Designing Fluid Interaction for Immersive Analytics using Hybrid User Interfaces in Mixed Reality

Doctoral Thesis

for obtaining the academic degree

Doctor of Natural Sciences (Dr. rer. nat.)

submitted by

Sebastian Hubenschmid

at the

Universität Konstanz

Faculty of Sciences

Department of Computer and Information Science

Konstanz, 2025

Date of the oral examination:	October 24, 2025
1. Reviewer:	Prof. Dr. Harald Reiterer
2. Reviewer:	Prof. Dr. Dieter Schmalstieg
Chair:	Prof. Dr. Falk Schreiber

Abstract

Immersive mixed reality environments are gradually becoming part of our everyday life. Yet, their sole reliance on spatial interaction paradigms presents inherent limitations for their broader applicability. One proposition, which is the focus of this dissertation, is to form a hybrid user interface by complementing mixed reality head-worn devices with familiar everyday devices, such as smartphones, tablets, and desktop computers. Although the concept of hybrid user interfaces has persisted for decades, no coherent delineation has emerged, leading to a distinct lack of consistent terminologies and design principles.

To take a step towards understanding the potential of hybrid user interfaces in general, this dissertation explores their potential for fluid interaction in immersive analytics. Hybrid interfaces are especially well-suited for immersive analytics, where the benefits of immersive 3D visualization must be balanced with the precision and familiarity of 2D interaction. Thus, this dissertation addressed the overarching research goal of: *How can hybrid user interfaces support fluid interaction in immersive analytics workflows?*

The first part of this dissertation addresses the foundational gap in prior literature. It introduces the concept of complementary interfaces to generally describe meaningful device combinations. It then contextualizes the concept of hybrid user interfaces within the current research landscape, thereby extracting key dimensions for describing and classifying hybrid user interfaces.

The second part presents the design and empirical evaluation of four hybrid user interface exemplars, each providing a unique perspective on the overarching research goal. The exemplars explore different design principles of hybrid user interfaces, such as enabling seamless transitions between devices, applying effective task allocation across heterogeneous environments, or creating novel interaction techniques enabled through such meaningful device combinations. The evaluation of each prototype contributes context-specific design insights and research implications that are synthesized into broader principles.

Overall, this dissertation revisits and contextualizes the fragmented concept of hybrid user interfaces within the current research landscape, providing design guidelines for achieving fluid interaction. While immersive analytics provides a fitting use case, the resulting guidelines extend to a wider range of scenarios, thus providing a foundation for the future of hybrid user interfaces.

Kurzfassung

Immersive Mixed Reality Umgebungen spielen eine immer größere Rolle in unserem Alltag. Die Abhängigkeit von räumlichen Interaktionsparadigmen bringt jedoch Einschränkungen für die breitere Anwendbarkeit mit sich. Ein Ansatz, mit dem sich diese Dissertation befasst, besteht darin, hybride Benutzungsschnittstellen zu bilden, indem Mixed-Reality-Headsets durch gängige Alltagsgeräte wie Smartphones, Tablets und Desktop-Computer ergänzt werden. Obwohl das Konzept hybrider Benutzungsschnittstellen bereits seit mehreren Jahrzehnten existiert, fehlt es bislang an einer kohärenten Abgrenzung, was zu einem Mangel an konsistenter Terminologie und fundierten Gestaltungsprinzipien geführt hat.

Um dieses Potential besser zu erfassen, untersucht diese Dissertation zuerst deren Eignung für das Konzept von „*fluid interaction*“ in „*immersive analytics*“. Hybride Benutzungsschnittstellen sind besonders geeignet für „*immersive analytics*“, wo immersiver 3D-Visualisierung mit der Präzision und Vertrautheit zweidimensionaler Interaktionen balanciert werden müssen. Diese Dissertation addressiert daher das übergreifende Forschungsziel: *Wie können hybride Benutzungsschnittstellen „fluid interaction“ in „immersive analytics“ Workflows unterstützen?*

Der erste Teil dieser Dissertation widmet sich der grundlegenden Forschungslücke in der bisherigen Literatur. Zunächst wird das Konzept komplementärer Schnittstellen eingeführt, um allgemeine bedeutsame Gerätekombinationen zu beschreiben. Anschließend wird das Konzept hybrider Benutzungsschnittstellen im aktuellen Forschungsstand verortet, wobei zentrale Dimensionen zur Beschreibung und Klassifikation solcher Schnittstellen extrahiert werden.

Der zweite Teil stellt das Design und die empirische Evaluation von vier exemplarischen hybriden Benutzungsschnittstellen vor, die jeweils eine spezifische Perspektive auf das übergeordnete Forschungsziel bieten. Die Beispiele untersuchen unterschiedliche Gestaltungsprinzipien hybrider Schnittstellen, wie etwa nahtlose Übergänge zwischen Geräten, effektive Aufgabenverteilung in heterogenen Umgebungen oder neuartige Interaktionstechniken, die durch solche bedeutsamen Gerätekombinationen ermöglicht werden. Die Evaluation der jeweiligen Prototypen liefert kontextspezifische Gestaltungserkenntnisse und Forschungsschlüsse, die zu übergreifenden Prinzipien zusammengeführt werden.

Insgesamt kontextualisiert diese Dissertation das bislang fragmentierte Konzept hybrider Benutzungsschnittstellen im Lichte aktueller Forschung und formuliert Gestaltungsempfehlungen zur Ermöglichung von „*fluid interaction*“. Obwohl „*immersive analytics*“ einen passenden Anwendungsfall darstellt, sind die entwickelten Richtlinien auf eine Vielzahl weiterer Szenarien übertragbar und bieten somit eine Grundlage für die Zukunft hybrider Benutzungsschnittstellen.

Acknowledgements

I owe gratitude to those who aided me on my path.

To the thesis committee.

I. Harald Reiterer
For starting me on this journey and giving me the support to grow into an independent researcher.

II. Dieter Schmalstieg
For the insightful discussions that continue beyond this thesis.

III. Falk Schreiber
For conducting the examination of this work with precision and fairness.

To the HCI Group in Konstanz and colleagues.

IV. Simon Butscher
For guiding me towards HCI and our first publication, which led to this thesis.

V. Jonathan Wieland
For making the workdays more enjoyable and the steady support inside and outside academia.

VI. Johannes Zagermann
For openness and ubiquitous support, from the very start to the very end (and beyond).

VII. Daniel Immanuel Fink
For balance and reliable support in shared efforts.

VIII. Anke Reinschlüssel
For practical insights into academic pursuits.

IX. Abdelrahman Zaky, Maximilian Dürr, Jens Müller, Tiare Feuchtner, Ingrid Henschler
For continuous positivity that created an enjoyable work atmosphere.

X. Student assistants
For their hard work that strengthened my research: Tobias Böhnemann, Robin Erb, Ziyao He, Patrick Koston, Kai Leibold, Daniel Leicht, Lisa-Maria Reutlinger, Paul Rinau, Simon Röhrle, Simon Schwitz, Moritz Skowronski, Cedric Wiese, Yasin Yagdi, and all others I may have forgotten.

XI. Coauthors
For enriching our work and integrating me into their own works (in no particular order): Katja Vock, Johannes Fuchs, Philipp Fleck, Aimée Sousa Calepso, Michael Sedlmair, Andrea Batch, Niklas Elmqvist, Steven Feiner, Jens Emil Grønbæk, Benjamin Lee, Raimund Dachselt, Jens Grubert, Markus Tatzgern, Rasmus Lunding, Kaj Grønbæk, Janis Fiedler, Kathrin Wunsch, Britta Renner, Alexander Woll, Hyunsung Cho, Akihiro Kiuchi, David Lindlbaur, Priscilla Balestrucci, Sven Mayer, Marc O. Ernst, Albrecht Schmidt.
Marc Satkowski and Julián Méndez deserve special thanks for their outstanding contributions to our survey.

To friends and family.

XII. Peter Schmitt-Förster, Henryc Fels, Felix Bauer
For their company throughout the year, despite my complaints.

XIII. Jannis Weiser, Natascha Weiser, Tobias Zimmermann
For social distractions outside the digital world.

XIV. My family
For constant support.

Publications

Parts of this dissertation have been previously published in the following peer-reviewed venues. Reused material is highlighted at the beginning of each corresponding chapter.

