empirical and experimental research in a two-staged process. First, it integrates findings and limitations individually for each of the two research objectives: Spatial Navigation (RO1) and Cross-Device Interaction (RO2). In a second and final step, it integrates findings of both research objectives together and summarizes their overall conclusions. It eventually abstracts results to design guidelines (DG), in particular, to guide the design of fun and joyful UbiComp experiences. These guidelines are used to conceptually re-design Blended Shelf and Integrative Workplace from Chapter 3 (page 57). In an outlook, it finally proposes future research leaning towards ecologically valid studies that are aligned with the thinking of embodied interaction.

8.1 Thesis Summary

The goal of this thesis is to shed light into current UbiComp experiences. It is ought to find hidden potentials of UbiComp by the guidance of beliefs on embodied cognition and recent theories on human spatial memory. It seeks for possibilities that exploit users' pre-existing knowledge for human-computer interaction. This knowledge is often deeply embodied in our everyday activities, which we subconsciously apply when sufficiently practiced. Ideally, such subconscious actions provide — yet unexplored — potentials to build fun and joyful UbiComp experiences. Therefore, two fundamental and high-level challenges were tackled within this thesis:

- Enable users to exploit pre-existing knowledge to navigate and interact in virtual information spaces.
- Seek for opportunities to utilize commodity and off-the-shelf hardware to enable users to work across multiple mobile devices.

As a starting point, Chapter 2 establishes a theoretical background. It briefly motivates Weiser's seminal vision of a world of ubiquitous computing and discusses opinions on its success and failure with potential reasons for shortcomings. It depicts often contradicting views of researchers, in particular, opinions of HCI researchers. Together with the embodied interaction theory (Dourish, 1999), the Reality-Based Interaction framework (Jacob et al., 2008) and the Blended Interaction framework (Jetter et al., 2014), the thorough review of these opinions helped to identify potentials for new UbiComp experiences by moving computing technology out of users' center of their conscious attention (Weiser and Brown, 1996). Ideally, this allows users to access computing technology when needed rather than letting technology dictate them how and when to use it (Rogers, 2006).

These potentials for new UbiComp experiences are further strengthened in Chapter 3 (page 57). This Chapter 3 reveals issues of knowledge work activities in academic libraries. Data gathered in field studies and its subsequent analysis, empirically validated the need for physical shelf-browsing and fluid configuration of work artifacts.

Consecutive Chapters 4-7 operationalize these potentials and research their benefits in controlled experiments. They further seek for an understanding of the importance of space as a cognitive resource by observing various spatial and cross-device interactions and their impact on users' performance (e.g., navigation and object recall) and subjective workloads (e.g., physical and mental demand). Their findings will be integrated in a two-staged process in the following sections. First, it integrates findings and limitations separately for each of the two research objectives: Spatial Navigation (RO1) and Cross-Device Interaction (RO2). Second, it integrates findings of both research objectives together and summarizes their overall findings.

8.2 Integration of Findings

Integration of findings begins with Chapters 4 and 5. They study egocentric spatial navigation with a single device and potential implications for the design of dynamic peepholes. In particular dynamic peepholes to navigate in virtual information spaces such as digital library collections. Then, it integrates findings from Chapters 6 and 7 emphasizing on interaction with information spaces that are scattered across multiple devices. It eventually gives implications for the design of cross-device interactions.

8.2.1 Findings from Spatial Navigation

Spatial navigation in virtual information spaces such as dynamic peephole navigation offers familiar physical navigation. However, the size of the screen for peephole navigation also comes at a different cost.

As shown in *Chapter 4 – Spatial Navigation – Peephole Size & Navigation Behavior*, a larger peephole screen size reduces the need for slow physical panning and search in favor of a faster visual scanning of the display's content. It further allows for recognition rather than recall from spatial memory because it reveals more visual features that support user orientation all at once. In real-world systems, however, larger peepholes or large displays increase cost, energy consumption, and weight, and the devices become more cumbersome to carry and handle.

Alternatives are small and lightweight handheld projectors which can produce a relatively large peephole. However, some practical problems come to play. For instance, hand jittering produces a shaking display making it hard to read projected content. Often it is difficult to find surfaces in the right size and lighting conditions for projection. Last, projecting content in public spaces inhibit privacy, which resonates with privacy concerns raised by participants of the Blended Shelf study (see Chapter 3.3.2, page 79).

Designers must make concessions due to these constraints. They want users to experience the benefits of larger peepholes while avoiding the many disadvantages that result from using and handling larger devices or mobile projections. Therefore answering the question of how small peepholes can become without overburdening their users during map navigation is of high practical relevance.

In *Chapter 4 – Spatial Navigation – Peephole Size & Navigation Behavior*, we found that a tablet-sized peephole is already a "sweet spot" between peephole size and both user navigation performance and user task load. A smartphone-sized peephole is too small and outperformed by all larger sizes. Unsurprisingly, our research revealed that larger peepholes significantly improve learning speed, navigation speed, and reduce task load. What is surprising though and contradicting to some Fitts' Law peephole target acquisition models, this added benefit diminishes with growing sizes, and peephole sizes greater than a tablet screen do not pay off regarding navigation performance or task load anymore.

Our additional study in *Chapter 5 – Spatial Navigation – Spatial Memory* uses the "sweet-spot" tablet-sized screen to compare traditional multi-touch navigation to dynamic peephole navigation. It revealed a significantly better navigation performance for peephole navigation in terms of path length (47%) and task time (34%). Importantly, participants reported a better user experience and a significantly lower mental demand and frustration. They often used their physical position as a spatial cue to maintain their global orientation in the virtual information space.

Limitations

Although our studies on spatial navigation are the first of their kind, they have some limitations that we are aware of, and that could be addressed in future research efforts.

With the study presented in *Chapter 4 – Spatial Navigation – Peephole Size & Navigation Behavior*, we deliberately controlled device-specific properties such as weight or resolution. Only using peephole size as independent variable increased the internal validity but also decreased the external validity. It would be interesting to repeat the experiment using real-world physical devices to see if the same results can be replicated or if the differences in device-specific properties such as resolution, weight, or latency outweigh differences in peephole size.

In *Chapter 5 – Spatial Navigation – Spatial Memory* all participants chose a horizontal orientation of the tablet device in the multi-touch condition (see Figure 5.3, page 123). Participants put the tablet on the mobile desk in a horizontal position in front of them rather than aligning it vertically with the large vertical screen. This probably could have some effect on their mental demand because of the mental adjustments of the different orientations. However, this might also have implications on users’ subjective workload as eventually physical demand and effort increase due to unnatural position of the tablet. This assumption is guided by Müller et al. who investigated the effect of orientation of a peephole device on users’ subjective workload. They find that holding a peephole in a horizontal orientation has a significantly lower physical demand but also higher mental demand. They hypothesize “that the increased mental demand results from an additional degree of freedom (yaw rotation)” (Müller et al., 2015). However, Müller et al. did not have an extra large screen and therefore the characteristics of tablet-to-large-display alignment and its impact on users’ performance needs to be explored in a separate study.

8.2.2 Findings from Cross-Device Interaction

Interaction across one’s personal device ecosystem or moving virtual objects from one’s device to another person’s device is an often mentioned problem in HCI literature (Dearman and Pierce, 2008; Santosa and Wigdor, 2013; Cecchinato et al., 2016). However, technology for such cross-device interaction is either low-cost but does not allow for spatial tracking of devices (Hamilton and Wigdor, 2014; Chi and Li, 2015) or requires instrumentation of devices (Schmitz et al., 2010) or rooms with expensive tracking systems (Marquardt et al., 2012a). Moreover, most of the cross-device interaction systems are research prototypes and closed source and thus not available to conduct ecologically valid user studies. Align with the often unavailability of cheap or low-cost tracking systems are missing guiding principles for cross-device interactions (Oulasvirta, 2008).

In *Chapter 6 – Cross-Device Interaction – Enabling Technology*, we described HuddleLamp, a sensing technology in the form of a desk lamp with an integrated RGB-D camera that tracks the movements of multiple mobile displays on a table. It allows for various kinds of scenarios. For example, single user scenarios such as interaction across multiple documents when reading and writing across them. It also allows for around-the-table collaboration among several users and tablets. Our approach of hybrid sensing implemented in HuddleLamp accounts for both scenarios.

HuddleLamp is the first mobile device tracking system that comes with a reasonable effort of environment instrumentation and works with off-the-shelf hardware. Only, a single low-cost RGB-D camera needs to be mounted above a table, i.e., in a desk lamp. Its tracking system detects multiple mobile screens by exploiting their optical properties in the IR range. It additionally uses RGB images and virtually displayed fiducial markers to track screens better and to distinguish them from other objects or users’ hands.

After a technical evaluation of our hybrid sensing approach, we also introduced HuddleLamp's web-based architecture and JavaScript API for ad-hoc collaboration that enables users to add or remove displays and reconfigure them in space at any time without installing any software. We discussed how this could be used to create large multi-device tiled displays for multi-user and multi-touch interaction. Beyond our five demonstrated examples of different interaction techniques, we believe that HuddleLamp's setup and hybrid sensing tracking will allow the rapid exploration of future cross-device interactions supporting group collaborations. Therefore, we made HuddleLamp software freely available and offer it as open source project to the public and for the research community[1].

To understand the potentials of HuddleLamp and how its spatially-aware multi-device tracking spawns new UbiComp experiences, we explored the practicality of spatial cues for cross-device sensemaking. In particular to find solutions and answers to problems raised in *Chapter 3.2.4 – Fluid Configuration of Work Artifacts*.

For example, *Chapter 7 – Cross-Device Interaction – Understanding Spatial Cues* presents the results of a two-phased study exploring the design space of mobile cross-device interactions. In this Chapter, we first describe our results from a gesture elicitation study in which 71% of the elicited cross-device interactions were spatially-aware. We discuss how participants strongly associated cross-device tasks with interacting and thinking in space (resonating with (Kirsh, 2010)). Based on the users' suggestions, we implemented two spatially-aware interaction techniques and one spatially-agnostic technique, comparing them in a controlled experiment. The results show that spatially-aware techniques, when designed with care, are preferred by users and can decrease their mental demand, effort, and frustration during mobile cross-device interactions.

Limitations

During the exploration of the design space of HuddleLamp and the following user study on cross-device gestures, we also identified limitations.

There is a small but noticeable delay between the physical movement of a screen and the corresponding reaction of the UI. While the vision processing works at the maximum frame rate of the camera (30 fps), the web socket connection, the synchronization of devices with Meteor, and particularly the different rendering performances of browsers and devices induce a noticeable latency. Best results were achieved using Safari Mobile on Apple iPad 2/3/Air, Apple iPhone 4/5S, and

---

[1] HuddleLamp project website – `http://www.huddlelamp.org` (last accessed: May 29th, 2016)

Microsoft Surface 2 Pro with Chrome. Less successful were tests with mobile devices of older generations (e.g. Apple iPad 1, Apple iPhone 3G) and surprisingly with Google Nexus 5 & 10. However, this limitation will naturally fade away with next generations of wireless communication protocols (e.g., Li-Fi[2]), mobile device hardware with powerful onboard graphic chips, and browsers with better support for GPU rendering.

The recognition of multi-touch gestures such as pinch-to-zoom is limited when they span device boundaries. Currently, gestures are only recognized if all fingers are on a single device. They cannot be performed by putting one finger on one device and other fingers on a different device. Nonetheless, for each device, they are correctly recognized and processed, and their results are immediately synchronized with all other devices.

An apparent limitation of HuddleLamp compared to traditional tabletops is the absence of a large screen that also displays touch-enabled interactive content around the mobile screens on a table. Does this inhibit its usefulness in real-world scenarios? This disadvantage has to be weighed against HuddleLamp's advantages regarding low-cost portability, ad-hoc interaction, and distributed rendering. Furthermore, it supports high-resolution output and responsive input on a capacitive multi-touch screen, sometimes with a pressure-sensitive stylus (e.g. on a Microsoft Surface tablet). This is often not the case for large tabletops with optical touch detection.

A limitation of our cross-device gesture study is its focus on two-dimensional device configurations on a desk. Results are not generalizable to other spatial configurations such as handheld devices or see-through tangible lenses (Spindler et al., 2009). More future work is needed to study cross-device techniques for such more complex 3D device configurations.

8.2.3 Overall Integration of Findings

Findings concerning both research objectives can be finally integrated into overall results of research. They contribute to both, the deductive and inductive approach applied in this thesis. In *Chapter 2 – Theoretical Background*, we presented recent emerging theories, models, and frameworks in HCI that particularly reason for a mutual interplay between body and mind. This unfolds new opportunities for human-computer interaction beyond the antiquary use of keyboard and mouse.

---

[2] Light Fidelity (Li-Fi) is a wireless communication technology that exploits light for data transmission; thereby achieving "lightning" fast communication rates.

With this research we contribute to the manifestation of embodied cognition theories as action and thinking in space improves users' performance in various ways. For example, as shown in research on spatial navigation, whole body movements improve users' navigation performance over traditional multi-touch interaction when navigating in both familiar and unknown information spaces. Or when working across multiple tablets and moving virtual objects between them. Spatial interaction and spatially-aware interaction techniques are preferred by most users and considered as joyful and close to reality. At the same time, spatial interactions decrease their mental demand and frustration. Resonating with this, our cross-device object movement study clearly revealed that users think and act in space (Kirsh, 2010; Kirsh, 2011). But the design of a spatial and cross-device interaction technique plays an important role on whether users experience it as "natural" (Jetter et al., 2014) or cumbersome.

8.3 Design Guidelines

Finally, all findings can be transformed into two design guidelines **(DG)** for future spatial and cross-device applications. Eventually, these guidelines are applied by researchers and practitioners to develop UbiComp experiences that increase users' task performance, lower their individual workloads such as mental demand, effort, and frustration. At the same time, these guidelines can lead to an increased cumulative value of mobile devices that are already around us. These two design guidelines are applied in a re-design of Blended Shelf and Integrative Workplace to close with the open issues raised in Chapter 3 (page 57).

DG1: Consider tablet-sized peephole interaction to navigate in large virtual information spaces when privacy, navigation performance, and cognitive demand are important requirements.

Due to their large screen real-estate, wall-sized displays have the advantage of viewing an entire information space, or at least significant portions of it, all at once. Users thereby can step back and overview display's content to recognize and access objects of interest. They do not rely on recalling object locations from memory. Large displays, therefore, support recognition rather than recall[3]. However, as revealed by the Blended Shelf user study in Chapter 3 (page 57), a large display also comes with the cost of losing privacy. Inherently, a smaller screen better supports privacy during exploration and navigation of an information space when compared to a wall-sized display.

---

[3] "Minimize the user's memory load by making objects, actions, and options visible." (Nielsen Norman Group) – `https://www.nngroup.com/articles/ten-usability-heuristics/` (last accessed: May 29th, 2016)

However, a smaller screen also comes with a cost. Users might have to manually scan the off-screen content using view management techniques such as multi-touch navigation or peephole navigation. Despite this seemingly obvious drawback, in Chapter 4 (page 95) we found that a relatively small tablet-sized peephole display leads to a similar task performance for map navigation when compared to a wall-sized display, and once users familiarized themselves with the information space. Thereby, the benefit of a large and often costly display diminishes over time.

Alternative to peephole navigation, multi-touch navigation allows users to navigate in virtual information spaces using traditional drag-to-pan and pinch-to-zoom touch gestures. However, as shown in Chapter 5 (page 119), with this navigation technique users often lose global orientation in the large information space. Whereas with egocentric peephole navigation they can concur from their physical position to the virtual position. This spatial cue enables them to maintain their global orientation in the virtual information space. Also, they navigate in the information space more efficiently with an egocentric peephole navigation than with multi-touch navigation.

In consequence, peephole interaction is an alternative to large wall-sized screens when privacy is a requirement. It also is superior to traditional multi-touch navigation and leads to better navigation performance and reduces users' cognitive demand. As hinted with our initial experiment E2 in Chapter 5 (page 119), peephole interaction potentially can better exploit long-term spatial memory when compared to traditional multi-touch interaction.

Therefore, we recommend using tablet-sized peephole interaction to navigate in large virtual information spaces when privacy, navigation performance and cognitive demand are important requirements. Especially, in those contexts that require the user to invest considerable cognitive resources for application-level tasks (Andrews et al., 2010). For example, during time-critical decision-making, sensemaking processes, or more generally most higher-order learning tasks (Ohlsson, 1995). However, this must be balanced with the greater physical demand which prohibits the use of peephole navigation when navigation operations are executed very frequently over a longer period and thus result in physical strain and fatigue.

DG2: Design for fun and joyful cross-device experiences through spatially-aware cross-device interactions.

The need for parallel and sequential use of multiple mobile devices has been discovered by HCI researchers (Dearman and Pierce, 2008; Jokela et al., 2015a; Jokela et al., 2015b; Cecchinato et al., 2016) and industry[4] alike. As shown in Chapter 3 (page 57) with the Integrative Workplace user study, this is also very similar for knowledge workers who fluidly arrange work artifacts on their workspaces to compare documents or for cross-referencing. Despite the reported daily sequential and parallel use of mobile devices, HCI still provides little support for cross-device interactions. Even worse, the appropriate design of cross-device interactions and whether they **need** to be spatially-aware is still an open question.

In Chapter 7 (page 169), we took an initial stab and investigated in this question. We found that users expect cross-device interactions to be spatially-aware (71.1% of all suggestions) and when implemented with great care are the most favored cross-device techniques. Beyond user preference, spatially-aware techniques such as the presented edge bubbles technique can even outperform non-spatial or spatially-agnostic cross-device techniques regarding users' subjective workload such as mental demand, effort, and frustration.

However, it is important to mention that the "design" of a spatially-aware cross-device interaction technique plays a significant role in (i) users' preference and (ii) users' subjective workloads. This was clearly shown in Chapter 7 (page 169) with the significant differences between both spatially-aware techniques: edge bubbles and radar view. The edge bubbles technique was superior to the radar view technique because of its egocentric and more direct manipulation (Hutchins et al., 1985). Its spatial representation did **not** require users to switch mentally between an egocentric and a top-down view. Top-down views such as radar views seem to introduce a cognitive load that can entirely diminish their advantages over simple spatially-agnostic menus.

In consequence, we recommend egocentric spatially-aware cross-device interactions. They are preferred by users, considered as fun and joyful cross-device experiences, and can reduce users' cognitive load. Enabling technologies such as HuddleLamp, which have been considered as futuristic and far beyond current sensing technologies just two years ago (Hamilton and Wigdor, 2014), can now provide the necessary sensing to enable such fun and joyful cross-device experiences at low cost and with little instrumentation effort. Ideally, such technologies are available in libraries and offices, e.g., as replacement of a light bulb in desk lamps. However, as shown in our experiment in Chapter 7 (page 169), spatially-aware techniques have to be designed with care. Otherwise, non-spatial techniques can perform equally well or better.

---

[4] The New Multi-screen World: Understanding Cross-platform Consumer Behavior – `https://think.withgoogle.com/databoard/media/pdfs/the-new-multi-screen-world-study_research-studies.pdf` (last accessed: February 5th, 2016)

At the time of writing this thesis, this design recommendation has already been applied by Wozniak et al. in their RAMPARTS system (Wozniak et al., 2016) confirming that users think and act in space. Participants in their study explicitly mentioned that the ability to arrange artifacts in space was good and makes the task easier. This manifests our recommendation for spatially-aware cross-device interactions also for sensemaking tasks.

8.4 Transfer Findings to Envision Future Knowledge Work

In a final step within this thesis, I will use the two design guidelines and exemplify the utility of them by doing a re-design of Blended Shelf and Integrative Workplace on a conceptual level. This re-design closes with the open issues postulated in Chapter 3 (page 57) by illustrating it in the following knowledge work scenario (see Figures 8.1-8.5).

Knowledge workers often sit at their workplace to read and write documents or cross-reference between them. With the proliferation of lightweight tablets and ebook readers, for most people they became a de facto replacement for books. Modern devices have a large capacity and therefore can store a vast amount of books on a single appliance. Moreover, they provide their users with digital functions such as search, highlighting tools, and bookmarking. In the re-design, I consider tablet devices and smartphones (tablets' smaller companions) as modular building blocks of a future system. For two reasons, they can be (i) physically arranged and stacked in space and (ii) fluidly assembled to larger compositions when needed. HuddleLamp detects such physical device compositions by analyzing the distance and orientation of mobile devices to each other.

In Figure 8.1A (page 202), a user composes a word processor based on two tablets. The right tablet allows for handwritten pen input while the second tablet to the left immediately transforms ink strokes into a machine readable text. Thereby, the user can compare handwriting with system's transformation results and correct when necessary without the need to manually switch between both. Moreover, views on both tablets are coordinated. For example, highlighting a word on one tablet also emphasizes the word on the other tablet or scrolling on one tablet also scrolls content on the other tablet to the corresponding position.

While the two tablets centered in the user's workspace, other tablets are arranged in close proximity. They hold relevant information for the task at hand and allow the user to think and act in space. For example, the tablet on the far right of the

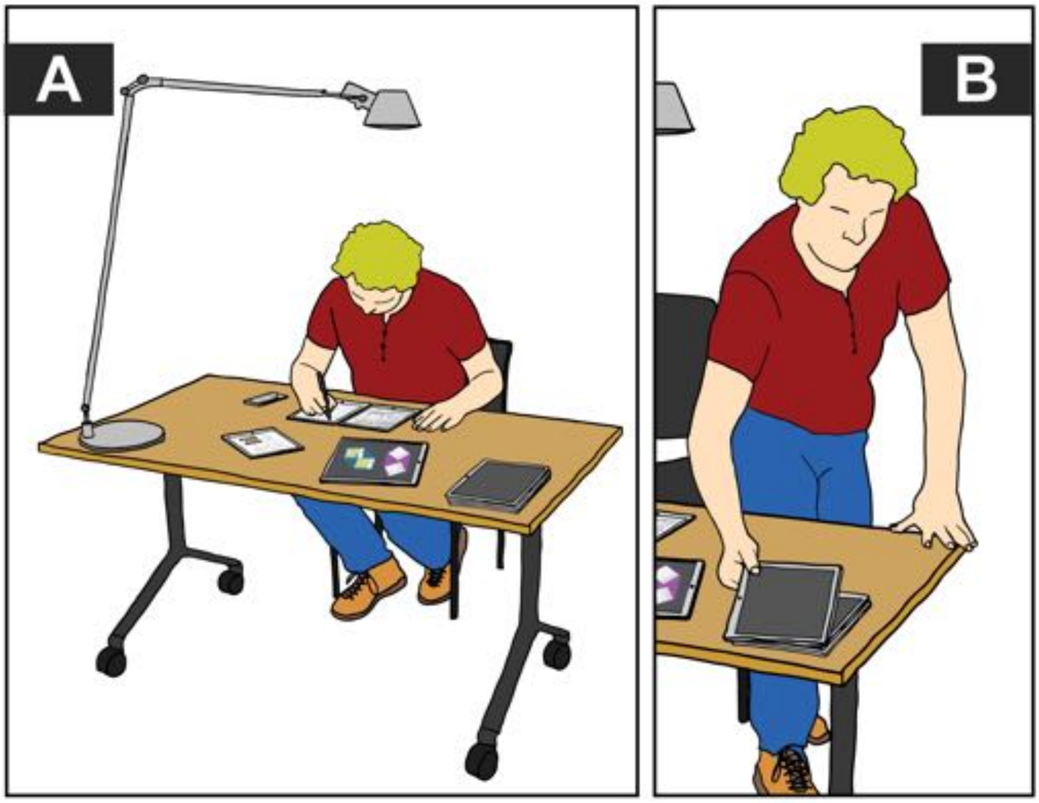


Figure 8.1.: A knowledge workplace with multiple mobile devices. (A) The devices are arranged in space to spatially reflect a user's current thoughts or to fluidly configure the workspace according to a user's current needs. (B) A user grabs a tablet from the stack of idling devices.

user visualizes book content from which the user might want to quote. The larger tablet on the top left of the user acts as information container and holds notes or excerpted information that might be relevant in future. Information on this tablet could also result from a collaborative co-located Web search (Rädle et al., 2013b) or sensemaking (Zagermann et al., 2016) activity conducted earlier with co-workers.

As one essential part of the knowledge work process, knowledge workers need to acquire new information. Our study in Chapter 3 (page 57) revealed that library users still consider bookshelves as a valuable tool for bibliographic search because of its inherent quality of shelf browsing. The re-design, therefore, also provides a reality-based exploration of digital library collections using peephole interaction. To access the virtual information space, the user grabs a so far idling tablet from the device stack (far left end of the table) (see Figure 8.1B, page 202). The user transitions from *reading and writing* to *bibliographic search*. Thereby, supporting an entire workflow as suggested by Blended Interaction (Jetter et al., 2014)). As a by-product of this transition, the workspace on the table remains unchanged and will be in the same state when the user returns after a search. Also, as an additional benefit, tedious "ALT-Tab" switching between WIMP applications will be obsolete.

On the just grabbed tablet, the user initiates a keyword search. As a result of this search, a digital bookshelf is created similar to Blended Shelf and with a default media arrangement (e.g., classification of a user's preferred library) (see Figure 8.2C, page 203). Of course, media on the bookshelf can be rearranged by other facets at any time such as publication year, publication date, author name, or even their physical size or cover color. Peephole interaction allows to physically explore the digital bookshelf (see Figure 8.2D, page 203). Books and other media on the shelf are directly accessible (see Figure 8.2E and F, page 203). A tap selects a book or media and displays its content. For example, it opens a book in a book reader view or runs movies in a video player. In addition to search, a user can also create personal bookshelves and manually arrange books on the shelves just like with their real-world counterparts. As a digital power, however, personal bookshelves can be shared with others. For instance, a professor sharing a course reserve collection with students. Interestingly, today's tablet technologies such as Google Project Tango[5] would allow for such spatial navigation without the need to instrument an environment or devices.

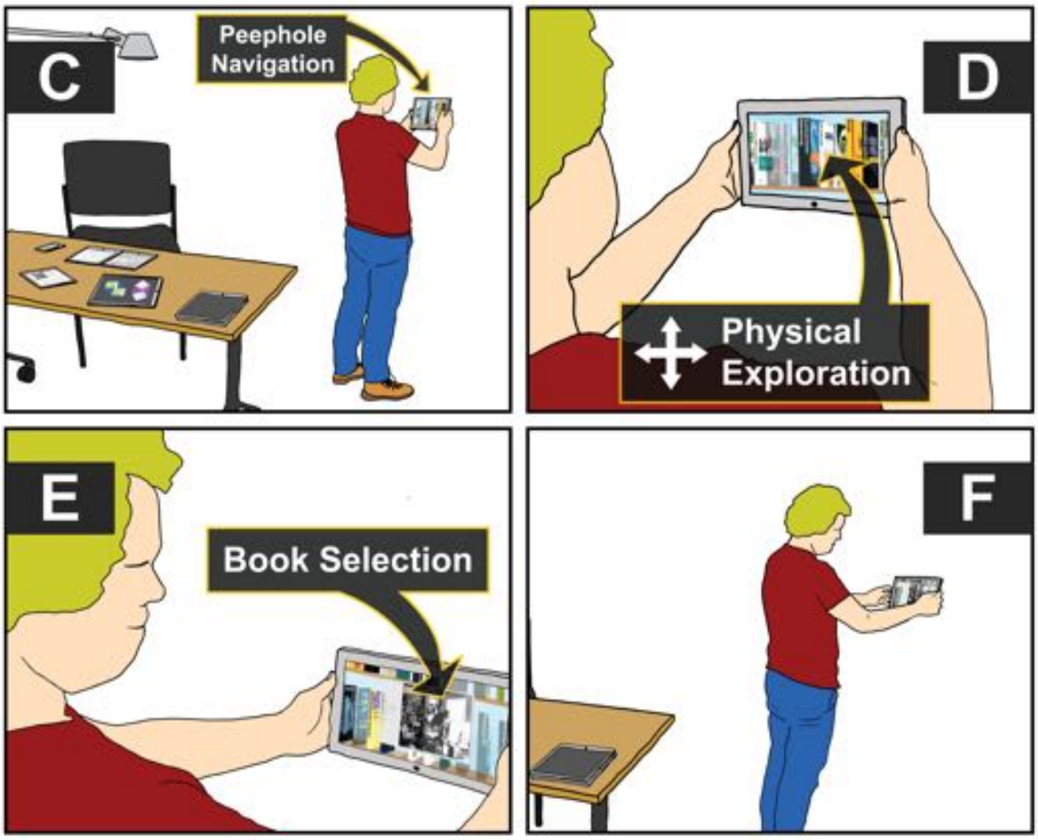


Figure 8.2.: Books and other digital media are browsed by physical interaction using peephole navigation.

---

[5] Project Tango technology gives a mobile device the ability to navigate the physical world similar to how we do as humans. – `https://www.google.com/atap/project-tango/` (last accessed: May 2nd, 2016)

Eventually, the information need is satisfied, and the user returns to the table with the tablet in book view. He moves two tablets aside and places the new digital book between them. From this book, the user can excerpt by dragging text and dropping it on the edge bubble that points towards the top left tablet (see Figure 8.3G, page 204). With the same interaction, the user can directly quote text and cite it in the handwritten document. Therefore, he drags the text from the source (book) and drops it on the edge bubble that points towards the word processor tablet. With the pen, he indicates the location at which the quote and citation should be inserted.

Figure 8.3.: Information is shared across devices using spatially-aware cross-device interactions such as edge bubbles.

Another advantage of spatial, cross-device interactions is the possibility to use devices' distance to implement spatial instrumental interaction (Beaudouin-Lafon, 2000). Figure 8.4 (page 205) and Figure 8.5 (page 205) exemplify two potential use cases for spatial, cross-device instrumental interaction. The smartphone, therefore, functions as the instrument, which also adjusts to the current context. For example, the context depends on the distance of the smartphone to any other device. In the first example and when the smartphone is placed next to the tablet in a book view, it shows a search input box. This search box enables the user to search for keywords within the book. Search results are presented on the smartphone and a selection of a result opens the corresponding book page on the tablet Figure 8.4 (page 205). In the second example, the smartphone acts as color palette to change the pen tip (see Figure 8.5., page 205) For example, alter the pen's function and color to a text marker to highlight text in a document.

This scenario implements proposed design guidelines as mentioned earlier. It demonstrates the applicability of them to the problem space explored in Chapter 3 (page 57)

Figure 8.4.: Devices such as smartphones act as tools or instruments to manipulate views or substances displayed on spatially distant devices. In this case, a smartphone triggers keyword search in the next-by tablet and further list all search results. Tapping on a search result opens the corresponding page on the tablet.

Figure 8.5.: Similar to the keyword search Figure 8.4, (page 205) the smartphone acts as color palette to define the pen tip color when writing on the tablet next-by.

and proposes design solutions for therein identified problems. Future studies are needed to empirically validate this re-design and to gain further insights on such spatial navigation and cross-device interactions for knowledge work.

8.5 Contributions Overview

Before concluding with future work, Table 8.1 (page 213) summarizes all thesis findings in an overview. Columns represent the different research phases in this thesis work beginning with: (i) field studies in the Library of the University of Konstanz and analysis (Chapter 3), (ii) design-oriented research through user studies with fully functional research prototypes (Chapter 3), (iii) research-oriented design through controlled lab experiments (Chapters 4-7), and a transfer of overall research findings to envision future knowledge workspaces (current Chapter). The rows represent a horizontal path through this thesis that connects identified problems with design solutions and findings from controlled experiments. From left to right it presents (i) tackled research objectives and problem space, (ii) a summary for each research phase, and (iii) particular findings and issues. Issues are highlighted with a yellow cell background.

		Library Context and Analysis through Field Studies	Design-oriented Research through Research Prototypes	Research-oriented design through Experiments	Transfer of Findings to Envision Future Knowledge Workspaces
Bibliographic Search — Spatial Navigation (RO1)	**Summary**	Physical libraries and especially bookshelves offer a natural physical exploration of library collections and also inherently support serendipitous discoveries (see Section 3.2.4, page 70).	Blended Shelf mimics a shelf-like browsing experience to explore digital library collections. Its visual design and interaction design was inspired by the Reality-Based Interaction Framework (Jacob et al., 2008) (see Section 3.3.2, page 79).	Peephole interaction can be an alternative approach to large screens or multi-touch interaction and to explore entire library collections in physical space (see Chapters 4 and 5).	A future system could allow for physical exploration of library collection through peephole interaction (see Figure 8.2C-F, page 203).
	Findings	In contrast to OPACs and their convergent (goal-directed) search behavior, bookshelves offer a divergent often explorative search behavior (see Section 3.2.4, page 70).	Blended Shelf combines convergent and divergent search in one user interface, which was considered as aesthetically pleasing by users (see Section 3.3.2, page 79).	Users conceive peephole interaction as “natural” and as easy to understand. (see Chapter 5, page 119)	An ideal system combines advantages of Blended Shelf and a “natural” and easy to understand peephole interaction.

Table 8.1.: Overall summary of field study findings and issues, findings of the experimental research, and contributions and transfer to envision future knowledge workplaces. Issues are highlighted with a yellow cell background.

		Library Context and Analysis through Field Studies	**Design-oriented Research through Research Prototypes**	**Research-oriented design through Experiments**	**Transfer of Findings to Envision Future Knowledge Workspaces**
		Current library systems do not unify both search behaviors, convergent and divergent, in a single system (see Section 3.1.2, page 59).			
		Only media in a tangible form are accessible by physical exploration (see Section 3.2.4, page 70).			
		Library users recognize shelf-browsing as valuable experience (see Section 3.2.4, page 70).	Users consider Blended Shelf as complement to shelf browsing (see Section 3.3.2, page 79).	A physical egocentric navigation exploits users' spatial memory, lowers their mental demand, and increases their experiences (see Chapter 5, page 119).	Future virtual bookshelves explorable by peephole interaction could exploit spatial memory and eventually replace their physical counterparts.
			Interaction, however, was often not immediately understood by users (see Section 3.3.2, page 79).		