Journal Publications

- **Sebastian Hubenschmid***, Marc Satkowski*, Johannes Zagermann*, Julián Méndez*, Niklas Elmqvist, Steven Feiner, Tiare Feuchtner, Jens Emil Sloth Grønbæk, Benjamin Lee, Dieter Schmalstieg, Raimund Dachselt, and Harald Reiterer. "Hybrid User Interfaces: Past, Present, and Future of Complementary Cross-Device Interaction in Mixed Reality." Submitted to: *IEEE Transactions on Visualization and Computer Graphics*. 2025

  Responsibilities: Marc Satkowski, Johannes Zagermann, Julián Méndez, and I contributed equally to the literature analysis and writing. Julián Méndez created the wonderful sketches for our taxonomy. Niklas Elmqvist, Steven Feiner, Tiare Feuchtner, Jens Emil Sloth Grønbæk, Benjamin Lee, and Dieter Schmalstieg contributed to the writing of the paper and offered feedback. Raimund Dachselt and Harald Reiterer contributed to the writing and supervised the work.

- Rasmus Lunding, **Sebastian Hubenschmid**, Tiare Feuchtner, and Kaj Grønbæk. "ARTHUR: Authoring Human–Robot Collaboration Processes with Augmented Reality Using Hybrid User Interfaces." In: *Virtual Reality* 29.2 (2025), p. 73. ISSN: 1434-9957. DOI: 10.1007/s10055-025-01149-6

  Responsibilities: Rasmus Lunding spearheaded the writing, design, technical realization, and study conduct of this work. Tiare Feuchtner and I supported the writing of the paper. Tiare Feuchtner and Kaj Grønbæk supervised this work.

- Johannes Zagermann*, **Sebastian Hubenschmid***, Priscilla Balestrucci, Tiare Feuchtner, Sven Mayer, Marc O. Ernst, Albrecht Schmidt, and Harald Reiterer. "Complementary Interfaces for Visual Computing." In: *it - Information Technology* 64.4–5 (2022), pp. 145–154. ISSN: 2196-7032. DOI: 10.1515/itit-2022-0031

  Responsibilities: Johannes Zagermann and I equally spearheaded the writing of the paper. Priscilla Balestrucci, Tiare Feuchtner, Sven Mayer, Marc O. Ernst, and Albrecht Schmidt contributed to the writing of the paper. Harald Reiterer contributed to the writing and supervised the work.

* Contributed equally

Conference Publications

- **Sebastian Hubenschmid**, Johannes Zagermann, Robin Erb, Tiare Feuchtner, Jens Grubert, Markus Tatzgern, Dieter Schmalstieg, and Harald Reiterer. “SpatialMouse: A Hybrid Pointing Device for Seamless Interaction Across 2D and 3D Spaces.” In: *Proceedings of the 2025 31st ACM Symposium on Virtual Reality Software and Technology*. New York, NY, USA: ACM, 2025, pp. 1–13. ISBN: 979-8-4007-2118-2. DOI: 10.1145/3756884.3766047

  Responsibilities: I lead the writing and design, and conducted parts of the data analysis of this work. Johannes Zagermann provided support in all areas. Robin Erb spearheaded the implementation, study conduct, and parts of the data analysis. Tiare Feuchtner, Jens Grubert, Markus Tatzgern, and Dieter Schmalstieg provided feedback throughout the project and contributed to the writing of the paper. Harald Reiterer contributed to the writing and supervised the project.

- Jonathan Wieland, Hyunsung Cho, **Sebastian Hubenschmid**, Akihiro Kiuchi, Harald Reiterer, and David Lindlbauer. “Push2AR: Enhancing Mobile List Interactions Using Augmented Reality.” In: *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2024, pp. 671–680. ISBN: 979-8-3315-1647-5. DOI: 10.1109/ISMAR62088.2024.00082

  Responsibilities: Jonathan Wieland spearheaded the writing, design, technical realization, study conduct, and data analysis of this work. Hyunsung Cho supported the writing and study conduct. Akihiro Kiuchi and I supported the writing and technical realization. Harald Reiterer and David Lindlbauer supervised this work.

- **Sebastian Hubenschmid**, Johannes Zagermann, Daniel Leicht, Harald Reiterer, and Tiare Feuchtner. “ARound the Smartphone: Investigating the Effects of Virtually-Extended Display Size on Spatial Memory.” In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI ’23. Hamburg Germany: ACM, 2023, pp. 1–15. ISBN: 978-1-4503-9421-5. DOI: 10.1145/3544548.3581438

  Responsibilities: I contributed to the design of the study, co-supervised the implementation, analyzed parts of the study, and spearheaded the writing. Johannes Zagermann co-supervised the design and implementation, and contributed to formulating the research questions. Daniel Leicht was responsible for formulating research questions, implementing the prototype, conducting the study, and analyzing the study data. Harald Reiterer and Tiare Feuchtner supervised this work.

- **Sebastian Hubenschmid***, Jonathan Wieland*, Daniel Immanuel Fink*, Andrea Batch, Johannes Zagermann, Niklas Elmqvist, and Harald Reiterer. “ReLive: Bridging In-Situ and Ex-Situ Visual Analytics for Analyzing Mixed Reality User Studies.” In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI ’22. New Orleans LA USA: ACM, 2022, pp. 1–20. ISBN: 978-1-4503-9157-3. DOI: 10.1145/3491102.3517550

  Responsibilities: Jonathan Wieland, Daniel Fink, and I contributed equally to the writing, design, implementation, and analysis. Andrea Batch supported the study and contributed to the writing, design, and implementation. Johannes Zagermann provided assistance during the writing, design, analysis, and study. Niklas Elmqvist and Harald Reiterer supervised the project.

- Katja Vock, **Sebastian Hubenschmid**, Johannes Zagermann, Simon Butscher, and Harald Reiterer. “IDIAR: Augmented Reality Dashboards to Supervise Mobile Intervention Studies.” In: *Mensch und Computer 2021*. MuC ’21. New York, NY: ACM, 2021, pp. 248–259. ISBN: 978-1-4503-8645-6. DOI: 10.1145/3473856.3473876

  Responsibilities: Katja Vock spearheaded the design, implementation, study conduct, and data analysis of this work. Johannes Zagermann and I contributed to the writing of the paper. Simon Butscher supported the writing and supervised the design and study of this work. Harald Reiterer supervised the project.

- **Sebastian Hubenschmid**, Johannes Zagermann, Simon Butscher, and Harald Reiterer. “STREAM: Exploring the Combination of Spatially-Aware Tablets with Augmented Reality Head-Mounted Displays for Immersive Analytics.” In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI ’21. New York, NY, USA: Association for Computing Machinery, 2021, pp. 1–14. ISBN: 978-1-4503-8096-6. DOI: 10.1145/3411764.3445298

  Responsibilities: I spearheaded the writing, design, implementation, study conduct, and data analysis of this work. Johannes Zagermann and Simon Butscher supported the writing, design, and study. Harald Reiterer supervised the project.

  Note: This publication builds on work previously submitted as Master’s thesis at the University of Konstanz, titled “*MIDAiR – Multimodal Interaction for Visual Data Analysis in Augmented Reality*”. While the publication was prepared afterwards, it reuses substantial parts of the analysis and study. However, the study analysis, design insights, and research implications were revised and refined for the publication.

Workshop Submissions

- **Sebastian Hubenschmid**, Johannes Zagermann, Niklas Elmqvist, Tiare Feuchtner, Jens Grubert, Markus Tatzgern, Dieter Schmalstieg, and Harald Reiterer. “Revisiting Hybrid Input Devices for Immersive Analytics.” In: *2025 IEEE Conference on Human Factors in Immersive Analytics (HFIA)*. IEEE, 2025. DOI: 10.1109/HFIA68651.2025.00011

  Responsibilities: I spearheaded the writing of this work. Johannes Zagermann supported the ideation and writing of the paper, sketching out the proposed interaction techniques. Niklas Elmqvist, Tiare Feuchtner, Jens Grubert, Markus Tatzgern, Dieter Schmalstieg, and Harald Reiterer contributed to the ideation and writing of the paper. Harald Reiterer also supervised the project.

- **Sebastian Hubenschmid***, Daniel Immanuel Fink*, Johannes Zagermann, Jonathan Wieland, Harald Reiterer, and Tiare Feuchtner. “Colibri: A Toolkit for Rapid Prototyping of Networking Across Realities.” In: *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Sydney, Australia: IEEE, 2023, pp. 9–13. ISBN: 9798350328912. DOI: 10.1109/ISMAR-Adjunct60411.2023.00010

  Responsibilities: I spearheaded the writing and implementation of Colibri. Daniel Fink spearheaded the implementation of Colibri and contributed to the writing. Johannes Zagermann and Jonathan Wieland contributed to the writing. Harald Reiterer and Tiare Feuchtner supervised this work.