Table 8.1.: Overall summary of field study findings and issues, findings of the experimental research, and contributions and transfer to envision future knowledge workplaces. Issues are highlighted with a yellow cell background.

		Library Context and Analysis through Field Studies	Design-oriented Research through Research Prototypes	Research-oriented design through Experiments	Transfer of Findings to Envision Future Knowledge Workspaces
		Shelf-browsing is only supported in open-stack libraries (see Section 3.2.4, page 70).	The 3D visualization of Blended Shelf allows users to explore entire library collection including electronic as well as closed-stack media (see Section 3.3.2, page 79).	Small tablet-sized peephole are a "sweet spot" for peephole interaction and provide sufficient privacy (see Chapter 4, page 95).	Library users often own mobile devices such as tablets. They could use these tablets to privately explore library collections. Also when users bring their personal devices, libraries do not have to buy expensive hardware and maintain large displays for exploring library collections.
			Because of the large display, however, users raised privacy concerns when searching for and browsing through socially inept or sensitive literature (see Section 3.3.2, page 79).		
			Hardware for Blended Shelf is arguably expensive and furthermore needs to be maintained by library staff or external contractors (see Section 3.3.2, page 79).		

Table 8.1.: Overall summary of field study findings and issues, findings of the experimental research, and contributions and transfer to envision future knowledge workplaces. Issues are highlighted with a yellow cell background.

		Library Context and Analysis through Field Studies	**Design-oriented Research through Research Prototypes**	**Research-oriented design through Experiments**	**Transfer of Findings to Envision Future Knowledge Workspaces**
Read & Write — Cross-Device Interaction (RO2)	**Summary**	The use of and interaction in space is an important cognitive resource during work desk activities (Kirsh, 2010).	Integrative Workplace blends the qualities of digital and physical work artifacts. It enables users to apply a digital function to analog and digital media alike. The design was inspired by the Blended Interaction framework (Jetter et al., 2014).	Cross-device interaction is an alternative approach to highly instrumented workplaces such as Integrative Workplace (see Chapters 6 and 7).	A future system could allow for fluid configuration of work artefacts to enable users to think and act in space (Kirsh, 2010) (see Figure 8.3, page 204).

Table 8.1.: Overall summary of field study findings and issues, findings of the experimental research, and contributions and transfer to envision future knowledge workplaces. Issues are highlighted with a yellow cell background.

		Library Context and Analysis through Field Studies	**Design-oriented Research through Research Prototypes**	**Research-oriented design through Experiments**	**Transfer of Findings to Envision Future Knowledge Workspaces**
	Findings	Knowledge workers think and act in space. For example, by stacking books to keep track of search trajectories or to spatially reflect current thoughts (see Section 3.2.4, page 74).	With Integrative Workplace, users can arrange physical as well as digital work artifacts in space (see Section 3.3.3, page 85).	With HuddleLamp, work artefacts can be arranged in physical space similar to Integrative Workplace (see Chapter 6, page 143). However, in contract to Integrative Workplace, unused artifacts can be put aside or stacked for later use.	A future system should support different working styles. It should be flexible enough to adapt to the user's current task at hand. For example, a sensemaking task eventually leads to different device configurations than reading and writing with multiple documents.
		With prevalent knowledge work tools, it is cumbersome to arrange digital media in physical space and keep track of their locations (see Section 3.2.4, page 74).			
		Knowledge workers often work with different media, both analog and digital (see Section 3.2.4, page 74).	Integrative Workplace allows users to search and excerpt content from both digital and printed documents using the same interaction techniques (see Section 3.3.3, page 85).	HuddleLamp allows for interactions across devices. For example, through drag-and-flick gestures (see Section 6.3.1, page 160) or above the surface interactions (see Section 6.3.1, page 162).	A future system should allow distributing content to several devices or let users store information on other devices for later use (see Figure 8.3, page 204). Fluid device configuration would even

Table 8.1.: Overall summary of field study findings and issues, findings of the experimental research, and contributions and transfer to envision future knowledge workplaces. Issues are highlighted with a yellow cell background.

		Library Context and Analysis through Field Studies	**Design-oriented Research through Research Prototypes**	**Research-oriented design through Experiments**	**Transfer of Findings to Envision Future Knowledge Workspaces**
		The current gap between analog and digital world prevent knowledge workers to easy transition between both. For example, to digitally excerpt from a book using well-known drag and drop interactions (see Section 3.2.4, page 74).		Users understand spatial interactions such as edge bubbles. They even prefer such interaction over spatially-agnostic interactions (see Chapter 7, page 169).	enable new interaction paradigms such as instrumental interaction (Beaudouin-Lafon, 2000) where devices change their roles and function according to their location in space or distance to other devices (see Figure 8.4, page 205) and (see Figure 8.5, page 205).

Table 8.1.: Overall summary of field study findings and issues, findings of the experimental research, and contributions and transfer to envision future knowledge workplaces. Issues are highlighted with a yellow cell background.

		Library Context and Analysis through Field Studies	Design-oriented Research through Research Prototypes	Research-oriented design through Experiments	Transfer of Findings to Envision Future Knowledge Workspaces
			Markers need to be printed on physical documents in order to identify them and track their location and orientation over time (see Section 3.3.3, page 85).	HuddleLamp provides a low-cost solution for tracking multiple mobile devices. It only requires little instrumentation of the environment (see Chapter 6, page 143).	When computing power, tracking, and a camera is integrated into a single lightbulb, a low-cost tracking could be installed by non-expert users. Thereby, enabling cross-device experiences could be as simple as screwing in a lightbulb.
			Hardware for Integrative Workplace is expensive to buy and maintain. Moreover, a physical table needs to be instrumented in advance (see Section 3.3.3, page 85).		

Table 8.1.: Overall summary of field study findings and issues, findings of the experimental research, and contributions and transfer to envision future knowledge workplaces. Issues are highlighted with a yellow cell background.

8.6 Future Directions

This research on spatial navigation and cross-device interactions answered research questions raised at the beginning of this thesis. But it also uncovered new possibilities for future research directions on the way. So far, all of the research were conducted as lab experiments. However, we also learned from ecological psychology that human behavior depends on the environment in which a study is conducted. Instead of solely focussing on basic research, research could be extended to "the wild" by embracing methods from ecological psychology. The following list samples possible projects that could be subject to future research.

8.6.1 Study on Long-Term Spatial Memory

Despite the fact that the results of our spatial memory study in Chapter 5 (page 119) show a significant difference in navigation performance, we could not observe any statistically significant differences for the spatial memory. However, the second experiment E2 in Chapter 5 (page 119) hinted at a potential effect of the greater proprioceptive and kinesthetic feedback on long-term spatial memory, which could be investigated in a follow-up study and eventually longitudinal. Future work could examine whether the implicitly better spatial memory during the navigation task for the peephole navigation and not significantly different spatial memory in the specific recall task is due to muscle memory. Moreover, this could also be used to investigate how peephole navigation performs in collaborative environments, for example, multi-display environments for mixed-focus collaboration (Tang et al., 2006).

The ShelfHole concept (see Figure 8.6, page 215) could serve as real-world application providing an ecologically valid task. As illustrated in Figure 8.6 (page 215), spatial tracking of off-the-shelf devices could be implemented by exploiting low-cost augmented reality (AR) tracking technology such as Studierstube (Schmalstieg et al., 2002) and a marker wall or using Google Project Tango tablets (Scan QR Code #19 or enter “19” in the *MediaBrowser* application).

8.6.2 Peephole Navigation in Collaborative Mixed-Reality Environments

Social interaction plays a crucial role in mixed-focus collaboration. Building on work by Billinghurst et al. (Billinghurst et al., 2015), we could learn and understand the subtleties of social interactions and groups’ communication behaviors during collab-

Video

Code #19

Figure 8.6.: The ShelfHole concept. A low-cost peephole interaction for reality-based exploration of library collections.

orative, spatial tasks in mixed-reality environments (e.g., urban planning). We could investigate on research questions like "What problems occur when multiple users concurrently manipulate virtual elements?" and "How do they mitigate when such problems occur?" Together with Müller et al., we took an initial stab on observing what effect virtual objects as spatial cues have in collaborative mixed-reality environments. Thereby, we particularly focussed on how spatial cues shape communication behavior and user task load during collaboration (Müller et al., 2016).

Investigations on long-term and short-term spatial memory could also be extended to the third dimension. With emerging mixed-reality technologies such as Google Project Tango or Microsoft HoloLens, further research is necessary to give explanation to questions like "How is (long-term) spatial memory affected when virtual elements are no longer visually distinguishable from physical elements?", "What role does missing haptic feedback for virtual elements play for spatial memory?", or "Are there any differences in how humans cognitively process mixed-reality information?"

8.6.3 Research Cross-Device Interactions in a variety of Application Domains

With our work on cross-device interaction, we have only scratched the surface of what can be done in future with ad-hoc communities of spatially-aware and reconfigurable displays and devices. We consider HuddleLamp only as the first step towards realizing this greater vision. One of the great advantages that we see in

HuddleLamp or similar approaches is that they make digital tools for co-located collaboration available to the masses. In future, the HuddleLamp's tracking camera could be integrated into a lightbulb together with a processing unit for computer vision tracking. Figure 8.7 (page 217) shows another use case for HuddleLamp (Scan QR Code #20 or enter "20" in the *MediaBrowser* application). It illustrates the HuddleLamp Design Tank, a mobile trolley that unfolds opportunities for ad hoc creative design sessions, similar to IdeaVis or AffinityTable (Geyer, 2013). Thereby, these technologies can be used where motion-capturing tracking systems and large tabletops or surfaces are too expensive to buy, setup, and maintain.

Software

Code #21

As the first step in this direction, we implemented Connichiwa, a versatile framework to build cross-device experiences that work with off-the-shelf devices and does not require additional hardware (Schreiner et al., 2015a) (Scan QR Code #21 or enter "21" in the *MediaBrowser* application) . Consequentially, we are now working on applications for engaging a great variety of users, e.g., playful exploration of ebook collections in public libraries (Rädle et al., 2012a) or community centers, collaborative search tools for schools (Rädle et al., 2013b; Zagermann et al., 2015), novel kinds of informational or therapeutic games (Marwecki et al., 2013), or in museums (Klinkhammer et al., 2011).

Video

Code #20

Figure 8.7.: HuddleLamp integrated in a design tank. The rendering of the 3D model (top) and the result implemented by the local workshops of the University of Konstanz (bottom).

Interview Guidlines

A

Interviewfragebogen

Interviewpartner zuerst die Fragen ohne Hilfestellung beantworten lassen und anschließend, wenn erforderlich, per Nachfrage und mit Beispielen vertiefen.

Studierende im UB-Hauptgebäude oder vor der Mensa ansprechen: „Entschuldigung, dürfte ich fragen in welchem Semester Du bist?“ Nur Studierende ab dem 3. FS interessant.

1. In welchem Fachsemester bist Du und welchem Fachbereich gehörst Du an?
2.
   a) Nenne bitte so vollständig wie möglich alle verschiedene Gruppenarbeitssituationen in denen Du während Deines Studiums mit Kommilitonen zusammengearbeitet hast? (zB Lerngruppen, Referatsgruppen, Examensvorbereitung usw.)
   b) Wenn Du durch Freunde anderer Fachbereiche von anderen Gruppenarbeitssituationen gehört hast, nenne bitte auch diese.
   c) Bitte spezifiziere das jeweils zu erarbeitende Endergebnis.
3. Welche Hilfsmittel & Materialien habt Ihr in den Gruppenarbeitssituationen eingesetzt? (zB Bücher, Vorlesungen digital od. ausgedruckt, Laptop)
4. Wo haben Eure Treffen vorwiegend stattgefunden und warum? (Ruhe, Umgebung, bestimmte Hilfsmittel wie Tafel, Whiteboard, Kaffeemaschine)
5. Wenn Ihr Euch getroffen habt, was waren die Hauptgründe Eurer Treffen? (zB Planung, Strukturierung, Arbeitsteilung, Kommunikation, kollaboratives Erarbeiten/ Schreiben, Erklären, Diskutieren usw.)
6. Kannst Du Dich an Hilfsmittel, Materialien, Techniken o.ä. Erinnern, das Euch im Moment der Gruppenarbeit gefehlt hatte?
7. Wo gab es die größten Probleme während der Zusammenarbeit? Was waren positive Aspekte?
8. Was könnte Euch in diesem Moment die Zusammenarbeit besonders erleichtern ohne Euch die Arbeit abzunehmen?

Wenn es nicht zur Sprache kam:

- auf die verschiedenen genutzten und/oder nicht genutzten Medien der Bibliothek eingehen.
- auf die Literatursuche, -bewertung und -relevanz eingehen. (Wie wird gesucht? Was sind Eure Erfahrungen?)
- auf die Erfahrung mit Smartphones und Computern eingehen.
- Im letzten Punkt auf unterstützende Maßnahmen eingehen: Strukturierungshilfe, Recherchehilfe

B

Paper Questionnaire

blended
Library

Das Ziel unseres Projektes ist es Konzepte für „die Bibliothek und den Arbeitsplatz der Zukunft“ zu entwickeln. Es sollen sowohl Suchprozesse, wissenschaftliches Bearbeiten von Dokumenten als auch Zusammenarbeit in der Bibliothek zukünftig stärker unterstützt werden. Ferner möchten wir einen Eindruck davon bekommen, was den potentiellen Benutzern, also Dir, wichtig ist. Wir suchen Deine Konzepte, Bedürfnisse und Ideen, um die Bibliothek der Zukunft als ganzheitliches System zu gestalten.

Unter den Teilnehmern aus Konstanz verlosen wir am Ende der Umfrage Kaffee-Gutscheine.

25x Kaffee-Gutschein (Seezeit / Campus Café)

Wir freuen uns auf Ihre Teilnahme!

EvaSys	Blended Library - Bedarfsanalyse	Electric Paper

Universität Konstanz
AG Reiterer

Roman Raedle
Bedarfsanalyse

Markieren Sie so: ☐ ☒ ☐ ☐ ☐ Bitte verwenden Sie einen Kugelschreiber oder nicht zu starken Filzstift. Dieser Fragebogen wird maschinell erfasst.
Korrektur: ☐ ■ ☐ ☒ ☐ Bitte beachten Sie im Interesse einer optimalen Datenerfassung die links gegebenen Hinweise beim Ausfüllen.

1. Demographische Daten

1.1 Wie alt sind Sie?

1.2 Studieren Sie derzeit?
☐ Ja ☐ Nein

1.3 Wenn Sie derzeit **studieren**: Welchen Studiengang habe Sie? (bitte ausgeschrieben)

1.4 Wenn Sie derzeit **studieren**: In welchem Hochschul-Semester sind Sie?

1.5 Wenn Sie derzeit **studieren**: Für welchen Abschluss sind Sie aktuell eingeschrieben (bei Sonstiges bitte ausgeschrieben)
☐ Bachelor ☐ Master ☐ Diplom
☐ Magister ☐ Doktor ☐ Sonstiges

1.6 Wenn Sie derzeit **nicht studieren**: Welchen Beruf haben Sie?

1.7 In welcher Stadt studieren beziehungsweise arbeiten Sie (bitte ausgeschrieben)

1.8 Was ist Ihre Muttersprache?
☐ Deutsch ☐ Englisch ☐ Spanisch
☐ Französisch ☐ Russisch ☐ Türkisch
☐ Andere

1.9 Welche weiteren Sprachen sprechen Sie?
☐ Deutsch ☐ Englisch ☐ Französisch
☐ Spanisch ☐ Russisch ☐ Türkisch
☐ Andere

1.10 Wie viele Stunden arbeiten Sie täglich im Durchschnitt am Computer, Tablet-PC oder Smartphone?

2. Rechercheverhalten

2.1 Welche Medien(arten) verwenden Sie beim wissenschaftlichen Arbeiten?

☐ Bücher
☐ eBooks
☐ Elektronische Texte
☐ Zeitschriften (Printmedien)
☐ Vorlesungsaufzeichnungen
☐ Videos (DVD, VHS)
☐ Sonstiges

2.2 In welchen Katalogen/Suchmasken suchen Sie nach weiterführender Literatur?

☐ Suche im Bücher-Regal
☐ Web-Suchmaschine
☐ Lokaler Katalog (Libero)
☐ KonSearch
☐ Fachdatenbank
☐ Sonstige

2.3 Welchen Dienst oder welches Programm nutzen Sie zur Unterstützung bei der Literaturverwaltung?

☐ Citavi
☐ EndNote
☐ Mendeley
☐ Bibliographix
☐ RefWorks
☐ Sonstige

2.4 Warum nutzen Sie die oben ausgewählten Programme beziehungsweise was gefällt Ihnen an diesen?

2.5 Suchen Sie allein oder gemeinsam mit Anderen nach relevanter Literatur?

☐ Allein
☐ Gemeinsam
☐ Beides

2.6 Suchen Sie gezielt nach nur einem spezifischen Medium oder gleich nach mehreren relevanten Medien?

☐ Ein Medium
☐ Mehrere
☐ Beides, je nach Situation
☐ Sonstiges

2.7 Wie verfahren Sie mit relevanten Informationen aus **analogen** Quellen?

☐ Kopieren
☐ Digitale Abschrift anfertigen
☐ Handschriftlich zusammenfassen
☐ Einscannen
☐ Sonstiges

2.8 Wie verfahren Sie mit relevanten Informationen aus **digitalen** Quellen?

☐ Ausdrucken
☐ Digital bearbeiten
☐ Copy & Paste in anderes Dokument
☐ Sonstiges

3. Lesen & Bearbeiten

3.1 Mit welchen Materialien und Hilfsmitteln arbeiten Sie beim wissenschaftlichen Arbeiten und beim Verwalten Ihrer Quell-Literatur?

☐ Stift & Papier ☐ Computer ☐ Tablet-PC / Smartphone

Stift & Papier

3.2 In welchen Situationen und warum arbeiten Sie wissenschaftlich mit Stift & Papier?

3.3 Welche Probleme treten hierbei auf und was stört Sie?

Computer

3.4 In welchen Situationen und warum arbeiten Sie wissenschaftlich mit dem Computer?

3.5 Welche Programme nutzen Sie hierfür an Ihrem Computer?

3.6 Welche Probleme treten hierbei auf und was stört Sie?

Tablet-PC / Smartphone

3.7 In welchen Situationen und warum arbeiten Sie wissenschaftlich mit dem Tablet-PC / Smartphone?

3.8 Welche Programme nutzen Sie hierfür an Ihrem Tablet-PC / Smartphone?

3.9 Welche Probleme treten hierbei auf und was stört Sie?

3. Lesen & Bearbeiten [Fortsetzung]

3.10 Was fehlt Ihnen beim wissenschaftlichen Arbeiten und dem Verwalten der Quell-Literatur?

Folgende Gründe werden für das Lesen angeführt. Bitte lesen Sie diese zunächst durch.
Unabhängig von der Häufigkeit, sortieren Sie diese bitte nach Ihrer persönlichen Wichtigkeit von (1-10) ohne eine Spalte doppelt zuzuordnen.

3.11 **Cross-Referenzierung**: Lesen, um Informationen der Dokumente in Texte zu integrieren	sehr unwichtig	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	sehr wichtig
3.12 **Lesen, um Fragen zu suchen/zu beantworten**: Zielgerichtete Suche, um mit Hilfe der Informationen Entscheidungen zu treffen		☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	
3.13 **Diskussionsunterstützung**: Lesen, um auf Texte während Diskussionen verweisen zu können		☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	
3.14 **Skimming**: schnelles Durchblättern, Überfliegen eines Textes, um schnell eine grobe Idee des Inhaltes zu bekommen		☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	
3.15 **Text prüfen**: Lesen, um Text zu editieren oder kritisch zu prüfen		☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	
3.16 **Selbst-Informieren**: Informationen sammeln, um Wissen ohne spezifisches Ziel / Anwendungsbereich zu generieren, häufig genutzt um ein Thema erst einmal zu verstehen		☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	
3.17 **Eigenen Text lesen zur Erinnerung**: eigene Notizen & Texte durchlesen, um die nächsten Schritte zu ermitteln		☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	
3.18 **Identifikation**: Lesen, um zu verstehen, um was für eine Art Dokument/Literatur es sich handelt		☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	
3.19 **Lernen**: Lesen, um sich zu einem späteren Zeitpunkt auf Informationen zu beziehen oder diese anzuwenden		☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	
3.20 **Zuhören unterstützen**: Lesen um das Zuhören während eines Vortrages zu unterstützen		☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	

3. Lesen & Bearbeiten [Fortsetzung]

3.21 Wie verfahren Sie, wenn Sie in einer Quell-Literatur einen Verweis zu einer anderen Literaturquelle einsehen müssen?

☐ Lesen Sie die Quelle zuerst zu Ende und lesen dann die referenzierte Literatur?
☐ Suchen Sie zunächst die Referenz-Literatur heraus und lesen anschließend die ursprüngliche Literatur weiter?
☐ Beides
☐ Sonstiges

3.22 Weshalb suchen Sie nach weiteren Quellen?

☐ Terminologie-Probleme oder (Wort-)Definitionen
☐ Zitate nachvollziehen
☐ Inhalte nachvollziehen
☐ Weitere Quellen für die weitere Recherche des Themas
☐ Thema vertiefen
☐ Überblick über Thema verschaffen
☐ Sonstiges

3.23 Wie schätzen Sie Ihre Lesegeschwindigkeit generell ein?

sehr langsam ☐ ☐ ☐ ☐ ☐ sehr schnell ☐ keine Angabe

3.24 Sehen Sie bezüglich Ihrer generellen Lesegeschwindigkeit Verbesserungspotenzial?

☐ Ja
☐ Nein

3.25 **Weitere Ideen, Anregungen, Wünsche oder Anforderungen?**
Geplant ist ein System, das Sie möglichst ganzheitlich und ohne Unterbrechungen beim Lesen und Bearbeiten von Quell-Literatur beim wissenschaftlichen Arbeiten unterstützen soll, was wünschen Sie sich hierfür?

4. Weitere Studien

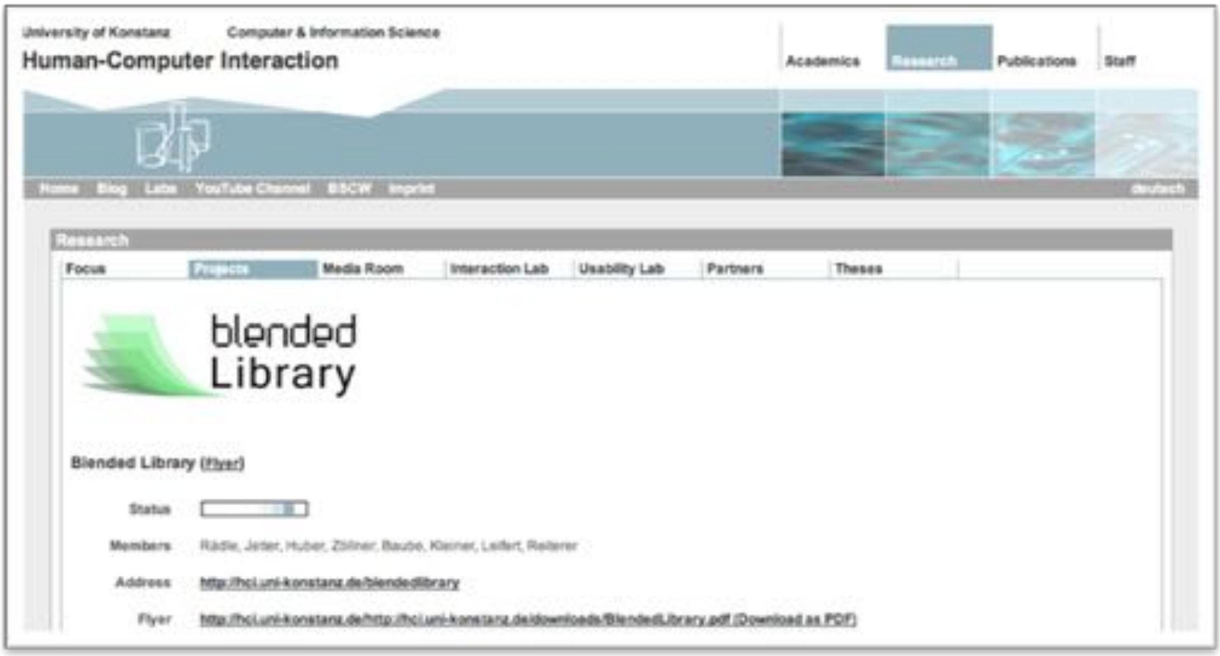
4.1 **Haben Sie Interesse an weiteren Studien für die Blended Library teilzunehmen, sind an den Ergebnissen dieser Studie interessiert oder möchten an der Verlosung teilnehmen?**
Hinterlassen Sie uns Ihre E-Mail Adresse damit wir Sie dementsprechend kontaktieren können.

blended
Library

Vielen Dank für Ihre Teilnahme!

Wir möchten uns ganz herzlich für Ihre Mithilfe bedanken. Besuchen Sie auch regelmäßig unsere Webseite, um die Fortschritte des Projekts Blended Library zu verfolgen.

Projektwebseite Blended Library - http://hci.uni-konstanz.de/blendedlibrary



Blended Library (Flyer)

Status

Members Rädle, Jetter, Huber, Zöllner, Baube, Kleiner, Leifert, Reiterer

Address http://hci.uni-konstanz.de/blendedlibrary

Flyer http://hci.uni-konstanz.de/http://hci.uni-konstanz.de/downloads/BlendedLibrary.pdf (Download as PDF)

Bibliography

Anderson, J. R. (1996). "ACT: A simple theory of complex cognition." In: *American Psychologist* 51.4, pp. 355–365. DOI: `10.1037/0003-066X.51.4.355` (cit. on p. 36).

Andrews, C. Endert, A. North, C. (2010). "Space to Think: Large, High-Resolution Displays for Sensemaking". In: *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*. New York, New York, USA: ACM Press, p. 55. DOI: `10.1145/1753326.1753336` (cit. on pp. 4–6, 33, 46, 76, 120, 172, 199).

Arai, T. Aust, D. Hudson, S. E. (1997). "PaperLink: A Technique for Hyperlinking from Real Paper to Electronic Content". In: *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '97*. New York, New York, USA: ACM Press, pp. 327–334. DOI: `10.1145/258549.258782` (cit. on p. 86).

Atkinson, R. Shiffrin, R. (1968). "Human memory: A proposed system and its control processes". In: *The psychology of learning and motivation*. Ed. by K. W. Spence and J. T. Spense. Vol. 2. New York: Academic Press, Inc., pp. 89–195 (cit. on p. 47).

Aylett, M. P. Quigley, A. J. (2015). "The Broken Dream of Pervasive Sentient Ambient Calm Invisible Ubiquitous Computing". In: *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '15*. New York, NY, USA: ACM, pp. 425–435. DOI: `10.1145/2702613.2732508` (cit. on pp. 26, 31, 32, 35).

Baddeley, A. (2000). "The episodic buffer: a new component of working memory?" In: *Trends in Cognitive Sciences* 4.11, pp. 417–423. DOI: `10.1016/S1364-6613(00)01538-2` (cit. on p. 48).

Baddeley, A. D. Hitch, G. (1974). "Working memory". In: *The psychology of learning and motivation* 8, pp. 47–89 (cit. on pp. 47, 49).

Balakrishnan, R. (2004). ""Beating" Fitts' law: virtual enhancements for pointing facilitation". In: *International Journal of Human-Computer Studies* 61.6, pp. 857–874. DOI: `10.1016/j.ijhcs.2004.09.002` (cit. on p. 128).

Ball, R. North, C. (2008). "The effects of peripheral vision and physical navigation on large scale visualization". In: *Proceedings of Graphics Interface 2008 - GI '08*. Toronto, Ont., Canada: Canadian Information Processing Society, pp. 9–16 (cit. on pp. 6, 113, 120).

Ball, R. North, C. Bowman, D. A. (2007). "Move to Improve: Promoting Physical Navigation to Increase User Performance with Large Displays". In: *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '07*. New York, New York, USA: ACM Press, p. 191. DOI: `10.1145/1240624.1240656` (cit. on pp. 6, 120, 138, 139).

Ballendat, T. Marquardt, N. Greenberg, S. (2010). "Proxemic Interaction: Designing for a Proximity and Orientation-Aware Environment". In: *ACM International Conference on Interactive Tabletops and Surfaces - ITS '10*. New York, New York, USA: ACM Press, p. 121. DOI: `10.1145/1936652.1936676` (cit. on p. 163).

Bangor, A. Kortum, P. T. Miller, J. T. (2008). "An Empirical Evaluation of the System Usability Scale". en. In: *International Journal of Human-Computer Interaction* 24.6, pp. 574–594. DOI: `10.1080/10447310802205776` (cit. on p. 88).

Baudisch, P. Rosenholtz, R. (2003). "Halo: a technique for visualizing off-screen objects". In: *Proceedings of the conference on Human factors in computing systems - CHI '03*. 5. New York, New York, USA: ACM Press, p. 481. DOI: `10.1145/642611.642695` (cit. on p. 182).

Beaudouin-Lafon, M. (2000). "Instrumental Interaction: An Interaction Model for Designing Post-WlMP User Interfaces". In: *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '00*. New York, New York, USA: ACM Press, pp. 446–453. DOI: `10.1145/332040.332473` (cit. on pp. 3, 204, 212).

Bederson, B. B. (2011). "The promise of zoomable user interfaces". In: *Behaviour & Information Technology* 30.6, pp. 853–866. DOI: `10.1.1.147.6981` (cit. on pp. 120, 121).

Bell, G. Dourish, P. (2006). "Yesterday's tomorrows: notes on ubiquitous computing's dominant vision". In: *Personal and Ubiquitous Computing* 11.2, pp. 133–143. DOI: `10.1007/s00779-006-0071-x` (cit. on pp. 26, 29, 31).

Bezerianos, A. Isenberg, P. (2012). "Perception of Visual Variables on Tiled Wall-Sized Displays for Information Visualization Applications". In: *IEEE Transactions on Visualization and Computer Graphics* 18.12, pp. 2516–2525. DOI: `10.1109/TVCG.2012.251` (cit. on p. 6).

Biehl, J. T. Bailey, B. P. (2004). "ARIS: an interface for application relocation in an interactive space". In: *In Proceedings of Graphics Interface 2004 - GI '04*. Waterloo, Ontario, Canada: Canadian Human-Computer Communications Society, pp. 107–116 (cit. on pp. 172, 174, 180, 182).

Bier, E. A. Stone, M. C. Pier, K. Buxton, W. DeRose, T. D. (1993). “Toolglass and magic lenses: the see-through interface”. In: *Proceedings of the 20th annual conference on Computer graphics and interactive techniques - SIGGRAPH ’93*. Mmm. New York, New York, USA: ACM Press, pp. 73–80. DOI: `10.1145/166117.166126` (cit. on p. 138).

Billinghurst, M. Clark, A. Lee, G. (2015). “A Survey of Augmented Reality”. In: *Foundations and Trends in Human-Computer* 8.2, pp. 73–272. DOI: `10.1561/1100000049` (cit. on p. 214).

Bjorneborn, L. (2008). *Serendipity dimensions and users’ information behaviour in the physical library interface*. en (cit. on pp. 72, 73, 80).

Bødker, S. (2006). “When second wave HCI meets third wave challenges”. In: *Proceedings of the 4th Nordic conference on Human-computer interaction changing roles - NordiCHI ’06*. New York, New York, USA: ACM Press, pp. 1–8. DOI: `10.1145/1182475.1182476` (cit. on p. 77).

Bødker, S. Klokmose, C. N. (2016). “Dynamics , Multiplicity and Conceptual Blends in HCI”. In: *Proceedings of the 34th Annual ACM Conference on Human Factors in Computing Systems - CHI ’16* (cit. on p. 42).

Brooke, J. (1996). *SUS: A quick and dirty usability scale*. Tech. rep. (cit. on p. 87).

Butz, A. Krüger, A. (2014). *Mensch-Maschine-Interaktion*. De Gruyter, p. 235 (cit. on pp. 61, 64, 67).

Buxton, B. (2007). *Sketching User Experiences: Getting the Design Right and the Right Design*. First. Morgan Kaufmann (cit. on pp. 77, 78).

Cao, X. Balakrishnan, R. (2006). “Interacting with dynamically defined information spaces using a handheld projector and a pen”. In: *Proceedings of the 19th annual ACM symposium on User interface software and technology - UIST ’06*. New York, New York, USA: ACM Press, p. 225. DOI: `10.1145/1166253.1166289` (cit. on pp. 83, 84, 111).

Cao, X. Li, J. J. Balakrishnan, R. (2008). “Peephole Pointing: Modeling Acquisition of Dynamically Revealed Targets”. In: *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI ’08*. New York, New York, USA: ACM Press, p. 1699. DOI: `10.1145/1357054.1357320` (cit. on pp. 97, 112, 115).

Cecchinato, M. E. Cox, A. L. Bird, J. (2015). “Working 9-5?: Professional Differences in Email and Boundary Management Practices”. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI ’15*. New York, New York, USA: ACM Press, pp. 3989–3998. DOI: `10.1145/2702123.2702537` (cit. on p. 34).

Cecchinato, M. E. Sellen, A. Shokouhi, M. Smyth, G. (2016). “Finding Email in a Multi-Account, Multi-Device World”. In: *Proceedings of the 34th Annual ACM Conference on Human Factors in Computing Systems - CHI ’16* (cit. on pp. 9, 34, 195, 200).