- **Sebastian Hubenschmid***, Johannes Zagermann*, Daniel Fink*, Jonathan Wieland*, Tiare Feuchtner*, and Harald Reiterer*. “Towards Asynchronous Hybrid User Interfaces for Cross-Reality Interaction.” In: *ISS'21 Workshop Proceedings: "Transitional Interfaces in Mixed and Cross-Reality: A New Frontier?"* 2021. DOI: 10.18148/kops/352-2-84mm0sggczq02

  Responsibilities: All authors contributed equally to this work.

Co-Authored Publications Not Present in This Thesis

- Christoph Anthes, Mark Billinghurst, Uwe Grünefeld, Hans-Christian Jetter, Hai-Ning Liang, Frank Maurer, David Aigner, Craig Anslow, Guillaume Bataille, Abraham Campbell, Judith Friedl-Knirsch, Alexander Gall, Renan Guarese, **Sebastian Hubenschmid**, Yue Li, Fabian Pointecker, Andreas Riegler, Daniel Roth, Rishi Vanukuru, Nanjia Wang, Lingyun Yu, Johannes Zagermann, and Daniel Zielasko. “Grand Challenges in Cross Reality.” Submitted to: *IEEE TVCG*. 2025
- Janis Fiedler*, Kathrin Wunsch*, **Sebastian Hubenschmid**, Harald Reiterer, Britta Renner, and Alexander Woll. “SMARTFAMILY2.0: A refined mobile health intervention to promote physical activity and healthy eating in a family setting – A randomized-controlled trial.” In: *JMIR mHealth and uHealth* (2025). DOI: 10.2196/65558
- Rasmus Lunding, **Sebastian Hubenschmid**, and Tiare Feuchtner. “Proposing a Hybrid Authoring Interface for AR-Supported Human-Robot Collaboration.” In: *7th International Workshop on Virtual, Augmented, and Mixed-Reality for Human-Robot Interactions*. 2024
- Kathrin Wunsch, Janis Fiedler, **Sebastian Hubenschmid**, Harald Reiterer, Britta Renner, and Alexander Woll. “An mHealth Intervention Promoting Physical Activity and Healthy Eating in a Family Setting (SMARTFAMILY): Randomized Controlled Trial.” In: *JMIR mHealth and uHealth* 12 (2024). DOI: 10.2196/51201
- Johannes Zagermann, **Sebastian Hubenschmid**, Daniel Immanuel Fink, Jonathan Wieland, Harald Reiterer, and Tiare Feuchtner. “Challenges and Opportunities for Collaborative Immersive Analytics with Hybrid User Interfaces.” In: *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Sydney, Australia: IEEE, 2023. DOI: 10.1109/ismar-adjunct60411.2023.00044
- Philipp Fleck, Aimée Sousa Calepso, **Sebastian Hubenschmid**, Michael Sedlmair, and Dieter Schmalstieg. “RagRug: A Toolkit for Situated Analytics.” In: *IEEE Transactions on Visualization and Computer Graphics* 29.7 (2023), pp. 3281–3297. ISSN: 2160-9306. DOI: 10.1109/TVCG.2022.3157058
- Simon Butscher, **Sebastian Hubenschmid**, Jens Müller, Johannes Fuchs, and Harald Reiterer. “Clusters, Trends, and Outliers: How Immersive Technologies Can Facilitate the Collaborative Analysis of Multidimensional Data.” In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '18. Montreal, Canada: ACM, 2018, pp. 1–12. ISBN: 978-1-4503-5620-6. DOI: 10.1145/3173574.3173664
- **Sebastian Hubenschmid**, Simon Butscher, Johannes Zagermann, and Harald Reiterer. “Employing Tangible Visualisations in Augmented Reality with Mobile Devices.” In: *MultimodalVis '18 Workshop*. 2018, p. 4

Contributing Theses

The following student theses that I supervised contribute to the indicated chapters of this dissertation:

- **Chapter 5: Exemplar: ARound the Smartphone**
  Is partially based on:
  Leicht, Daniel (2022). "PhoneHole: Analyzing the Effects of Screen Size on Navigation and Spatial Memory in a Hybrid User Interface". Master Thesis. University of Konstanz.
- **Chapter 7: Exemplar: SpatialMouse**
  Is partially based on:
  Erb, Robin (2025). "Design and Evaluation of the Spatial Mouse – A Hybrid Pointing Device for Interaction across 2D and 3D Spaces". Bachelor Thesis. University of Konstanz.

Contents

Chapter 1

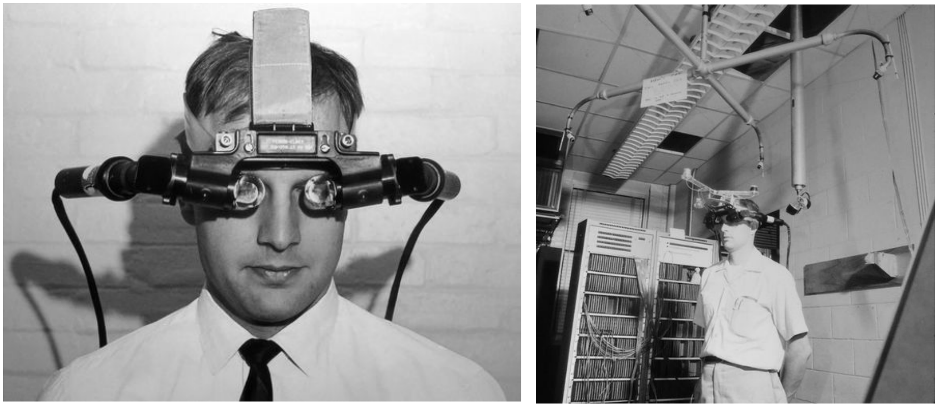
Introduction

In the past decade, there has been a renaissance of immersive mixed reality (MR) environments [276] driven by advancements in augmented reality (AR) and virtual reality (VR) hardware[1]. Once colossal frames of intricate machinery [376] (see Figure 1.1), head-worn devices (HWDs)[2] for immersive MR environments have now become affordable commodity devices set to pervade many aspects of our lives. Thanks to ongoing advancements in both hardware (e.g., increased fidelity of display technology) and software (e.g., foveated rendering, inside-out tracking), current HWDs are self-contained computing devices that have the potential to bring us closer to Weiser's vision of ubiquitous computing [405]. As such, these platforms have now evolved to the point where researchers can focus on the nuances of interaction design without being constrained by major technological limitations. Yet, despite their increasing commercial popularity, they are far from presenting a viable alternative to established computing platforms such as desktops and smartphones. What is holding them back?

At the time of writing, there are many factors that impede the viability of immersive MR environments. One factor can be attributed to lingering hardware limitations such as weight [416], heat [416], and vergeance-accomodation conflicts [221]. These limitations result in substantial discomfort issues [73, 310] and simulator sickness [209], making these immersive environments unattractive, especially for prolonged usage. Another factor is the limited scope of available applications, coupled with an isolated and ever-changing software ecosystem [63], which renders them impractical for any holistic or productive workflow. While these factors can be addressed given further iterations, I argue that one of the main factors is the almost exclusive reliance on spatial interaction paradigms.

[1]As the line between augmented and virtual reality becomes increasingly blurry, the term "*mixed reality*" in this work is taken to encompass both [368].

[2]Although the term "*head-mounted display*" (HMD) is more commonly used in prior literature, I believe that this term has now become archaic as the current hardware generation is no longer *mounted* to the user's head and consists of more than just *displays*. Instead, this thesis will use the term *head-worn devices* (HWD) to better reflect the state of contemporary hardware, but intentionally keep referring to *head-mounted displays* when referring to archaic hardware.



(a) Head-Mounted Display (b) Sword of Damocles

Figure 1.1: Sutherland's *head-mounted display* [376], published in 1968, marks one of the first steps towards AR hardware. Miniature CRT displays were *mounted* to the person's head, and a large frame mounted to the ceiling (often referred to as the "*sword of damocles*") was used to track the user's position.