Chattopadhyay, D. O’Hara, K. Rintel, S. Rädle, R. (2016). “Office Social: Presentation Interactivity for Nearby Devices”. In: *Proceedings of the 34th Annual ACM Conference on Human Factors in Computing Systems - CHI ’16*. New York, NY, USA: ACM Press (cit. on pp. xix, 90).

Chen, N. Guimbretiere, F. Sellen, A. (2012). “Designing a multi-slate reading environment to support active reading activities”. In: *ACM Transactions on Computer-Human Interaction* 19.3, pp. 1–35. DOI: `10.1145/2362364.2362366` (cit. on pp. 11, 170, 174, 175, 179).

Chen, N. Guimbretière, F. Sellen, A. (2013). “Graduate student use of a multi-slate reading system”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI ’13*. New York, New York, USA: ACM Press, p. 1799. DOI: `10.1145/2470654.2466237` (cit. on pp. 9, 11, 171).

Chi, P.-y. P. Li, Y. (2015). “Weave: Scripting Cross-Device Wearable Interaction”. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI ’15*. New York, NY, USA: ACM, pp. 3923–3932. DOI: `10.1145/2702123.2702451` (cit. on p. 195).

Chong, M. K. Mayrhofer, R. Gellersen, H. (2014). “A Survey of User Interaction for Spontaneous Device Association”. In: *ACM Computing Surveys* 47.1, pp. 1–40. DOI: `10.1145/2597768` (cit. on pp. 89, 152).

Craik, F. I. Lockhart, R. S. (1972). “Levels of processing: A framework for memory research”. In: *Journal of Verbal Learning and Verbal Behavior* 11.6, pp. 671–684. DOI: `10.1016/S0022-5371(72)80001-X` (cit. on p. 47).

Dachselt, R. (2014). “Multimodale Interaktion mit großen Displays”. In: *Informatik-Spektrum* 37.5, pp. 397–401. DOI: `10.1007/s00287-014-0798-0` (cit. on p. 7).

Dachselt, R. Buchholz, R. (2008). “Throw and Tilt – Seamless Interaction across Devices Using Mobile Phone Gestures”. In: Lecture Notes in Informatics Series (LNI), pp. 272–278 (cit. on p. 172).

Dachselt, R. Buchholz, R. (2009). “Natural throw and tilt interaction between mobile phones and distant displays”. In: *Proceedings of the 27th international conference extended abstracts on Human factors in computing systems - CHI EA ’09*. New York, New York, USA: ACM Press, p. 3253. DOI: `10.1145/1520340.1520467` (cit. on p. 172).

Dachselt, R. Sarmad, A.-S. (2011). “Interacting with Printed Books Using Digital Pens and Smart Mobile Projection”. In: *Proceedings of the Workshop on Mobile and Personal Projection (MP2) @ ACM CHI 2011* (cit. on p. 86).

Dam, A. (1997). “Post-WIMP user interfaces”. In: *Communications of the ACM* 40.2, pp. 63–67. DOI: `10.1145/253671.253708` (cit. on p. 2).

Dearman, D. Pierce, J. S. (2008). “It’s on my other computer!: computing with multiple devices”. In: *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI ’08*. New York, New York, USA: ACM Press, p. 767. DOI: `10.1145/1357054.1357177` (cit. on pp. 3, 9, 34, 195, 199).

Dearman, D. Guy, R. Truong, K. (2012). “Determining the orientation of proximate mobile devices using their back facing camera”. In: *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI ’12*. CHI ’12. New York, New York, USA: ACM Press, p. 2231. DOI: `10.1145/2207676.2208377` (cit. on p. 164).

Deutsche Akkreditierungsstelle Technik, (2009). *Leitfaden Usability: Gestaltungsrahmen für den Usability-Engineering-Prozess*. Frankfurt/Main (cit. on p. 61).

Dippon, A. Wiedermann, N. Klinker, G. (2012). “Seamless integration of mobile devices into interactive surface environments”. In: *Proceedings of the 2012 ACM international conference on Interactive tabletops and surfaces - ITS ’12*. New York, New York, USA: ACM Press, p. 331. DOI: `10.1145/2396636.2396693` (cit. on p. 148).

Dostal, J. Nacenta, M. Rädle, R. Reiterer, H. Stellmach, S. (2013). “ITS 2013 workshop on visual adaptation of interfaces”. In: *Proceedings of the 2013 ACM international conference on Interactive tabletops and surfaces - ITS ’13*. New York, New York, USA: ACM Press, pp. 491–492. DOI: `10.1145/2512349.2512353` (cit. on p. 17).

Dourish, P. (1999). “Embodied Interaction: Exploring the Foundations of a New Approach to HCI” (cit. on pp. 15, 26, 34, 192).

Dourish, P. (2001). *Where the Action Is: The Foundations of Embodied Interaction*. MIT Press, p. 256 (cit. on pp. 15, 34, 36, 37).

Dumas, B. Lalanne, D. Oviatt, S. (2009). “Multimodal Interfaces: A Survey of Principles, Models and Frameworks”. In: *Human Machine Interaction*, pp. 3–26 (cit. on pp. 26, 49).

Durrant, A. Golembewski, M. Kirk, D. S. Benford, S. Rowland, D. McAuley, D. (2011). “Exploring a digital economy design space in theme parks”. In: *Procedings of the Second Conference on Creativity and Innovation in Design - DESIRE ’11*. New York, New York, USA: ACM Press, p. 273. DOI: `10.1145/2079216.2079256` (cit. on p. 27).

Fallman, D. (2007). “Why Research-Oriented Design Isn’t Design-Oriented Research: On the Tensions Between Design and Research in an Implicit Design Discipline”. In: *Knowledge, Technology & Policy* 20.3, pp. 193–200. DOI: `10.1007/s12130-007-9022-8` (cit. on pp. 77, 78).

Fauconnier, G. Turner, M. (1998). “Conceptual integration networks”. In: *Cognitive Science* 22.2, pp. 133–187. DOI: `10.1016/S0364-0213(99)80038-X` (cit. on pp. 5, 39, 40, 43).

Fauconnier, G. Turner, M. (2003). *The Way We Think: Conceptual Blending And the Mind’s Hidden Complexities*. Basic Books, Reprint edition, p. 464 (cit. on pp. 5, 40).

Fitzmaurice, G. W. (1993). “Situated information spaces and spatially aware palmtop computers”. In: *Communications of the ACM* 36.7, pp. 39–49. DOI: `10.1145/159544.159566` (cit. on pp. 52, 82, 83, 111, 112, 138).

Frisch, M. Dachselt, R. (2013). “Visualizing Off-Screen Elements of Node-Link Diagrams”. In: 12(2), pp. 133–162 (cit. on p. 182).

Gallud, José A., Ricardo Tesoriero, and Victor M.R. Penichet, eds. (2011). *Distributed User Interfaces*. Human-Computer Interaction Series. London: Springer London. DOI: `10.1007/978-1-4471-2271-5` (cit. on p. 9).

Gebhardt, C. (2013). “Integrative Workplace: Employing Reality-based Interaction to join Digital and Analog Media at a Workplace”. Master Thesis. University of Konstanz (cit. on pp. xxv, 86–88).

Gebhardt, C. Rädle, R. Reiterer, H. (2014b). “Employing Blended Interaction to Blend the Qualities of Digital and Physical Books”. en. In: *Mensch & Computer 2014 - Tagungsband*. Ed. by Andreas Butz, Michael Koch, and Johann Schlichter. De Gruyter Oldenbourg, pp. 35–44 (cit. on p. xx).

Gebhardt, C. Rädle, R. Reiterer, H. (2014c). “Integrative workplace: studying the effect of digital desks on users’ working practices”. In: *Proceedings of the extended abstracts of the 32nd annual ACM conference on Human factors in computing systems - CHI EA ’14*. New York, New York, USA: ACM Press, pp. 2155–2160. DOI: `10.1145/2559206.2581186` (cit. on pp. xxi, 55, 59).

Geyer, F. (2013). “Interactive Spaces for Supporting Embodied Collaborative Design Practices”. In: *PhD Thesis* (cit. on pp. 38, 216).

Geyer, F. Pfeil, U. Budzinski, J. Höchtl, A. Reiterer, H. (2011a). “AffinityTable - A Hybrid Surface for Supporting Affinity Diagramming”. In: *Proceedings of 13th IFIP TC13 Conference on Human-Computer Interaction - INTERACT ’13*. Springer, pp. 477–484 (cit. on p. 76).

Geyer, F. Pfeil, U. Höchtl, A. Budzinski, J. Reiterer, H. (2011b). “Designing reality-based interfaces for creative group work”. In: *Proceedings of the 8th ACM conference on Creativity and cognition - C&C ’11*. New York, New York, USA: ACM Press, p. 165. DOI: `10.1145/2069618.2069647` (cit. on p. 76).

Geyer, F. Budzinski, J. Reiterer, H. (2012). “IdeaVis: A Hybrid Workspace and Interactive Visualization for Paper-based Collaborative Sketching Sessions”. In: *Proceedings of the 7th Nordic Conference on Human-Computer Interaction Making Sense Through Design - NordiCHI ’12*. New York, New York, USA: ACM Press, p. 331. DOI: `10.1145/2399016.2399069` (cit. on p. 76).

Gopnik, A. Meltzoff, A. N. (1997). *Words, Thoughts, and Theories*. MIT Press, p. 350 (cit. on p. 84).

Greenberg, S. Marquardt, N. Ballendat, T. Diaz-Marino, R. Wang, M. (2011). “Proxemic Interactions: The New Ubicomp?” In: *interactions* 18.1, p. 42. DOI: `10.1145/1897239.1897250` (cit. on pp. 8, 11, 27, 45, 162, 163, 170).

Greenberg, S. Boring, S. Vermeulen, J. Dostal, J. (2014a). “Dark patterns in proxemic interactions”. In: *Proceedings of the 2014 conference on Designing interactive systems - DIS ’14*. New York, New York, USA: ACM Press, pp. 523–532. DOI: `10.1145/2598510.2598541` (cit. on pp. 18, 26).

Greenberg, S. Honbaek, K. Quigley, A. Reiterer, H. Rädle, R. (2014b). *Proxemics in Human-Computer Interaction (Dagstuhl Seminar 13452)*. eng. Tech. rep. 11, p. 57. DOI: `10.4230/DagRep.3.11.29` (cit. on pp. xxiii, 17).

Gutwin, C. Greenberg, S. (1998). “Design for individuals, design for groups: tradeoffs between power and workspace awareness”. In: *Proceedings of the 1998 ACM conference on Computer supported cooperative work - CSCW ’98*. New York, New York, USA: ACM Press, pp. 207–216. DOI: `10.1145/289444.289495` (cit. on pp. 43, 44).

Haber, J. Nacenta, M. a. Carpendale, S. (2014). “Paper vs. tablets: the effect of document media in co-located collaborative work”. In: *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces - AVI ’14*. September 2010. New York, New York, USA: ACM Press, pp. 89–96. DOI: `10.1145/2598153.2598170` (cit. on p. 76).

Hajizadehgashti, S. (2012). *Investigating the Impact of Table Size on External Cognition in Collaborative Problem-Solving Tabletop Activities*. en (cit. on p. 45).

Hamilton, P. Wigdor, D. J. (2014). “Conductor: enabling and understanding cross-device interaction”. In: *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI ’14*. CHI ’14. New York, New York, USA: ACM Press, pp. 2773–2782. DOI: `10.1145/2556288.2557170` (cit. on pp. 3, 9, 11, 89, 163, 170, 171, 175, 179, 181, 195, 200).

Harboe, G. Huang, E. M. (2015). “Real-World Affinity Diagramming Practices: Bridging the Paper-Digital Gap”. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI ’15*. New York, New York, USA: ACM Press, pp. 95–104. DOI: `10.1145/2702123.2702561` (cit. on p. 76).

Harper, R. Rodden, T. Rogers, Y. Sellen, A. (2008). *Being human: Human-computer interaction in the year 2020*. Tech. rep. Cambridge, England: Microsoft Research Ltd., Cambridge (cit. on pp. 2, 3).

Harrison, C. Benko, H. Wilson, A. D. (2011). “OmniTouch: Wearable Multitouch Interaction Everywhere”. In: *Proceedings of the 24th annual ACM symposium on User interface software and technology - UIST ’11*. New York, New York, USA: ACM Press, p. 441. DOI: `10.1145/2047196.2047255` (cit. on p. 163).

Hart, S. Staveland, L. (1988). “Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research”. In: *Human mental workload*. Ed. by A. Hancock and N. Meshkati. Vol. 1. 3. Amsterdam: North Holland Press (cit. on pp. 97, 102, 186).

Hartson, R. Pyla, P. (2012). “The UX Book: Process and Guidelines for Ensuring a Quality User Experience”. In: (cit. on p. 61).

Head, A. J. Eisenberg, M. B. (2009). *Lessons Learned: How College Students Seek Information in the Digital Age* (cit. on p. 71).

Henze, N. Boll, S. (2010). “Evaluation of an off-screen visualization for magic lens and dynamic peephole interfaces”. In: *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services - MobileHCI ’10*. New York, New York, USA: ACM Press, p. 191. DOI: `10.1145/1851600.1851632` (cit. on pp. 84, 101, 111, 112).

Heydekorn, J. Frisch, M. Dachselt, R. (2011). “Evaluating a User-Elicited Gesture Set for Interactive Displays”. en. In: *Mensch & Computer 2011: überMEDIEN|ÜBERmorgen*. Oldenbourg Verlag, pp. 191–200 (cit. on p. 173).

Hilliges, O. Izadi, S. Wilson, A. D. Hodges, S. Garcia-Mendoza, A. Butz, A. (2009). “Interactions in the air: adding further depth to interactive tabletops”. In: *Proceedings of the 22nd annual ACM symposium on User interface software and technology - UIST ’09*. UIST ’09. New York, New York, USA: ACM Press, p. 139. DOI: `10.1145/1622176.1622203` (cit. on pp. 162, 165).

Hinckley, K. (2003). “Synchronous gestures for multiple persons and computers”. In: *Proceedings of the 16th annual ACM symposium on User interface software and technology - UIST ’03*. New York, New York, USA: ACM Press, pp. 149–158. DOI: `10.1145/964696.964713` (cit. on pp. 27, 171, 174).

Hinckley, K. Ramos, G. Guimbretiere, F. Baudisch, P. Smith, M. (2004). “Stitching: pen gestures that span multiple displays”. In: *Proceedings of the working conference on Advanced visual interfaces - AVI ’04*. New York, New York, USA: ACM Press, p. 23. DOI: `10.1145/989863.989866` (cit. on pp. 11, 27, 163, 171, 174).

Hollan, J. Hutchins, E. Kirsh, D. (2000). “Distributed cognition: toward a new foundation for human-computer interaction research”. In: *ACM Transactions on Computer-Human Interaction* 7.2, pp. 174–196. DOI: `10.1145/353485.353487` (cit. on p. 162).

Holmquist, L. E. Mattern, F. Schiele, B. Alahuhta, P. Beigl, M. Gellersen, H.-W. (2001). "Smart-Its Friends: A Technique for Users to Easily Establish Connections between Smart Artefacts". In: *Proceedings of Third International Conference on Ubiquitous Computing*. UbiComp '01. Springer, p. 116 (cit. on p. 163).

Houben, S. Marquardt, N. (2015). "WatchConnect". In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*. New York, New York, USA: ACM Press, pp. 1247–1256. DOI: `10.1145/2702123.2702215` (cit. on p. 28).

Huang, D.-Y. Lin, C.-P. Hung, Y.-P. Chang, T.-W. Yu, N.-H. Tsai, M.-L. Chen, M. Y. (2012). "MagMobile: enhancing social interactions with rapid view-stitching games of mobile devices". In: *Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia - MUM '12*. MUM '12. New York, New York, USA: ACM Press, p. 1. DOI: `10.1145/2406367.2406440` (cit. on p. 164).

Huber, J. Steimle, J. Mühlhäuser, M. (2011). "A model of embodied dynamic peephole pointing for hidden targets". In: *Proceedings of the 25th BCS Conference on Human-Computer Interaction - BCS-HCI '11*. Swinton, UK, UK: British Computer Society, pp. 315–320 (cit. on pp. 97, 112, 115).

Hutchins, E. L. Hollan, J. D. Norman, D. A. (1985). "Direct Manipulation Interfaces". In: *Human-Computer Interaction* 1.4, pp. 311–338. DOI: `10.1207/s15327051hci0104_2` (cit. on pp. 5, 35–37, 200).

Isenberg, P. Fisher, D. Paul, S. A. Morris, M. R. Inkpen, K. Czerwinski, M. (2012). "Co-Located Collaborative Visual Analytics around a Tabletop Display". In: *IEEE Transactions on Visualization and Computer Graphics* 18.5, pp. 689–702. DOI: `10.1109/TVCG.2011.287` (cit. on p. 76).

Ishii, H. Lakatos, D. Bonanni, L. Labrune, J.-B. (2012). "Radical atoms: beyond tangible bits, toward transformable materials". In: *interactions* 19.1, p. 38. DOI: `10.1145/2065327.2065337` (cit. on p. 36).

Jacob, R. J. Girouard, A. Hirshfield, L. M. Horn, M. S. Shaer, O. Solovey, E. T. Zigelbaum, J. (2008). "Reality-based interaction: a framework for post-WIMP interfaces". In: *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08*. New York, New York, USA: ACM Press, p. 201. DOI: `10.1145/1357054.1357089` (cit. on pp. 15, 26, 33, 37–39, 44, 46, 192, 207).

Jakobsen, M. Hornbæk, K. (2012). "Proximity and physical navigation in collaborative work with a multi-touch wall-display". In: *Proceedings of the 2012 ACM annual conference extended abstracts on Human Factors in Computing Systems Extended Abstracts - CHI EA '12*. New York, New York, USA: ACM Press, p. 2519. DOI: `10.1145/2212776.2223829` (cit. on p. 120).

Jeanne Sholl, M. (2001). "The Role of a Self-Reference System in Spatial Navigation". In: pp. 217–232. DOI: `10.1007/3-540-45424-1_15` (cit. on pp. 49, 84).

Jetter, H.-C. (2013). “Design and Implementation of Post-WIMP Interactive Spaces with the ZOIL Paradigm”. In: p. 285 (cit. on pp. 3, 26).

Jetter, H.-C. Rädle, R. (2013). “Visual and Functional Adaptation in Ad-hoc Communities of Devices”. In: *International Workshop on Visual Adaptation of Interfaces (in conjunction with ITS ’13)*. St Andrews, Scotland, UK (cit. on pp. xxiv, 11, 89).

Jetter, H.-C. Reiterer, H. (2013). “Self-Organizing User Interfaces: Envisioning the Future of Ubicomp UIs”. In: *Workshop Blended Interaction (in conjunction with CHI ’13)*. Konstanz: University of Konstanz (cit. on pp. 9, 11).

Jetter, H.-C. Gerken, J. Zöllner, M. Reiterer, H. Milic-Frayling, N. (2011). “Materializing the Query with Facet-Streams – A Hybrid Surface for Collaborative Search on Tabletops”. In: *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI ’11*. New York, New York, USA: ACM Press, p. 3013. DOI: `10.1145/1978942.1979390` (cit. on pp. 46, 59, 76).

Jetter, H.-c. Geyer, F. Schwarz, T. Reiterer, H. (2012a). “Blended Interaction – Toward a Framework for the Design of Interactive Spaces”. In: *In proceedings of Workshop on Designing Collaborative Interactive Spaces AVI2012* (cit. on p. 44).

Jetter, H.-C. Zöllner, M. Gerken, J. Reiterer, H. (2012b). “Design and Implementation of Post-WIMP Distributed User Interfaces with ZOIL”. In: *International Journal of Human-Computer Interaction* 28.11, pp. 737–747. DOI: `10.1080/10447318.2012.715539` (cit. on p. 125).

Jetter, H.-C. Leifert, S. Gerken, J. Schubert, S. Reiterer, H. (2012c). “Does (multi-)touch aid users’ spatial memory and navigation in ’panning’ and in ’zooming & panning’ UIs?” In: *Proceedings of the International Working Conference on Advanced Visual Interfaces - AVI ’12*. New York, New York, USA: ACM Press, p. 83. DOI: `10.1145/2254556.2254575` (cit. on pp. 43, 50, 51, 129, 136).

Jetter, H.-C. Reiterer, H. Geyer, F. (2014). “Blended Interaction: understanding natural human–computer interaction in post-WIMP interactive spaces”. In: *Personal and Ubiquitous Computing* 18.5, pp. 1139–1158. DOI: `10.1007/s00779-013-0725-4` (cit. on pp. 3, 5, 35, 39–46, 81, 192, 198, 202, 210).

Jokela, T. Ojala, J. Olsson, T. (2015a). “A Diary Study on Combining Multiple Information Devices in Everyday Activities and Tasks”. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI ’15*. New York, New York, USA: ACM Press, pp. 3903–3912. DOI: `10.1145/2702123.2702211` (cit. on pp. 7, 9, 34, 89, 199).

Jokela, T. Chong, M. K. Lucero, A. Gellersen, H. (2015b). “Connecting devices for collaborative interactions”. In: *interactions* 22.4, pp. 39–43. DOI: `10.1145/2776887` (cit. on p. 199).

Jordà, S. Kaltenbrunner, M. Geiger, G. Alonso, M. (2006). “The reacTable: a tangible tabletop musical instrument and collaborative workbench”. In: *ACM SIGGRAPH 2006 Sketches on - SIGGRAPH ’06*. New York, New York, USA: ACM Press, p. 91. DOI: `10.1145/1179849.1179963` (cit. on p. 89).

Jörn Hurtienne, L. B. (2007). “Design for Intuitive Use - Testing image schema theory for user interface design”. In: *16 th International Conference on Engineering Design* (cit. on p. 5).

Jul, S. Furnas, G. W. (1998). “Critical zones in desert fog: aids to multiscale navigation”. In: *Proceedings of the 11th annual ACM symposium on User interface software and technology - UIST ’98*. New York, New York, USA: ACM Press, pp. 97–106. DOI: `10.1145/288392.288578` (cit. on p. 121).

Kaltenbrunner, M. Bencina, R. (2007). “reacTIVision: a computer-vision framework for table-based tangible interaction”. In: *Proceedings of the 1st international conference on Tangible and embedded interaction - TEI ’07*. New York, New York, USA: ACM Press, p. 69. DOI: `10.1145/1226969.1226983` (cit. on p. 89).

Kaufmann, B. Ahlström, D. (2012). “Revisiting Peephole Pointing: A Study of Target Acquisition with a Handheld Projector”. In: *Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services - MobileHCI ’12*. New York, New York, USA: ACM Press, p. 211. DOI: `10.1145/2371574.2371607` (cit. on pp. 83, 84, 97, 111, 112, 115).

Kaufmann, B. Ahlström, D. (2013). “Studying spatial memory and map navigation performance on projector phones with peephole interaction”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI ’13*. New York, New York, USA: ACM Press, p. 3173. DOI: `10.1145/2470654.2466434` (cit. on pp. 51, 52, 83, 84, 109, 111–113).

Keim, D. (2002). “Information visualization and visual data mining”. In: *IEEE Transactions on Visualization and Computer Graphics* 8.1, pp. 1–8. DOI: `10.1109/2945.981847` (cit. on p. 79).

Kendon, A. (2010). “Spacing and orientation in co-present interaction”. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 5967 LNCS, pp. 1–15. DOI: `10.1007/978-3-642-12397-9-1` (cit. on p. 163).

Kidd, A. (1994). “The marks are on the knowledge worker”. In: *Proceedings of the SIGCHI conference on Human factors in computing systems celebrating interdependence - CHI ’94*. New York, New York, USA: ACM Press, pp. 186–191. DOI: `10.1145/191666.191740` (cit. on pp. 76, 85, 172).

Kirsh, D. (2010). “Thinking with external representations”. In: *AI & Society* 25.4, pp. 441–454. DOI: `10.1007/s00146-010-0272-8` (cit. on pp. 36, 76, 85, 196, 198, 210).

Kirsh, D. (2011). “Thinking with the body”. In: *Proceedings of the 9th ACM SIGCHI Italian Chapter International Conference on Computer-Human Interaction Facing Complexity - CHItaly*, p. 15. DOI: `10.1145/2037296.2037303` (cit. on p. 198).

Kleiner, E. (2013). “Blended Shelf: Ein realitätsbasierter Ansatz zur Präsentation und Exploration von Bibliotheksbeständen”. Master Thesis. University of Konstanz (cit. on pp. xxv, 80–82).

Kleiner, E. Rädle, R. Reiterer, H. (2013). “Blended Shelf: Reality-based Presentation and Exploration of Library Collections”. In: *CHI ’13 Extended Abstracts on Human Factors in Computing Systems on - CHI EA ’13*. New York, New York, USA: ACM Press, p. 577. DOI: `10.1145/2468356.2468458` (cit. on pp. xxii, 55, 79, 80).

Klemmer, S. R. Newman, M. W. Farrell, R. Bilezikjian, M. Landay, J. A. (2001). “The Designers’ Outpost: A Tangible Interface for Collaborative Web Site Design”. In: *Proceedings of the 14th annual ACM symposium on User interface software and technology - UIST ’01*. New York, New York, USA: ACM Press, p. 1. DOI: `10.1145/502348.502350` (cit. on p. 76).

Klemmer, S. R. Hartmann, B. Takayama, L. (2006). “How bodies matter: five themes for interaction design”. In: *Proceedings of the 6th ACM conference on Designing Interactive systems - DIS ’06*. New York, New York, USA: ACM Press, p. 140. DOI: `10.1145/1142405.1142429` (cit. on p. 36).

Klinkhammer, D. Nitsche, M. Specht, M. Reiterer, H. (2011). “Adaptive personal territories for co-located tabletop interaction in a museum setting”. In: *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces - ITS ’11*. New York, New York, USA: ACM Press, p. 107. DOI: `10.1145/2076354.2076375` (cit. on p. 216).

Klokmose, C. N. Eagan, J. R. Baader, S. Mackay, W. Beaudouin-Lafon, M. (2015). “Webstrates: Shareable Dynamic Media”. In: *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology - UIST ’15*. New York, New York, USA: ACM Press, pp. 280–290. DOI: `10.1145/2807442.2807446` (cit. on p. 27).

Klokmose, C. N. (2007). “An instrumental paradigm for ubiquitous interaction”. In: *Proceedings of the 14th European conference on Cognitive ergonomics invent! explore! - ECCE ’07*. New York, New York, USA: ACM Press, p. 257. DOI: `10.1145/1362550.1362603` (cit. on p. 3).

Klokmose, C. N. Beaudouin-Lafon, M. (2009). “VIGO: Instrumental Interaction in Multi-Surface Environments”. In: *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09*. New York, New York, USA: ACM Press, p. 869. DOI: `10.1145/1518701.1518833` (cit. on p. 3).

Klum, S. Isenberg, P. Langner, R. Fekete, J.-D. Dachselt, R. (2012). “Stackables: Combining Tangibles for Faceted Browsing”. In: *Proceedings of the International Working Conference on Advanced Visual Interfaces - AVI ’12*. New York, New York, USA: ACM Press, p. 241. DOI: `10.1145/2254556.2254600` (cit. on p. 76).

Koike, H. Sato, Y. Kobayashi, Y. Tobita, H. Kobayashi, M. (2000). “Interactive textbook and interactive Venn diagram”. In: *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI ’00*. New York, New York, USA: ACM Press, pp. 121–128. DOI: `10.1145/332040.332415` (cit. on p. 85).

Koike, H. Sato, Y. Kobayashi, Y. (2001). “Integrating paper and digital information on EnhancedDesk: a method for realtime finger tracking on an augmented desk system”. In: *ACM Transactions on Computer-Human Interaction* 8.4, pp. 307–322. DOI: `10.1145/504704.504706` (cit. on p. 85).

Kray, C. Rohs, M. Hook, J. Kratz, S. (2008). “Group Coordination and Negotiation through Spatial Proximity Regions around Mobile Devices on Augmented Tabletops”. In: *3rd IEEE International Workshop on Horizontal Interactive Human Computer Systems*. TABLETOP ’08. IEEE, pp. 1–8 (cit. on p. 164).

Kray, C. Nesbitt, D. Dawson, J. Rohs, M. (2010). “User-defined gestures for connecting mobile phones, public displays, and tabletops”. In: *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services - MobileHCI ’10*. New York, New York, USA: ACM Press, p. 239. DOI: `10.1145/1851600.1851640` (cit. on pp. 172, 174, 175).

Lakoff, G. Johnson, M. (1999). *Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought*, p. 640 (cit. on p. 36).

Lazar, J. Feng, J. H. Hochheiser, H. (2010). “Research Methods in Human-Computer Interaction”. In: (cit. on p. 61).

Leifert, S. (2013). “Theorie und experimentelle Untersuchung des Einflusses von Interaktionsdesign auf das räumliche und inhaltliche Gedächtnis”. mastersthesis. University of Konstanz (cit. on p. 129).

Li, M. Kobbelt, L. (2012). “Dynamic tiling display: building an interactive display surface using multiple mobile devices”. In: *Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia - MUM ’12*. New York, New York, USA: ACM Press, p. 1. DOI: `10.1145/2406367.2406397` (cit. on pp. 9, 146, 164).

Liao, C. Tang, H. Liu, Q. Chiu, P. Chen, F. (2010). “FACT: fine-grained cross-media interaction with documents via a portable hybrid paper-laptop interface”. In: *Proceedings of the international conference on Multimedia - MM ’10*. New York, New York, USA: ACM Press, p. 361. DOI: `10.1145/1873951.1874001` (cit. on pp. 85, 86).

Linder, N. Maes, P. (2012). “The Design Evolution of LuminAR: A Compact and Kinetic Projected Augmented Reality Interface”. In: *Proceedings of the 2012 ACM annual conference extended abstracts on Human Factors in Computing Systems Extended Abstracts - CHI EA ’12*. New York, New York, USA: ACM Press, p. 1435. DOI: `10.1145/2212776.2212473` (cit. on p. 163).

Löchtefeld, M. Gehring, S. Krüger, A. (2011). "Distributed User Interfaces for Projector Phones". In: *Distributed User Interfaces: Designing Interfaces for the Distributed Ecosystem*. Ed. by José A. Gallud, Ricardo Tesoriero, and Victor M.R. Penichet. Human-Computer Interaction Series. London: Springer London. Chap. 13, pp. 113–123. DOI: `10.1007/978-1-4471-2271-5` (cit. on pp. 83, 84, 111).

Lu, Z. (2014). "User-defined Gesture for Cross-device Interaction". Unpublished Master Thesis. University College London, UK (cit. on p. 167).

Lucero, A. Keränen, J. Korhonen, H. (2010). "Collaborative use of mobile phones for brainstorming". In: *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services - MobileHCI '10*. New York, New York, USA: ACM Press, p. 337. DOI: `10.1145/1851600.1851659` (cit. on pp. 9, 11, 146, 171).

Lucero, A. Holopainen, J. Jokela, T. (2011). "Pass-them-around: collaborative use of mobile phones for photo sharing". In: *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11*. Vol. 54. New York, New York, USA: ACM Press, p. 1787. DOI: `10.1145/1978942.1979201` (cit. on pp. 9, 11, 163, 164, 171, 174).

Luger, E. (2012). "Consent reconsidered; reframing consent for ubiquitous computing systems". In: *Proceedings of the 2012 ACM Conference on Ubiquitous Computing - UbiComp '12*. New York, New York, USA: ACM Press, p. 564. DOI: `10.1145/2370216.2370310` (cit. on p. 27).

Luger, E. Rodden, T. (2013). "An informed view on consent for UbiComp". In: *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing - UbiComp '13*. New York, New York, USA: ACM Press, p. 529. DOI: `10.1145/2493432.2493446` (cit. on p. 27).

Marquardt, N. Jota, R. Greenberg, S. Jorge, J. A. (2011a). "The continuous interaction space: interaction techniques unifying touch and gesture on and above a digital surface". In: *In Proceedings of the 13th IFIP TC 13 international conference on Human-computer interaction - Volume Part III (INTERACT'11)*. Berlin, Heidelberg: Springer-Verlag, pp. 461–476 (cit. on p. 165).

Marquardt, N. Diaz-Marino, R. Boring, S. Greenberg, S. (2011b). "The proximity toolkit: prototyping proxemic interactions in ubiquitous computing ecologies". In: *Proceedings of the 24th annual ACM symposium on User interface software and technology - UIST '11*. New York, New York, USA: ACM Press, p. 315. DOI: `10.1145/2047196.2047238` (cit. on pp. 11, 27, 125, 146, 163).

Marquardt, N. Hinckley, K. Greenberg, S. (2012a). "Cross-device interaction via micro-mobility and f-formations". In: *Proceedings of the 25th annual ACM symposium on User interface software and technology - UIST '12*. New York, New York, USA: ACM Press, p. 13. DOI: `10.1145/2380116.2380121` (cit. on pp. 146, 163, 164, 175, 195).

Marquardt, N. Ballendat, T. Boring, S. Greenberg, S. Hinckley, K. (2012b). "Gradual engagement: facilitating information exchange between digital devices as a function of proximity". In: *Proceedings of the 2012 ACM international conference on Interactive tabletops and surfaces - ITS '12*. New York, NY, USA: ACM, p. 31. DOI: `10.1145/2396636.2396642` (cit. on p. 27).

Marwecki, S. Rädle, R. Reiterer, H. (2013). "Encouraging collaboration in hybrid therapy games for autistic children". In: *CHI '13 Extended Abstracts on Human Factors in Computing Systems on - CHI EA '13*. New York, New York, USA: ACM Press, p. 469. DOI: `10.1145/2468356.2468439` (cit. on pp. xxii, 216).