1.1 The Era of Spatial Interaction?

Interaction design is often driven by futuristic visions popularized in science fiction [212]. In the case of immersive MR environments, part of the persistent fascination with spatial interfaces may be attributed to depictions such as Steven Spielberg's *Minority Report* (see Figure 1.2). These visions continue to shape contemporary scientific literature as well as commercial development, leading to the proclamation of "*the era of spatial computing*"[3]. This paradigm encompasses a variety of different input modalities, such as VR controllers, free-hand interaction, and other multimodal combinations (e.g., gaze+pinch [313]). While they individually offer distinct trade-offs, they all strive towards the goal of enabling effective interaction within a 3D space: Spatial interaction typically leverages input with six degrees of freedom (6DoF) to establish one-to-one mappings for 3D manipulation that mimics real-world object handling and provides natural navigation [224]. While this is a natural fit for the increased spatial output capacities of MR HWDs, they fall short when applied to broader, more complex workflows that demand accuracy, prolonged usability, or seamless integration with existing tools and work practices. Based on my research, three major conceptual limitations collectively hinder the viability of spatial interaction for real-world scenarios: *ergonomics*, *accuracy*, and *legacy bias*.

Ergonomics. One major challenge of spatial interaction methods, such as free-hand interaction and VR controllers, is the ergonomics for prolonged usage in 3D spaces. The Consumed Endurance model [168] estimates that users

[3] https://www.apple.com/newsroom/2024/04/apple-vision-pro-brings-a-new-era-of-spatial-computing-to-business/, last accessed 2025-05-12.

Figure 1.2: Spatial interaction as envisioned in Steven Spielberg's *Minority Report*. Using tracked gloves, the user interacts with a natural mid-air interface through hand gestures, enabling operations such as zooming or scrubbing through visions of a crime scene that float in mid-air.

can hold up their arm for just about 90 s before experiencing shoulder pains. Recent studies [242, 243] suggest that even this model was overly optimistic, as it "*overestimates [endurance time] for low-exertion interactions*" [242]. The resulting shoulder fatigue, commonly referred to as "*gorilla arm syndrome*", is well-known in interacting with vertical touch displays [42, 168]. Given the similarities to free-hand interaction, it is unsurprising that the findings extend to spatial interaction [161, 307].

Accuracy. Another challenge in spatial interaction is the limited accuracy of current input modalities, particularly within 3D environments. This limitation can be attributed to two factors.

First, the lack of haptic feedback presents challenges even for simple tasks like pressing a virtual button due to the absence of mechanoreceptive cues [239]. This *touching the void* [75] issue frequently leads to confusion and overshooting errors [55]. Although prior research has explored different methods of providing haptic feedback to mid-air gestures and handheld controllers (e.g., electrical muscle stimulation [250]), a widespread off-the-shelf solution has yet to emerge.

Second, biomechanical constraints introduce unintended motion during spatial input. For example, fine motor control can be challenging without adequate arm support [78, 100] due to hand jitter [36], and biomechanical constraints may cause translations to be accompanied by unintentional rotations [169]. Further, activating a discrete input, for example by pressing a button on a VR controller or performing a pinching gesture, causes slight hand movement that distorts the intended action – a phenomenon commonly referred to as the *Heisenberg effect of spatial interaction* [48, 411]. It should be no surprise that 2D mouse input can be more accurate for 3D manipulations than spatial 6DoF inputs [31].

Legacy Bias. Lastly, it is essential to acknowledge the *legacy bias* as a limitation for spatial interaction: Users have "*an explicit desire to transfer their knowledge of past systems to new ones*" [280]. While spatial interaction has the potential for natural interaction paradigms such as embodied interaction, it requires the adoption of entirely new conceptual frameworks, which can introduce a steep learning curve. Yet, many common tasks such as editing documents or browsing the Web do not benefit from increased spatial capabilities, as they remain firmly grounded on 2D surfaces [115, 152]. This can be seen as a form of "*backwards compatibility*" for current work practices: Despite the advantages of spatial interaction, much of the users' work remains in 2D. Yet, spatial input methods already struggle with issues of accuracy and ergonomics in 3D contexts, which becomes even more pronounced when applied to 2D surfaces [122].

Although these conceptual limitations do not apply equally to all spatial input modalities, they present significant challenges in the broader adoption of MR HWDs for everyday workflows. As a result, MR HWDs are often used in an isolated environment for spatially focused tasks, rather than being integrated into existing workflows. Yet, given the increased spatial capabilities of MR HWDs, spatial interaction remains a necessity for MR environments. Instead of treating them as a "*one-size-fits-all*" solution [218], we should instead investigate approaches that leverage the strengths of spatial interaction while still allowing users to retain their established 2D workflows, for example, by tightly integrating other devices.

1.2 Towards Multi-Device Ecologies

Users nowadays interact with a diverse range of devices, each serving a specific role within their personal device ecology. Understanding each device's role within this ecology is not only a foundational requirement for designing user experiences today [165], but also the focus of many research areas such as distributed user interfaces [110] and cross-device interaction [56]. Thus, despite their potential, it seems unlikely that MR HWDs will supersede other devices in the near future, especially given the conceptual limitations of spatial interaction. Instead, we should consider their position within the user's overall device ecology and explore their integration with established devices, which, in turn, could inform solutions to overcome existing challenges by relying on prior knowledge. For example, the "*gorilla arm syndrome*" was initially researched in the context of vertical touch displays [42, 168], but its findings equally apply to mid-air interactions [161, 307]. Following this, research in the context of vertical displays could be similarly applied to MR environments, such as addressing the "*gorilla arm syndrome*" on a vertical display through the use of a mobile device [11, 42].

But what makes these multi-device ecologies worthwhile? Simply adding more devices can be counterproductive, as it might not fit to users' workflow or current activity [314]. Instead, prior work in the context of AR environments argues

Figure 1.3: The concept of *hybrid user interfaces* emerged from combining a small, yet high-resolution desktop with a head-mounted display that offers virtually unlimited space, but suffers from low visual acuity [122]. Image courtesy of Steven Feiner.

that "*[multiple] displays are particularly relevant to AR applications when they are complementary*" [351]. Indeed, in a successful multi-device ecology each device or interface component possesses *complementary* characteristics, filling a niche that was not suitably covered before, thus forming a *complementary interface* [424]. While such *complementary interfaces* apply to a wide range of application scenarios (e.g., complementary roles in collaboration [297]), they are especially compelling for combining heterogeneous device technologies in a unified interface, thus forming a *hybrid user interface*.

1.3 Hybrid User Interfaces in Mixed Reality

The concept of *hybrid user interfaces* emerged in 1991 and was motivated by the technological limitations of early computing hardware: At the time, desktop computers offered limited, yet high-resolution displays, while early head-mounted displays provided unlimited virtual display space but suffered from low visual acuity (see Figure 1.3). Feiner and Shamash recognized that combining these heterogeneous device technologies and "*[treating] these as complementary, rather than competing, technologies*" [122] creates a hybrid user interface that is greater than the sum of its parts. Although later chapters will explore the specific characteristics in more detail, we can surmise a preliminary definition as follows:

> **Preliminary Definition: Hybrid User Interface**
> Hybrid User Interfaces are an area of cross-device computing that leverages distinct benefits of heterogeneous display and device technology.

The general proposition of combining "*heterogeneous display and device technology*" [122] theoretically allows for an infinite integration of technologies, yet such heterogeneous combinations are especially compelling for MR environments: The ubiquity, convenience, and familiarity of conventional 2D platforms (e.g., smartphones, desktops) provide a perfect complement to the immersion and complexity of optical see-through (OST) and video see-through (VST) HWDs. Given the focus of this thesis on MR environments, one could be inclined to define distinct technological combinations as specific subsets, such as defining the combination of conventional 2D devices and MR HWDs as "*mixed reality hybrid user interfaces*" or "*augmented hybrid user interfaces*". However, this thesis adopts the broader term "*hybrid user interface*" to refer to these configurations, relying on the overall framing of this thesis to delineate its scope and findings and avoid unnecessarily specific terminology.

Although hybrid user interfaces originally emerged as a result of significant limitations of early head-mounted hardware, their principles remain as relevant for contemporary HWDs: Modern desktop computers and mobile devices continue to offer limited but high-resolution displays, whereas contemporary HWDs provide unlimited virtual display space, but remain partially constrained by limited visual acuity and their reliance on spatial interaction paradigms. By applying the concept of complementarity and forming a hybrid user interface, we can therefore address the aforementioned major conceptual limitations of spatial interaction: To improve **ergonomics**, spatial interaction paradigms can be used when appropriate, while still allowing users to fall back on more physically sustainable alternatives when needed, thereby reducing fatigue [182]; for enhanced **accuracy**, users may seamlessly switch to high-precision 2D devices for tasks that demand exact positioning or fine control [179]; lastly, to overcome **legacy bias**, hybrid user interfaces can extend familiar workflows rather than supersede them entirely, preserving established work practices while gradually introducing new mental models and interaction paradigms [407].