Mayhew, D. J. (1999). *The Usability Engineering Lifecycle: A Practitioner's Handbook for User Interface Design*. San Francisco, CA: Morgan Kaufmann (cit. on p. 61).

McNamara, T. P. Valiquette, C. M. (2004). "Remembering Where Things Are". In: *Human Spatial Memory: Remembering Where*. Ed. by Garry L. Allen. Mahwah, New Jersey, USA: Lawrence Erlbaum Associates, p. 366 (cit. on p. 49).

Mehra, S. Werkhoven, P. Worring, M. (2006). "Navigating on Handheld Displays: Dynamic versus Static Peephole Navigation". In: *ACM Transactions on Computer-Human Interaction* 13.4, pp. 448–457. DOI: `10.1145/1188816.1188818` (cit. on pp. 82, 83, 111).

Merrill, D. Kalanithi, J. Maes, P. (2007). "Siftables: towards sensor network user interfaces". In: *Proceedings of the 1st international conference on Tangible and embedded interaction - TEI '07*. TEI '07. New York, New York, USA: ACM Press, p. 75. DOI: `10.1145/1226969.1226984` (cit. on pp. 146, 164).

Morris, M. R. Lombardo, J. Wigdor, D. (2010). "WeSearch: Supporting Collaborative Search and Sensemaking on a Tabletop Display". In: *Proceedings of the 2010 ACM conference on Computer supported cooperative work - CSCW '10*. New York, New York, USA: ACM Press, pp. 401–410. DOI: `10.1145/1718918.1718987` (cit. on p. 76).

Morris, M. R. Danielescu, A. Drucker, S. Fisher, D. Lee, B. Schraefel, C. Wobbrock, J. O. (2014). "Reducing Legacy Bias in Gesture Elicitation Studies". In: *interactions* 21.3, pp. 40–45. DOI: `10.1145/2591689` (cit. on pp. 17, 174, 175, 179).

Morrison, A. Oulasvirta, A. Peltonen, P. Lemmela, S. Jacucci, G. Reitmayr, G. Näsänen, J. Juustila, A. (2009). "Like bees around the hive: a comparative study of a mobile augmented reality map". In: *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09*. New York, New York, USA: ACM Press, p. 1889. DOI: `10.1145/1518701.1518991` (cit. on pp. 83, 84, 101, 111).

Mueller, F. Stellmach, S. Greenberg, S. Dippon, A. Boll, S. Garner, J. Khot, R. Naseem, A. Altimira, D. (2014). "Proxemics play: understanding proxemics for designing digital play experiences". In: *Proceedings of the 2014 conference on Designing interactive systems - DIS '14*. New York, New York, USA: ACM Press, pp. 533–542. DOI: `10.1145/2598510.2598532` (cit. on p. 18).

Müller, J. Rädle, R. Jetter, H.-C. Reiterer, H. (2015). “An Experimental Comparison of Vertical and Horizontal Dynamic Peephole Navigation”. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI ’15*. New York, New York, USA: ACM Press, pp. 1523–1526. DOI: `10.1145/2702123.2702227` (cit. on pp. xx, 195).

Müller, J. Rädle, R. Reiterer, H. (2016). “Virtual Objects as Spatial Cues in Collaborative Mixed Reality Environments: How They Shape Communication Behavior and User Task Load”. In: *Proceedings of the 34th Annual ACM Conference on Human Factors in Computing Systems - CHI ’16*. New York, NY, USA: ACM Press (cit. on pp. xix, 215).

Murdock Jr. B. B. (1962). “The serial position effect of free recall”. In: *Journal of Experimental Psychology* 64.5, pp. 482–188 (cit. on p. 132).

Nacenta, M. Gutwin, C. Aliakseyeu, D. Subramanian, S. (2009). “There and Back Again: Cross-Display Object Movement in Multi-Display Environments”. In: *Human-Computer Interaction* 24.1, pp. 170–229. DOI: `10.1080/07370020902819882` (cit. on pp. 171, 174, 179).

Nacenta, M. A. Kamber, Y. Qiang, Y. Kristensson, P. O. (2013). “Memorability of predesigned and user-defined gesture sets”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI ’13*. New York, New York, USA: ACM Press, p. 1099. DOI: `10.1145/2470654.2466142` (cit. on pp. 9, 172, 174).

Nebeling, M. Teunissen, E. Husmann, M. Norrie, M. C. (2014). “XDKinect”. In: *Proceedings of the 2014 ACM SIGCHI symposium on Engineering interactive computing systems - EICS ’14*. New York, New York, USA: ACM Press, pp. 65–74. DOI: `10.1145/2607023.2607024` (cit. on p. 28).

Nielsen, M. Störring, M. Moeslund, T. B. Granum, E. (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. Vol. 2915. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 409–420. DOI: `10.1007/978-3-540-24598-8_38` (cit. on pp. 172, 173, 183).

O’Hara, K. Smith, F. Newman, W. Sellen, A. (1998). “Student readers’ use of library documents: implications for library technologies”. In: *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI ’98*. April. New York, New York, USA: ACM Press, pp. 233–240. DOI: `10.1145/274644.274678` (cit. on pp. 63, 64).

Ohlsson, S. (1995). “Learning to Do and Learning to Understand: A Lesson and a Challenge for Cognitive Modeling”. In: *Learning in humans and machines: Towards an interdisciplinary learning science,* pp. 37–62 (cit. on pp. 26, 47, 74, 75, 77, 199).

Ohta, T. Tanaka, J. (2012). “Pinch: An Interface That Relates Applications on Multiple Touch-Screen by ‘Pinching’ Gesture”. In: *Advances in Computer Entertainment*. ACE’12. Berlin: Springer Berlin Heidelberg, pp. 320–335. DOI: `10.1007/978-3-642-34292-9_23` (cit. on p. 163).

Olberding, S. Wessely, M. Steimle, J. (2014). “PrintScreen: Fabricating Highly Customizable Thin-film Touch-Displays”. In: *Proceedings of the 27th annual ACM symposium on User interface software and technology - UIST ’14*. New York, New York, USA: ACM Press, pp. 281–290. DOI: `10.1145/2642918.2647413` (cit. on pp. 32, 33).

Oudgenoeg-Paz, O. Leseman, P. P. M. Volman, M. C. J. M. (2014). “Can infant self-locomotion and spatial exploration predict spatial memory at school age?” In: *European Journal of Developmental Psychology* 11.1, pp. 36–48. DOI: `10.1080/17405629.2013.803470` (cit. on p. 49).

Oulasvirta, A. (2008). “When users "do" the Ubicomp”. In: *interactions* 15.2, p. 6. DOI: `10.1145/1340961.1340963` (cit. on pp. 8, 195).

Oulasvirta, A. Sumari, L. (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*. New York, New York, USA: ACM Press, p. 1127. DOI: `10.1145/1240624.1240795` (cit. on p. 3).

Oviatt, S. (2006). “Human-centered design meets cognitive load theory”. In: *Proceedings of the 14th annual ACM international conference on Multimedia - MULTIMEDIA ’06*. New York, New York, USA: ACM Press, p. 871. DOI: `10.1145/1180639.1180831` (cit. on p. 26).

Pahud, M. Hinckley, K. Iqbal, S. Sellen, A. Buxton, B. (2013). “Toward compound navigation tasks on mobiles via spatial manipulation”. In: *Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services - MobileHCI ’13*. New York, New York, USA: ACM Press, p. 113. DOI: `10.1145/2493190.2493210` (cit. on pp. 83, 84, 98, 110–113, 115).

Palfrey, J. (2015). *BiblioTech: why libraries matter more than ever in the age of Google*. New York, NY: Basic Books (cit. on p. 70).

Piazza, T. Zhao, S. Ramos, G. Yantaç, A. E. Fjeld, M. (2013a). “Dynamic Duo: Exploring Phone-Tablet, Combinations for Mobile Usage.” In: *Conference on Tangible Embedded and Embodied Interaction (TEI ’13) - Work in Progress*. New York: ACM (cit. on p. 164).

Piazza, T. Zhao, S. Ramos, G. Yantaç, A. E. Fjeld, M. (2013b). “Dynamic duo: phone-tablet interaction on tabletops”. In: *CHI ’13 Extended Abstracts on Human Factors in Computing Systems on - CHI EA ’13*. New York, New York, USA: ACM Press, p. 2803. DOI: `10.1145/2468356.2479520` (cit. on p. 164).

Pierce, J. S. Mahaney, H. E. Abowd, G. D. (2003). *Opportunistic Annexing for Handheld Devices: Opportunities and Challenges*. Tech. rep. (cit. on p. 11).

Preim, B. Dachselt, R. (2010). *Interaktive Systeme*. eXamen.press. Berlin, Heidelberg: Springer Berlin Heidelberg. DOI: `10.1007/978-3-642-05402-0` (cit. on p. 16).

Rädle, R. (2013). “Design and evaluation of proxemics-aware environments to support navigation in large information spaces”. In: *CHI ’13 Extended Abstracts on Human Factors in Computing Systems on - CHI EA ’13*. New York, NY, USA: ACM, p. 1957. DOI: `10.1145/2468356.2468710` (cit. on pp. xxi, 23, 55).

Rädle, R. Heilig, M. Reiterer, H. (2011b). “Interactive Reading: Serendipity in the Context of the Blended Library”. In: *Proceedings of 1st Workshop on Encouraging Serendipity in Interactive Systems, 13th IFIP TC13 Conference on Human-Computer Interaction*. Ed. by Stephann Makri, Elaine .G. Toms, Lori McCay-Peet, and Ann Blandford. Springer Verlag, pp. 24–26 (cit. on pp. xxiv, 59).

Rädle, R. Weiler, A. Huber, S. Jetter, H.-C. Mansmann, S. Reiterer, H. Scholl, M. H. (2012a). “eBook meets tabletop: using collaborative visualization for search and serendipity in on-line book repositories”. In: *Proceedings of the fifth ACM workshop on Research advances in large digital book repositories and complementary media - BooksOnline ’12*. New York, New York, USA: ACM Press, p. 3. DOI: `10.1145/2390116.2390120` (cit. on pp. xxiv, 59, 216).

Rädle, R. Jetter, H.-C. Reiterer, H. (2012d). “TwisterSearch: Supporting Social Search with Tabletop and Mobile Displays”. In: *Mensch & Computer 2012 – Workshopband: interaktiv informiert – allgegenwärtig und allumfassend!?* Ed. by Harald Reiterer and Oliver Deussen. 2009. München: Oldenbourg Verlag, pp. 509–512 (cit. on p. xxii).

Rädle, R. Jetter, H.-C. Butscher, S. Reiterer, H. (2013a). “The effect of egocentric body movements on users’ navigation performance and spatial memory in zoomable user interfaces”. In: *Proceedings of the 2013 ACM international conference on Interactive tabletops and surfaces - ITS ’13*. New York, New York, USA: ACM Press, pp. 23–32. DOI: `10.1145/2512349.2512811` (cit. on pp. xxi, 15, 17, 23, 111, 117).

Rädle, R. Jetter, H.-C. Reiterer, H. (2013b). “TwisterSearch: A distributed user interface for collaborative Web search”. In: *Distributed User Interfaces: Collaboration and Usability*. Ed. by María D. Lozano, Jose A. Gallud, Ricardo Tesoriero, and Victor M.R. Penichet. London, UK: Springer London. Chap. 5, pp. 4–7. DOI: `10.1007/978-1-4471-5499-0` (cit. on pp. xxi, 59, 76, 202, 216).

Rädle, R. Jetter, H.-C. Müller, J. Reiterer, H. (2014a). “Bigger is not always better: display size, performance, and task load during peephole map navigation”. In: *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI ’14*. New York, New York, USA: ACM Press, pp. 4127–4136. DOI: `10.1145/2556288.2557071` (cit. on pp. xx, 16, 17, 93).

Rädle, R. Jetter, H.-C. Marquardt, N. Reiterer, H. Rogers, Y. (2014b). “Demonstrating HuddleLamp: Spatially- Aware Mobile Displays for Ad-hoc Around-the-Table Collaboration”. In: *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces - ITS ’14*. New York, New York, USA: ACM Press, pp. 435–438. DOI: `10.1145/2669485.2676584` (cit. on pp. xxii, 141).

Rädle, R. Jetter, H.-C. Marquardt, N. Reiterer, H. Rogers, Y. (2014c). “HuddleLamp: Spatially-Aware Mobile Displays for Ad-hoc Around-the-Table Collaboration”. In: *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces - ITS ’14*. New York, New York, USA: ACM Press, pp. 45–54. DOI: `10.1145/2669485.2669500` (cit. on pp. xx, 17, 141).

Rädle, R. Jetter, H.-C. Schreiner, M. Lu, Z. Reiterer, H. Rogers, Y. (2015). “Spatially-aware or Spatially-agnostic?: Elicitation and Evaluation of User-Defined Cross-Device Interactions”. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI ’15*. New York, New York, USA: ACM Press, pp. 3913–3922. DOI: `10.1145/2702123.2702287` (cit. on pp. xx, 16, 17, 167).

Ramos, G. Hinckley, K. Wilson, A. Sarin, R. (2009). “Synchronous Gestures in Multi-Display Environments”. In: *Human–Computer Interaction* 24.1-2, pp. 117–169. DOI: `10.1080/07370020902739288` (cit. on p. 163).

Rapp, S. Michelitsch, G. Osen, M. Williams, J. Barbisch, M. Bohan, R. Valsan, Z. Emele, M. (2004). “Spotlight Navigation: Interaction with a handheld projection device (Video paper)”. In: *Proceedings of 3rd International Conference on Pervasive Computing - Pervasive ’05*. Vienna, Austria (cit. on pp. 83, 111).

Rauber, A. Bina, H. (1999). “A metaphor graphics based representation of digital libraries on the World Wide Web: using the libViewer to make metadata visible”. In: *Proceedings. Tenth International Workshop on Database and Expert Systems Applications. DEXA 99*. IEEE, pp. 286–290. DOI: `10.1109/DEXA.1999.795181` (cit. on p. 79).

Reetz, A. Gutwin, C. Stach, T. Nacenta, M. Subramanian, S. (2006). “Superflick: a natural and efficient technique for long-distance object placement on digital tables”. In: *Proceedings of Graphics Interface 2006 - GI ’06*. Toronto, Ontario, Canada: Canadian Information Processing Society, pp. 163–170 (cit. on pp. 172, 174, 180, 182).

Reiterer, H. Rädle, R. Butscher, S. Müller, J. (2016). “Blended Library – neue Zugangswege zu den Inhalten wissenschaftlicher und öffentlicher Bibliotheken”. In: *Bibliothek Forschung und Praxis* 40.1. DOI: `10.1515/bfp-2016-0010` (cit. on pp. xix, 55).

Rekimoto, J. (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` (cit. on p. 174).

Rekimoto, J. Saitoh, M. (1999). “Augmented surfaces: a spatially continuous work space for hybrid computing environments”. In: *Proceedings of the SIGCHI conference on Human factors in computing systems the CHI is the limit - CHI ’99*. CHI ’99. New York, New York, USA: ACM Press, pp. 378–385. DOI: `10.1145/302979.303113` (cit. on pp. 147, 164, 171, 172).

Robertson, G. Czerwinski, M. Larson, K. Robbins, D. C. Thiel, D. Dantzich, M. (1998). “Data Mountain: Using Spatial Memory for Document Management”. In: *Proceedings of the 11th annual ACM symposium on User interface software and technology - UIST ’98*. New York, New York, USA: ACM Press, pp. 153–162. DOI: `10.1145/288392.288596` (cit. on pp. 5, 50).

Rogers, Y. (2005). “New theoretical approaches for human-computer interaction”. In: *Annual Review of Information Science and Technology* 38.1, pp. 87–143. DOI: `10.1002/aris.1440380103` (cit. on p. 75).

Rogers, Y. (2006). “Moving on from Weiser’s Vision of Calm Computing: engaging UbiComp experiences”. In: Lecture Notes in Computer Science 4206. Ed. by Paul Dourish and Adrian Friday, pp. 404–421. DOI: `10.1007/11853565_24` (cit. on pp. 26, 31, 33, 35, 192).

Rogers, Y. (2012). “HCI Theory: Classical, Modern, and Contemporary”. In: *Synthesis Lectures on Human-Centered Informatics* 5.2, pp. 1–129. DOI: `10.2200/S00418ED1V01Y201205HCI014` (cit. on p. 77).

Rogers, Y. Sharp, H. Preece, J. (2011). “Interaction Design: Beyond Human - Computer Interaction”. In: (cit. on pp. 61, 67, 69).

Rohs, M. Essl, G. (2006). “Which one is better?: information navigation techniques for spatially aware handheld displays”. In: *Proceedings of the 8th international conference on Multimodal interfaces - ICMI ’06*. New York, New York, USA: ACM Press, p. 100. DOI: `10.1145/1180995.1181016` (cit. on pp. 83, 101, 112, 113).