While these theoretical benefits of hybrid user interfaces align well with findings from other research areas such as multimodal interaction, their design, implementation, and evaluation present unique challenges. Designers must enable seamless transitions between devices, manage shifts in visual attention, and ensure that interaction feels coherent across devices [187, 424]. This highlights the need for clear design principles to inform and support the development of hybrid user interfaces. However, quantifying the efficacy of design principles within an ecologically valid scenario is equally challenging, especially as comparative evaluations may not be possible as they, too, depend on highly specific design choices. Instead, adhering to established principles for interaction design, such as *fluid interaction*, can provide a valuable foundation for informing and shaping the design of hybrid user interfaces.

1.4 Revisiting Fluid Interaction for Immersive Analytics

Fluid interaction [111] is an interaction design concept in the field of information visualization, aimed at enabling seamless interactions and engaging user experiences. The concept of fluid interaction is defined by three characteristics: (1) promoting a state of flow, (2) supporting direct manipulation, and (3) minimizing the gulfs of action by reducing the cognitive and physical effort required to bridge intention and execution. By supporting these characteristics, an application can "*transform the sensemaking process into an efficient, illuminating, and even enjoyable experience*" [111]. Because fluid interaction has traditionally been applied in the context of 2D visualization, its extension to MR environments opens new opportunities for hybrid user interfaces to bridge conventional 2D visualizations with immersive 3D visualizations.

Although the use of 3D in information visualization has long been considered controversial [50, 360, 379], research in the context of immersive analytics [76] demonstrates its advantages in MR environments. Using immersive analytics, each visualization paradigm therefore offers distinct benefits: Conventional 2D visualizations are well-established and effective for providing an overview of dense information, which requires precise interaction. In contrast, immersive 3D visualizations are well-suited for inherently spatial data (e.g., motion trajectories) or providing environmental context (e.g., situated analytics [112]).

This dichotomy between 2D and 3D paradigms makes hybrid user interfaces a compelling platform for applying the principles of fluid interaction to the area of immersive analytics: By "*[treating] these as complementary, rather than competing,*" [122] visualization paradigms, hybrid user interfaces allow users to benefit from both. For example, providing a complementary 3D view to a 2D view has been shown to increase precision in spatial tasks [388]. By applying fluid interaction to immersive analytics using hybrid user interfaces, we can develop initial design insights and research implications for the use of hybrid user interfaces. This, in turn, creates the foundation for the broader applicability of hybrid user interfaces.

1.5 Research Objectives

Hybrid user interfaces represent a promising solution for bridging the gap between traditional 2D interactions and emerging 3D spatial environments. But without firm design principles and in the face of countless possible device combinations, the design space of hybrid user interfaces can appear bewilderingly large. To build an initial understanding of how such interfaces can be effectively designed, my goal is to investigate hybrid user interfaces for immersive analytics through the lens of fluid interaction, leading to the overarching research goal:

> **Research Goal**
> How can hybrid user interfaces support fluid interaction in immersive analytics workflows?

To unravel this exploratory research goal, this thesis is split into three parts.

Part I: Theory. First, to understand how hybrid user interfaces can effectively support fluid interaction, a clear definition of what constitutes a hybrid user interface is required. This thesis will therefore first examine the foundations of hybrid user interfaces, integrating them into adjacent research domains, such as cross-device interaction [56], distributed user interfaces [110], cross reality systems [362], and complementary interfaces [422]. Building on this foundation, this thesis then surveys how hybrid user interfaces are used in existing literature, thereby establishing a taxonomy as a framework for the remainder of this thesis.

Part II: Exemplars. Next, this thesis presents an in-depth exploration of the design, implementation, and evaluation of different hybrid user interface artefacts. These case studies are guided by three research objectives (ROs) that address the core challenges and opportunities of hybrid user interfaces:

> **RO1: Transitioning Between Devices**
> How can we design transitions between tightly coupled devices to support fluid interaction in hybrid user interfaces?

> **RO2: Task Allocation**
> How can an immersive analytics workflow be effectively distributed across heterogeneous devices in a hybrid user interface?

> **RO3: Interaction Techniques**
> What novel interaction techniques emerge from the use of complementary devices in hybrid user interfaces?

Each exemplar contributes to one or more of these research objectives, thereby yielding context-specific design insights that inform the overarching research goal.

Part III: Synthesis. While Part II provides an in-depth exploration of individual hybrid user interface prototypes, this part summarizes findings across the discussed prototypes to address the overarching research goal. By abstracting from the specific use cases and contexts in which the prototypes were developed, this dissertation identifies broader opportunities and challenges that transcend individual implementations. Thus, this thesis contributes to a more general understanding of how to design hybrid user interfaces to support fluid interaction within immersive analytics workflows.

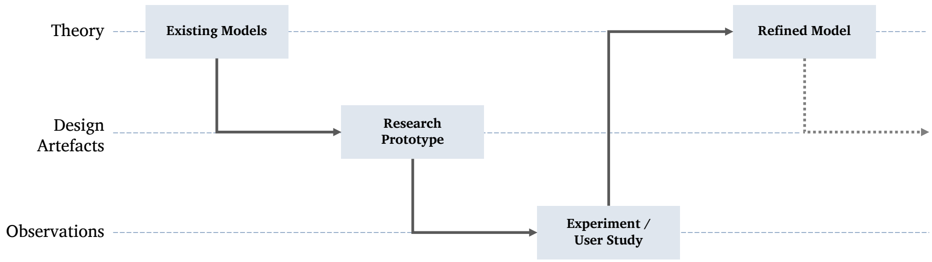


Figure 1.4: High-level research approach followed in this thesis as described by Mackay and Fayard [261].

1.6 Methodology

This thesis adopts a mixed-methods triangulation approach as described by Mackay and Fayard [261] (see Figure 1.4). My research started with a general understanding of existing models, such as cross-device interaction, multimodal interaction, and cross reality systems. These theories provided the foundation for the development of subsequent design artifacts, which either evaluate or explore novel design solutions for hybrid user interfaces. Each prototype was assessed through human-centered user studies, generating both empirical contributions and theoretical insights that, in turn, informed the design of future artifacts. This iterative, design-informed experimentation reflects the principle of using experiments as design-informing activities [298].

For each research prototype, I applied the validated solution design pattern [392]: First, a research gap or design challenge was identified based on prior theories coupled with a comprehensive literature review; next, a proposition in the form of a research prototype (i.e., design artifact) was created to explore a possible solution or compare different approaches; and lastly, the proposition was validated in an experiment or user study.

Evaluation methods were tailored to the goals of each prototype [147], but focus solely on human-centric evaluations, which is regarded as one of the "*core pillars of MR/AR evaluation*" [274]. A controlled experiment following best practices [175] was conducted when a clear comparative condition was feasible, providing both quantitative and qualitative data. To ensure methodological rigor and minimize confounding variables, these experiments were conducted using an abstract task rather than using immersive analytics scenarios. In contrast, exploratory usability studies were conducted to gain qualitative insights into the usage of novel interaction approaches. In these cases, immersive analytics tasks were chosen to enhance ecological validity and capture realistic user behavior.

Although my research followed this triangulation approach where possible, parts of my research, as well as the structure of this thesis, intentionally deviate from this chronological order. For example, the literature survey would have ideally been conducted at the very beginning of this thesis to frame subsequent research. However, the reality is that hybrid user interfaces have only recently gained enough traction in the research community to make such a survey meaningful (see Chapter 3). Nevertheless, the insights from this survey are presented upfront to provide an appropriate foundational understanding of hybrid user interfaces for the reader. Similarly, the exemplars are presented out of chronological order to improve the narrative flow of this thesis. Notably, STREAM (see Chapter 4) was developed for my Master's thesis and published afterwards with revised study analysis, design insights, and research implications.

1.7 Research Ethics

All empirical studies described in this thesis involved voluntary participation from human subjects. Each study received approval from the Ethics Committee of the University of Konstanz prior to its execution. During each study, we followed all necessary ethical and sanitary guidelines as well as best practices provided by the University of Konstanz. All participants received monetary compensation.

1.8 Contribution

The contributions of this work can be categorized into four high-level types [410]: *survey*, *theoretical*, *artifact*, and *empirical research contributions* (see Figure 1.5).

Survey Contributions. To understand how mixed reality hybrid user interfaces have been applied in prior work, this thesis presents an extensive literature survey. The survey helps to distinguish hybrid user interfaces from adjacent research areas, providing clarity on their unique position within the broader research landscape. In addition, the survey identifies current trends and research opportunities.