Rohs, M. Oulasvirta, A. (2008). “Target acquisition with camera phones when used as magic lenses”. In: *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI ’08*. New York, New York, USA: ACM Press, pp. 1409–1418. DOI: `10.1145/1357054.1357275` (cit. on pp. 113, 115).

Rohs, M. Schöning, J. Raubal, M. Essl, G. Krüger, A. (2007). “Map Navigation with Mobile Devices: Virtual versus Physical Movement with and without Visual Context”. In: *Proceedings of the ninth international conference on Multimodal interfaces - ICMI ’07*. New York, New York, USA: ACM Press, p. 146. DOI: `10.1145/1322192.1322219` (cit. on pp. 83, 84, 101, 111, 113).

Rosson, M. B. Carroll, J. M. (2002). *Usability Engineering: Scenario-based Development of Human-Computer Interaction*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc. (cit. on p. 61).

Rukzio, E. Holleis, P. Gellersen, H. (2012). “Personal Projectors for Pervasive Computing”. In: *IEEE Pervasive Computing* 11.2, pp. 30–37. DOI: `10.1109/MPRV.2011.17` (cit. on pp. 83, 84, 111).

Santosa, S. Wigdor, D. (2013). “A field study of multi-device workflows in distributed workspaces”. In: *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing - UbiComp ’13*. New York, New York, USA: ACM Press, p. 63. DOI: `10.1145/2493432.2493476` (cit. on p. 195).

Scarr, J. Cockburn, A. Gutwin, C. (2013). “Supporting and Exploiting Spatial Memory in User Interfaces”. English. In: *Foundations and Trends® in Human–Computer Interaction* 6.1, pp. 1–84. DOI: `10.1561/1100000046` (cit. on pp. 5, 49, 50).

Scharf, F. Wolters, C. Herczeg, M. Cassens, J. (2013). “Cross-Device Interaction : Definition, Taxonomy and Application”. In: *AMBIENT 2013: The Third International Conference on Ambient Computing, Applications, Services and Technologies*. Ed. by Maarten Weyn. IARIA. Porto, Portugal: IARIA, pp. 35–41 (cit. on p. 7).

Schmalstieg, D. Fuhrmann, A. Hesina, G. Szalavári, Z. Encarnação, L. M. Gervautz, M. Purgathofer, W. (2002). “The Studierstube Augmented Reality Project”. In: *Presence: Teleoperators and Virtual Environments* 11.1, pp. 33–54. DOI: `10.1162/105474602317343640` (cit. on p. 214).

Schmitz, A. Li, M. Schönefeld, V. Kobbelt, L. (2010). “Ad-Hoc Multi-Displays for Mobile Interactive Applications”. Englisch. In: *The Eurographics Association*. The Eurographics Association, pp. 45–52 (cit. on pp. 164, 195).

Schuler, D. (1994). “Social computing”. In: *Communications of the ACM* 37.1, pp. 28–29. DOI: `10.1145/175222.175223` (cit. on pp. 34, 36).

Schwarz, J. Klionsky, D. Harrison, C. Dietz, P. Wilson, A. (2012). “Phone as a Pixel: Enabling Ad-Hoc, Large-Scale Displays Using Mobile Devices”. In: *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI ’12*. New York, New York, USA: ACM Press, p. 2235. DOI: `10.1145/2207676.2208378` (cit. on pp. 157, 164).

Scott, C. L. Harris, R. J. Rothe, A. R. (2001). “Embodied Cognition Through Improvisation Improves Memory for a Dramatic Monologue”. en. In: *Discourse Processes* 31.3, pp. 293–305. DOI: `10.1207/S15326950dp31-3_4` (cit. on pp. 36, 46).

Sellen, A. J. Harper, R. H. R. (2001). *The Myth of the Paperless Office*. 1st. MIT Press, p. 245 (cit. on pp. 45, 76, 85).

Seyed, T. Burns, C. Costa Sousa, M. Maurer, F. Tang, A. (2012). “Eliciting usable gestures for multi-display environments”. In: *Proceedings of the 2012 ACM international conference on Interactive tabletops and surfaces - ITS ’12*. New York, New York, USA: ACM Press, p. 41. DOI: `10.1145/2396636.2396643` (cit. on pp. 172–175).

Shneiderman, B. (1983). “Direct Manipulation: A Step Beyond Programming Languages”. In: *Computer* 16.8, pp. 57–69. DOI: `10.1109/MC.1983.1654471` (cit. on p. 35).

Sholl, M. J. Nolin, T. L. (1997). “Orientation specificity in representations of place”. In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 23.6, pp. 1507–. DOI: `10.1037//0278-7393.23.6.1494` (cit. on pp. 49, 52).

Spindler, M. Stellmach, S. Dachselt, R. (2009). “PaperLens: Advanced Magic Lens Interaction Above the Tabletop”. In: *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces - ITS ’09*. New York, New York, USA: ACM Press, p. 69. DOI: `10.1145/1731903.1731920` (cit. on pp. 83, 84, 111, 123, 138, 165, 197).

Spindler, M. Martsch, M. Dachselt, R. (2012). “Going Beyond the Surface: Studying Multi-Layer Interaction Above the Tabletop”. In: *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI ’12*. New York, New York, USA: ACM Press, p. 1277. DOI: `10.1145/2207676.2208583` (cit. on p. 138).

Spindler, M. Büschel, W. Winkler, C. Dachselt, R. (2013). “Tangible displays for the masses: spatial interaction with handheld displays by using consumer depth cameras”. In: *Personal and Ubiquitous Computing* 18.5, pp. 1213–1225. DOI: `10.1007/s00779-013-0730-7` (cit. on p. 165).

Spindler, M. Schuessler, M. Martsch, M. Dachselt, R. (2014). “Pinch-drag-flick vs. spatial input: rethinking zoom & pan on mobile displays”. In: *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI ’14*. New York, New York, USA: ACM Press, pp. 1113–1122. DOI: `10.1145/2556288.2557028` (cit. on p. 84).

Steimle, J. Jordt, A. Maes, P. (2013). “Flexpad: Highly Flexible Bending Interactions for Projected Handheld Displays”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI ’13*. New York, New York, USA: ACM Press, p. 237. DOI: `10.1145/2470654.2470688` (cit. on pp. 148, 165).

Streitz, N. A. Geißler, J. Holmer, T. Konomi, S. Müller-Tomfelde, C. Reischl, W. Rexroth, P. Seitz, P. Steinmetz, R. (1999). “i-LAND”. In: *Proceedings of the SIGCHI conference on Human factors in computing systems the CHI is the limit - CHI ’99*. New York, New York, USA: ACM Press, pp. 120–127. DOI: `10.1145/302979.303010` (cit. on p. 45).

Tan, D. S. Pausch, R. Stefanucci, J. K. Proffitt, D. R. (2002). “Kinesthetic cues aid spatial memory”. In: *CHI ’02 extended abstracts on Human factors in computing systems - CHI ’02*. New York, New York, USA: ACM Press, p. 806. DOI: `10.1145/506443.506607` (cit. on pp. 5, 46, 50, 130).

Tan, D. S. Meyers, B. Czerwinski, M. (2004). “WinCuts: manipulating arbitrary window regions for more effective use of screen space”. In: *Extended abstracts of the 2004 conference on Human factors and computing systems - CHI ’04*. New York, New York, USA: ACM Press, p. 1525. DOI: `10.1145/985921.986106` (cit. on p. 8).

Tan, D. S. Gergle, D. Scupelli, P. Pausch, R. (2006). “Physically large displays improve performance on spatial tasks”. In: *ACM Transactions on Computer-Human Interaction* 13.1, pp. 71–99. DOI: `10.1145/1143518.1143521` (cit. on p. 6).

Tandler, P. Prante, T. Müller-Tomfelde, C. Streitz, N. Steinmetz, R. (2001). “Connectables: Dynamic Coupling of Displays for the Flexible Creation of Shared Workspaces”. In: *Proceedings of the 14th annual ACM symposium on User interface software and technology - UIST ’01*. UIST ’01. New York, New York, USA: ACM Press, p. 11. DOI: `10.1145/502348.502351` (cit. on p. 164).

Tang, A. Tory, M. Po, B. Neumann, P. Carpendale, S. (2006). “Collaborative coupling over tabletop displays”. In: *Proceedings of the SIGCHI conference on Human Factors in computing systems - CHI ’06*. New York, New York, USA: ACM Press, p. 1181. DOI: `10.1145/1124772.1124950` (cit. on pp. 45, 59, 120, 137, 214).

Thudt, A. Hinrichs, U. Carpendale, S. (2012). “The Bohemian Bookshelf: Supporting Serendipitous Book Discoveries through Information Visualization”. In: *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI ’12*. New York, New York, USA: ACM Press, p. 1461. DOI: `10.1145/2207676.2208607` (cit. on p. 79).

Tsang, M. Fitzmzurice, G. W. Kurtenbach, G. Khan, A. Buxton, B. Fitzmaurice, G. W. Kurtenbach, G. Khan, A. Buxton, B. (2002). “Boom Chameleon: Simultaneous capture of 3D viewpoint, voice and gesture annotations on a spatially-aware display”. In: *ACM Transactions on Graphics*. Vol. 22. 3. New York, New York, USA: ACM Press, pp. 111–120. DOI: `10.1145/571985.572001` (cit. on pp. 83, 101, 111).

Ullmer, B. Ishii, H. (2000). “Emerging frameworks for tangible user interfaces”. In: *IBM Systems Journal* 39.3.4, pp. 915–931. DOI: `10.1147/sj.393.0915` (cit. on pp. 34, 36).

Valdes, C. Eastman, D. Grote, C. Thatte, S. Shaer, O. Mazalek, A. Ullmer, B. Konkel, M. K. (2014). “Exploring the design space of gestural interaction with active tokens through user-defined gestures”. In: *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI ’14*. New York, New York, USA: ACM Press, pp. 4107–4116. DOI: `10.1145/2556288.2557373` (cit. on pp. 173, 175).

Voida, S. Podlaseck, M. Kjeldsen, R. Pinhanez, C. (2005). “A study on the manipulation of 2D objects in a projector/camera-based augmented reality environment”. In: *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI ’05*. New York, New York, USA: ACM Press, p. 611. DOI: `10.1145/1054972.1055056` (cit. on pp. 172, 173, 175).

Vyas, D. Veer, G. Heylen, D. Nijholt, A. (2009). “Space as a Resource in Creative Design Practices”. In: pp. 169–172. DOI: `10.1007/978-3-642-03658-3_23` (cit. on p. 76).

Wallace, J. R. Scott, S. D. MacGregor, C. G. (2013). “Collaborative sensemaking on a digital tabletop and personal tablets: prioritization, comparisons, and tableaux”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI ’13*. New York, New York, USA: ACM Press, p. 3345. DOI: `10.1145/2470654.2466458` (cit. on p. 76).

Wang, M. Boring, S. Greenberg, S. (2012). “Proxemic peddler: a public advertising display that captures and preserves the attention of a passerby”. In: *Proceedings of the 2012 International Symposium on Pervasive Displays - PerDis ’12*. New York, New York, USA: ACM Press, pp. 1–6. DOI: `10.1145/2307798.2307801` (cit. on p. 163).

Weigel, M. Mehta, V. Steimle, J. (2014). “More than touch: understanding how people use skin as an input surface for mobile computing”. In: *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI ’14*. New York, New York, USA: ACM Press, pp. 179–188. DOI: `10.1145/2556288.2557239` (cit. on p. 173).

Weiser, M. (1991). “The computer for the 21st century”. In: *Scientific American* 265.3, pp. 94–104. DOI: `10.1038/scientificamerican0991-94` (cit. on pp. 9, 26–29, 31–34, 46, 59, 81, 192).

Weiser, M. (1999). “The computer for the 21st century”. In: *ACM SIGMOBILE Mobile Computing and Communications Review* 3.3, pp. 3–11. DOI: `10.1145/329124.329126` (cit. on pp. 3, 33).

Weiser, M. Brown, J. S. (1996). *The Coming Age of Calm Technology* (cit. on p. 192).

Wellner, P. (1993). “Interacting with paper on the DigitalDesk”. In: *Communications of the ACM* 36.7, pp. 87–96. DOI: `10.1145/159544.159630` (cit. on pp. 85, 87).

Wigdor, D. Wixon, D. (2011). *Brave NUI World: Designing Natural User Interfaces for Touch and Gesture*. Morgan Kaufmann, p. 264 (cit. on p. 3).

Wigdor, D. Shen, C. Forlines, C. Balakrishnan, R. (2006). “Table-centric interactive spaces for real-time collaboration”. In: *Proceedings of the working conference on Advanced visual interfaces - AVI ’06*. New York, New York, USA: ACM Press, p. 103. DOI: `10.1145/1133265.1133286` (cit. on pp. 172, 174, 180, 182).

Wigdor, D. Jiang, H. Forlines, C. Borkin, M. Shen, C. (2009). “WeSpace: the design development and deployment of a walk-up and share multi-surface visual collaboration system”. In: *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09*. New York, New York, USA: ACM Press, p. 1237. DOI: `10.1145/1518701.1518886` (cit. on p. 8).

Wilson, A. Benko, H. Izadi, S. Hilliges, O. (2012). “Steerable augmented reality with the beamatron”. In: *Proceedings of the 25th annual ACM symposium on User interface software and technology - UIST ’12*. New York, New York, USA: ACM Press, p. 413. DOI: `10.1145/2380116.2380169` (cit. on p. 163).

Wilson, A. D. Benko, H. (2010). "Combining multiple depth cameras and projectors for interactions on, above and between surfaces". In: *Proceedings of the 23nd annual ACM symposium on User interface software and technology - UIST '10*. New York, New York, USA: ACM Press, p. 273. DOI: `10.1145/1866029.1866073` (cit. on pp. 163, 165).

Winkler, C. Löchtefeld, M. Dobbelstein, D. Krüger, A. Rukzio, E. (2014). "Surface-Phone: A Mobile Projection Device for Single- and Multiuser Everywhere Tabletop Interaction". In: *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*. New York, New York, USA: ACM Press, pp. 3513–3522. DOI: `10.1145/2556288.2557075` (cit. on p. 163).

Wobbrock, J. O. Morris, M. R. Wilson, A. D. (2009). "User-defined gestures for surface computing". In: *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09*. New York, New York, USA: ACM Press, p. 1083. DOI: `10.1145/1518701.1518866` (cit. on pp. 172, 174).

Wozniak, P. Goyal, N. Kucharski, P. Lischke, L. Mayer, S. Fjeld, M. (2016). "RAMPARTS: Supporting Sensemaking with Spatially-Aware Mobile Interactions". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing System* 2016. Ed. by ACM. DOI: `http://dx.doi.org/10.1145/2858036.2858491` (cit. on p. 201).

Wu, C.-S. A. Robinson, S. J. Mazalek, A. (2007). "WikiTUI: Leaving Digital Traces in Physical Books". In: *Proceedings of the international conference on Advances in computer entertainment technology - ACE '07*. New York, New York, USA: ACM Press, p. 264. DOI: `10.1145/1255047.1255116` (cit. on p. 85).

Wu, C.-S. A. Robinson, S. J. Mazalek, A. (2008). "Turning a page on the digital annotation of physical books". In: *Proceedings of the 2nd international conference on Tangible and embedded interaction - TEI '08*. New York, New York, USA: ACM Press, p. 109. DOI: `10.1145/1347390.1347414` (cit. on p. 85).

Yang, J. Wigdor, D. (2014). "Panelrama: enabling easy specification of cross-device web applications". In: *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*. New York, New York, USA: ACM Press, pp. 2783–2792. DOI: `10.1145/2556288.2557199` (cit. on pp. 3, 8).

Yee, K.-p. (2003). "Peephole displays: pen interaction on spatially aware handheld computers". In: *Proceedings of the conference on Human factors in computing systems - CHI '03*. 5. New York, New York, USA: ACM Press, p. 1. DOI: `10.1145/642611.642613` (cit. on pp. 82, 83, 111, 138).

Zadow, U. Büschel, W. Langner, R. Dachselt, R. I. M. L. (2014). "SleeD: Using a Sleeve Display to Interact with Touch-sensitive Display Walls". In: *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces - ITS '14*. New York, New York, USA: ACM Press, pp. 129–138. DOI: `10.1145/2669485.2669507` (cit. on p. 7).

Zagermann, J. Pfeil, U. Rädle, R. Jetter, H.-C. Klokmose, C. Reiterer, H. (2016). “When Tablets meet Tabletops: The Effect of Tabletop Size on Around-the-Table Collaboration with Personal Tablets”. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI ’16*. New York, New York, USA: ACM Press, pp. 5470–5481. DOI: `10.1145/2858036.2858224` (cit. on pp. xix, 202).

Zimmerman, J. Forlizzi, J. Evenson, S. (2007). “Research through design as a method for interaction design research in HCI”. In: *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI ’07*. New York, New York, USA: ACM Press, p. 493. DOI: `10.1145/1240624.1240704` (cit. on pp. 77, 78).

List of Figures

List of Tables