Theoretical Contributions. This thesis provides a twofold contribution to the theoretical understanding of hybrid user interfaces. First, it introduces the overarching concept of complementary interfaces [424], providing a lens to reason about meaningful combinations in multi-device ecologies. Second, it presents a taxonomy of hybrid user interfaces, contextualizing them within modern device ecologies and offering a structured framework through which future systems can be analyzed and designed.

Artifact Contributions. To investigate the potential of hybrid user interfaces, this thesis presents four research prototypes that were designed for immersive analytics. All research prototypes are available as open-source projects, allowing for their dissemination and reproduction within the research community.

Empirical Research Contribution. Each artifact in this thesis is accompanied by an empirical evaluation, resulting in findings that contribute towards a deeper understanding of hybrid user interfaces in practice. Individual findings are synthesized into design insights and research implications within their respective context, and are further synthesized into overarching principles that inform the broader design for supporting fluid interaction at the end of this thesis.

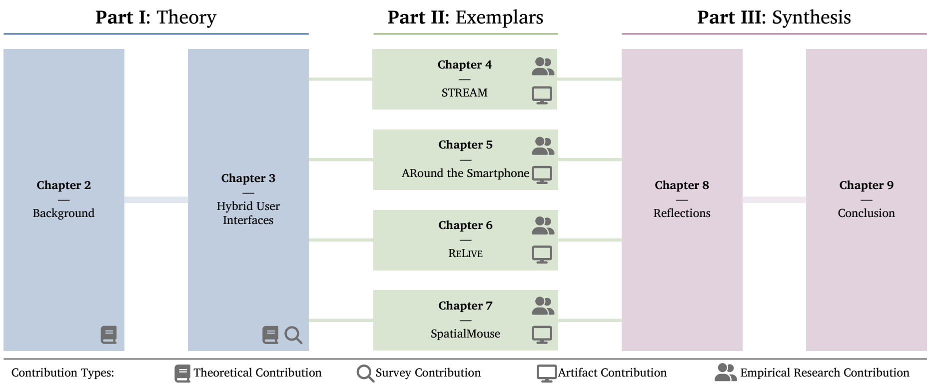


Figure 1.5: Overview of the structure and contribution types of this thesis.

1.9 Outline

This thesis is thematically organized into three main parts (see Figure 1.5). The following paragraphs provide a brief glimpse into the content of each part and chapter. Related publications are listed at the beginning of each chapter. The remainder of this thesis adopts the academic "*we*" instead of "*I*" to reflect the collaborative nature of the research and contributions.

Part I: Theory

Part I: Theory establishes the foundational concepts necessary for understanding hybrid user interfaces, including a terminology and taxonomy that serve as a framework throughout the remainder of the work.

Chapter 2 presents the **background** on related research fields that intersect with the concept of hybrid user interfaces. It provides an overview of distributed user interfaces, cross-device interaction, cross reality systems, and transitional interfaces, discussing how each relates to hybrid user interfaces. This chapter then introduces complementary interfaces as overarching conceptual model and outlines fluid interaction and immersive analytics as overall use case for this thesis.

Chapter 3 provides an in-depth investigation of the concept of **hybrid user interfaces**. This chapter explores the history of hybrid user interfaces, introduces a contemporary definition, and presents a taxonomy, which will serve as a framework for the remainder of this thesis.

Part II: Exemplars

Part II: Exemplars presents the design, implementation, and evaluation of four hybrid user interface prototypes. Each exemplar addresses research objectives that contribute to the overarching research goal by providing tangible design insights and research implications.

Chapter 4 introduces **STREAM**, a hybrid user interface that combines spatially-aware tablets with an immersive AR environment. STREAM highlights the potential of hybrid user interfaces in immersive analytics and demonstrates novel interaction techniques that facilitate seamless transitions between the AR environment and the tablet.

Chapter 5 presents **ARound the Smartphone**, a hybrid user interface that extends a smartphone with a virtual display through an AR HWD. This chapter presents an experiment investigating how different virtual screen sizes affect users' cognitive load, contributing to a better understanding of device transitions in hybrid user interfaces.

Chapter 6 presents **ReLive**, a hybrid user interface combining a desktop system with an immersive VR environment. This combination allows users to benefit from the strengths of both platforms: providing a computational notebook with aggregated statistics on a desktop and a reconstruction of the original context for sensemaking in VR. ReLive explores strategies for enabling smooth transitions between these otherwise isolated environments.

Chapter 7 introduces the **SpatialMouse**, a hybrid pointing device designed to enable seamless interaction across 2D and 3D spaces. Although not a hybrid *user interface* per se, it represents a potential hybrid *input device* that can support fluid interaction within hybrid user interfaces.

Part III: Synthesis

Part III: Synthesis builds on these findings by consolidating context-specific insights into overarching design principles aimed at enabling fluid interaction within immersive analytics workflows.

Chapter 8 provides a **reflection** on the design insights and research implications gained from the previous exemplars. By synthesizing these findings, the chapter outlines a set of general principles aimed at guiding the design of hybrid user interfaces, particularly in support of fluid interaction in immersive analytics.

Chapter 9 provides a **conclusion** by summarizing the key contributions of this thesis, outlining directions for future research, and discussing the future of hybrid user interfaces as we approach Sutherland's vision of the "*ultimate display*" [377].

Part I

THEORY

Chapter 2

Background

Chapter Content

Parts of this chapter are based on the following publications:

Sebastian Hubenschmid*, Marc Satkowski*, Johannes Zagermann*, Julián Méndez*, Niklas Elmqvist, Steven Feiner, Tiare Feuchtner, Jens Emil Sloth Grønbæk, Benjamin Lee, Dieter Schmalstieg, Raimund Dachselt, and Harald Reiterer. "Hybrid User Interfaces: Past, Present, and Future of Complementary Cross-Device Interaction in Mixed Reality." Submitted to: *IEEE Transactions on Visualization and Computer Graphics*. 2025

Johannes Zagermann*, **Sebastian Hubenschmid***, Priscilla Balestrucci, Tiare Feuchtner, Sven Mayer, Marc O. Ernst, Albrecht Schmidt, and Harald Reiterer. "Complementary Interfaces for Visual Computing." In: *it - Information Technology* 64.4–5 (2022), pp. 145–154. ISSN: 2196-7032. DOI: 10.1515/itit-2022-0031

* Contributed equally

2.1 Chapter Context

Hybrid user interfaces are closely intertwined with several adjacent research fields, sharing overlapping goals and concepts: Because they combine multiple devices, they represent a *distributed user interface*; their tightly-coupled interaction across devices aligns with *cross-device interaction*; their application in MR environments can be understood as part of *cross reality environments*; and the need to move between devices positions them within *transitional interfaces*. This chapter reviews these four adjacent research fields and discusses the position of hybrid user interfaces within this broader research landscape. It then introduces the concept of *complementary interfaces* as a theoretical lens for exploring how meaningful combinations can be designed across these domains. Finally, this chapter offers a brief overview of *immersive analytics* and *fluid interaction* to provide context for the investigated use case addressed in this thesis.

2.2 Distributed User Interfaces

The field of *distributed user interfaces* can be seen as a general foundation of any research on multi-device usage. Prior work by Elmqvist [110] defines the term as:

> **Definition: Distributed User Interface**
> "*A distributed user interface is a user interface whose components are distributed across one or more of the dimensions input, output, platform, space, and time*" [110].

In the context of hybrid user interfaces, all proposed dimensions are relevant, as they span multiple technologies and *distribute* the user interface accordingly. However, by combining heterogeneous technologies, hybrid user interfaces aim to distribute input and output to the most suitable device to achieve a goal while focusing on their complementary use. Therefore, *distributed user interfaces* can be seen as an overarching conceptual model, with hybrid user interfaces representing one possible technical realization of this concept that focuses on heterogeneous devices. Over time, this theoretical perspective of distributed user interfaces has been further consolidated under the more practical umbrella term of "*cross-device interaction*".

2.3 Cross-Device Interaction

Cross-device interaction focuses on research that "*transcends the individual device and user*" [56], unifying research that is focused on different kinds of multi-device environments, ranging from multi-monitor setups to ad-hoc mobile device ecologies. While there is no formal definition, cross-device interaction is often related to Weiser's vision of ubiquitous computing [405], studying devices such as smartphones, tablets, or larger interactive surfaces. Brudy et al. [56] offer a seminal overview of this field, introducing a device ontology and a taxonomy that outlines key characteristics of cross-device systems (see Figure 2.1). Their taxonomy includes the following dimensions:

- The **temporal** dimension distinguishes between *synchronous* (using devices simultaneously) and *asynchronous* use (using devices in sequence).
- The **configuration** dimension classifies multi-device arrangements, such as *mirrored* configurations, *spatial distributions* or *logical distributions*, *second screen* usage, and *migratory* or *cross-platform* arrangements. Note that configurations are closely tied to the temporal dimension: while most imply synchronous use, migratory and cross-platform configurations are inherently asynchronous.
- The **relationship** dimension describes user-device relationships, ranging from *single user*, multi-device scenarios (*1…m*) to collaborative *multi-user* settings with either single-device usage (*1…1*×*1…1*) or multi-device scenarios (*n…m*).
- The **scale** dimension describes to the physical proximity of devices relative to each user. This dimension shares close ties with proxemics [158], and is therefore classified into *near*, *personal*, *social*, and *public* scales.
- The **dynamics** dimension captures the degree of configurability in multi-device environments, ranging from *ad-hoc, mobile*, over *semi-fixed* to *fixed* dynamics. While *ad-hoc, mobile* configurations can be highly dynamic and easily reconfigured, *fixed* installations usually involve persistent arrangements with predetermined devices.

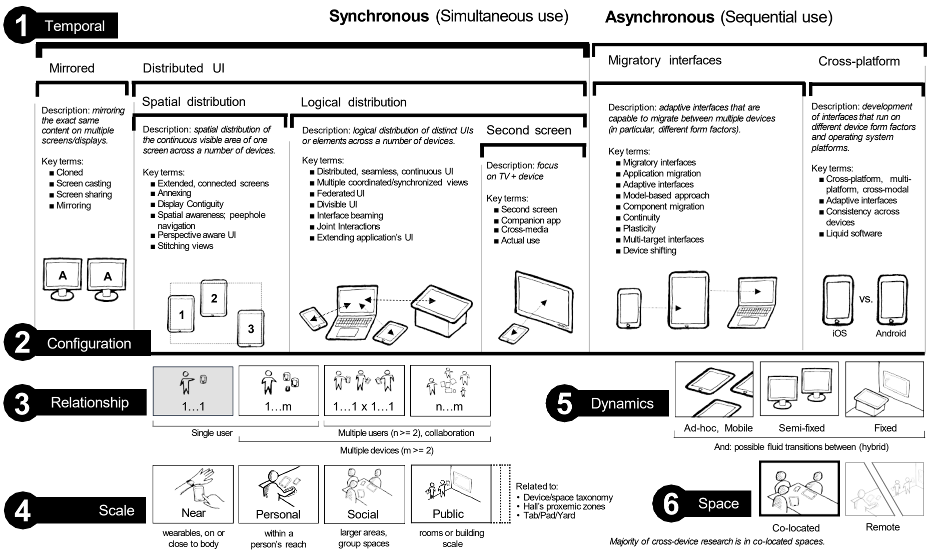


Figure 2.1: Taxonomy of cross-device interaction, adapted from Brudy et al. [56].

- The **space** dimension differentiates between *co-located* and *remote* interactions, linking cross-device research to research on computer-supported cooperative work (CSCW).

While their taxonomy and overall research focus are applicable to a wide range of devices, their main focus is on dynamic combinations of homogeneous devices: For example, head-worn MR devices and tangibles are a part of the cross-device ontology, but their taxonomy does not elaborate on combinations of heterogeneous (e.g., non-immersive and immersive) devices.

In addition, hybrid user interfaces differ in the relationship between devices: In cross-device interaction, additional devices are often introduced *ad hoc* to extend existing workflows. These devices offer optional functionality (e.g., externalizing content or distributing tasks across devices) but are not essential to the overall operation. As such, these systems exhibit a low functional codependency. In contrast, hybrid user interfaces are intentionally designed with tightly-coupled devices, leading to a high functional codependency. This relationship between devices is inherent to the interface itself: For example, one device may serve as input and another as output – creating a functional codependency between devices, such that the application is only complete once all devices are present.

Cross-device interaction can therefore be seen as an umbrella term that includes research not only on homogeneous but also heterogeneous combinations, as well as varying degrees of functional dependency between devices. Hybrid user interfaces represent a distinct subdomain that explicitly investigates heterogeneous device combinations with tightly-coupled devices.

2.4 Cross Reality Environments

Using the reality–virtuality (RV) continuum of Milgram and Kishino [276] as a foundation (see Figure 2.2), the research area of *cross reality* investigates the benefits of combining various points on the RV continuum. Assuming that various stages on this continuum (e.g., represented as AR or VR) are beneficial for solving a task, creating an environment that allows users to cross realities with different degrees of virtual content can be promising:

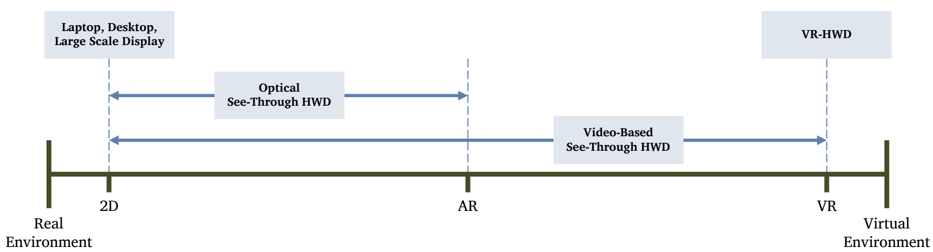


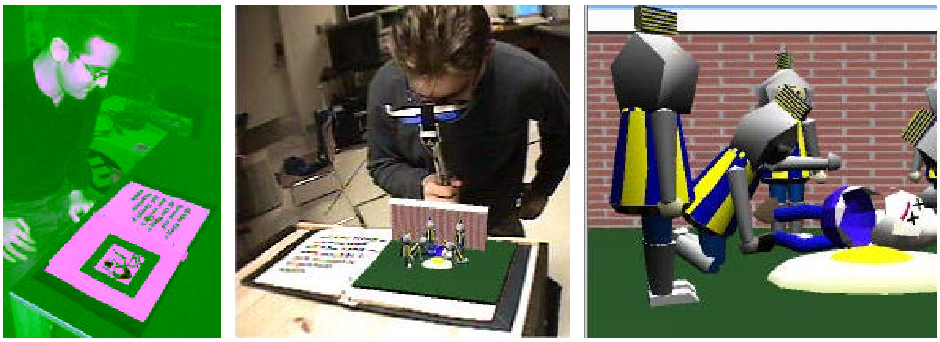
Figure 2.2: Milgram and Kishino's [276] reality–virtuality continuum, adapted from Anthes et al. [7]. Video-based see-through head-worn devices are the only devices that cover the full range of stages.

> **Definition: Cross Reality**
> "*Cross Reality allows users to transition between different stages of the [RV continuum], or interconnect different users at different stages of the [RV continuum]*" [7].

A wide range of applications emerges when systems leverage multiple stages along the RV continuum rather than being limited to a single point. Auda et al. [9] therefore further categorize such cross reality systems into three distinct types:

- **Transitional:** A user transitions between different stages on the RV continuum. For example, a user may transition from reality to VR (see Figure 2.3).
- **Substitutional:** Objects are repurposed for the user's current stage on the RV continuum. For example, a physical keyboard may be repurposed for VR to improve typing [270].
- **Multi-User:** Multiple users collaborating across different stages on the RV continuum. For example, one user may be in AR, while another user is in VR [297].

Key aspects of cross reality environments therefore include "*smooth transition[s] between systems using different degrees of virtuality*", allowing for "*collaboration between users using different systems with different degrees of virtuality*" [362]. Recent taxonomies by Auda et al. [9] and Wang and Maurer [399] further solidify concepts such as transitioning between different points on the RV continuum or concurrently using multiple distinct systems along the RV continuum. For a single user, such cross reality environments typically create a sequence of actions (e.g., switching devices). Depending on the transition, this could also be described as a migratory interface [56] or asynchronous hybrid user interface [187]. Given the emphasis on *transitioning* between devices, cross reality is also closely intertwined with the research area of transitional interfaces [71, 269].



(a) Reality (b) Augmented Reality (c) Virtual Reality

Figure 2.3: The MagicBook [39] demonstrates the use of multiple stages on the reality-virtuality continuum.

In the context of hybrid user interfaces, previous research mostly focuses on combining 2D devices with MR environments [43, 68, 122, 140], which can be seen as different stages on the RV continuum. In line with prior literature [7], cross reality can thus be viewed as an umbrella term for hybrid user interfaces, where a significant subset of research on hybrid user interfaces lies at the intersection between cross-device interaction and cross reality: While cross-device interaction focuses on the combination of *devices*, the area of cross reality concentrates on the combination of *realities*. Thus, the area of cross reality research encompasses both homogeneous (e.g., collaboration between HWDs in different realities) and heterogeneous device combinations (e.g., switching from desktop to VR). The latter – heterogeneous combinations across the RV continuum – aligns closely with the scope of hybrid user interfaces.

2.5 Transitional Interfaces

The MagicBook by Billinghurst, Kato, and Poupyrev [39] (see Figure 2.3) is often described as the first *transitional interface* [71, 269]: Here, a user can transition along the RV continuum – from browsing the physical book to a handheld AR display to immersive VR.

> **Definition: Transitional Interface**
> A transitional interface addresses the challenge of "*how do we effectively interact with, transition between, and collaborate across different types of environments?*" [146]

Similar to research on cross reality environments, transitional interfaces typically involve a sequence of activities (see Figure 2.4): A user solves one aspect of a given task in one environment and continues to work on the activity in another environment. The transition keeps the user oriented in the task space, choosing one environment (e.g., stage on the RV continuum or device) at a time. This transition may include explicit transitioning functions, which migrate the user's context and representation between environments, which may be further augmented with animations or visual cues to support continuity [234, 244, 355].

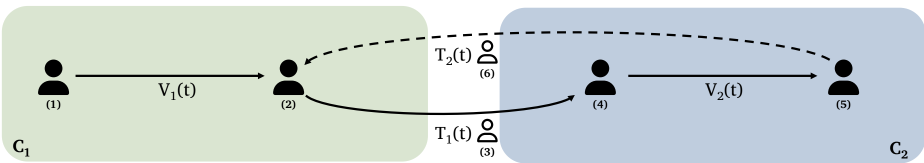


Figure 2.4: Step-by-step representation of a user transitioning between two contexts (C_1 and C_2). While the user can navigate within their respective context using a locomotion function $V(t)$, they can also *transition* between contexts using a transition $T(t)$, representing a *transitional interface*. Figure adapted from Grasset, Looser, and Billinghurst [146].

Although transitional interfaces can be regarded as a subset of cross reality, transitional interfaces focus on the design of transitions and their effect on users [269]. As such, transitional interfaces also apply to cross-device interaction, for example when transitioning between heterogeneous devices: While cross-device and cross reality systems describe *why* users transition between environments, transitional interfaces investigate *how* such transitions can be designed.

In the context of hybrid user interfaces, connecting to multiple stages on the RV continuum at the same time is a key ability (e.g., using a 2D device in an MR environment). In line with previous research, "*[hybrid user interfaces] and transitional interfaces [can be seen] as complementary*" [71], rather than competing: A hybrid user interface may use a transitional interface when switching between devices; likewise, a transitional interface may span across multiple heterogeneous devices to be switched between.

2.6 Complementary Interfaces

Traditional desktop interfaces rely on complementary input devices (e.g., mouse and keyboard) to perform tasks, such as pointing and text input. In contrast, many post-WIMP[1] [391] and ubiquitous computing interfaces [405] such as smartphones and tablets are self-contained, trading complementary peripherals with the convenience of built-in touch interaction and a combined input and output space. However, as task complexity increases, single devices may no longer be sufficient to adequately support users in their workflows: For example, research has shown that alternative input modalities can benefit our interaction (e.g., by improving spatial memory [423] or decreasing cognitive load [433]).

The need to go beyond individual devices, modalities, and environments therefore leads to the emergence of research areas such as distributed user interfaces, cross-device interaction, cross reality environments, and transitional interfaces, which all investigate the usage of different devices or environments. These research streams can be seen as manifestations of Mark Weiser's vision of the computer for the 21st century: "*specialized elements of hardware and software, connected […], will be so ubiquitous that no one will notice their presence*" [405]. The technological and methodological advances within the last decades allow researchers to design and evaluate new interaction paradigms beyond the boundaries of a single device and modality, leading to a variety of combinations of interfaces that can be used seamlessly in concert. Yet, little consideration has been given *when* such combinations are useful: "*without the adequate relation between them, these environments are simply mixed*" [210]. Indeed, handling multiple devices can increase cognitive load [322], with high transaction costs [176], and users are often not aware of the benefits of including additional devices into their workflow [314]. What makes these multi-device ecologies worthwhile?

[1] post-"**W**indows **I**cons **M**enus **P**ointer".

Simply adding more devices can be counterproductive, as it might not fit to users' workflow or current activity [314]. In a successful multi-device ecology, each device or interface component possesses complementary characteristics, filling a niche that was not suitably covered before. Attributing unique roles, properties, and purposes to each device and modality can lead to a worthwhile combination of interfaces that can overcome the mentioned issues. These meaningful combinations of devices and modalities can be called *complementary interfaces*:

> **Definition: Complementary Interface**
> "*A Complementary Interface is a meaningful combination of interfaces that support users in their current task at hand*" [420].

By distributing interaction across devices and modalities, we establish a *symbiosis of interfaces*, where each component purposefully increases the quality of interaction and further supports users in their current activity. Hence, complementary interfaces are an umbrella term that includes combinations of not only homogeneous and heterogeneous device classes, but also input (e.g., interaction techniques) and output modalities (e.g., visually or auditory). Importantly, complementary interfaces always feature some degree of heterogeneity in the involved components that complement each other to support the overall system functionality in solving the task at hand. These degrees of heterogeneity may lie in the input or output modality, location (e.g., screen space or input space), or dimensionality of data visualization (e.g., 2D and 3D).

In the context of hybrid user interfaces, complementary interfaces can therefore be considered as an overarching concept. Feiner and Shamash [122] suggest to treat multiple technologies in a *complementary* way, thereby taking advantage of their individual benefits. This can be seen as one of the key aspects of hybrid user interfaces, for example by complementing 2D visual and interaction spaces with MR environments. The core ideas of complementary interfaces are thus typically part of hybrid user interfaces, in that hybrid user interfaces represent a technical realization of this concept.

2.7 Immersive Analytics

Immersive analytics [76] investigates how technologies such as VR can support analytical reasoning and sense-making (see Figure 2.5). While the field encompasses a broad spectrum of novel technologies such as haptics and fabrication, this thesis focuses specifically on MR environments.

Definition: Immersive Analytics
"*Immersive Analytics investigates how new interaction and display technologies [such as MR HWDs] can be used to support analytical reasoning and decision making. The aim is to [...] allow users to immerse themselves in their data*" [76].

Although 3D visualizations have long been viewed with skepticism [50, 360, 379], there is an increasing amount of research indicating the benefits of 3D visualizations in MR (e.g., [22, 381, 418]). For example, 3D visualizations are well-suited for inherently spatial data (e.g., volumetric structures, motion trajectories) or providing environmental context (e.g., situated analytics [112]). In addition, Kraus et al. demonstrate the usefulness of immersion for 3D cluster identification [222] and showed that 3D heatmaps can outperform 2D heatmaps when comparing single data items [223].

In the past decades, different 3D visualizations in MR have been explored, such as flight trajectories [188], 3D parallel coordinates [66], interactively connecting and linking together different axes [22, 89], link routing between different visualizations in a 3D space [318], or 3D geotemporal visualizations [371]. Consequently, frameworks that facilitate the creation of immersive visualizations have emerged, such as DXR [361], IATK [88], VRIA [64], or RagRug [128].

However, the increased complexity of immersive visualizations has been identified as a grand challenge in immersive analytics [114]. It is thus unsurprising that much of the analytical workflow continues to rely on 2D environments, despite the potential of MR. As a result, prior work has increasingly explored hybrid approaches, combining 2D and 3D representations. For instance, 3D visualizations can enhance spatial precision when used as a complement to 2D displays [388],

(a) (b)

Figure 2.5: (a) Immersive analytics investigates how emergent technologies such as mixed reality can facilitate analytical reasoning and sense-making using, for example, stereoscopic 3D visualizations in collaborative settings [66]. (b) Immersive analytics and situated analytics [112] also benefit from the environmental context, such as visualizing data close to a referent within the user's environment [128].

while many others have demonstrated the benefits of pairing desktop-based analysis with immersive environments [186, 197, 289, 387].

Of course, the idea of combining different representations to leverage their individual strengths is not new: For example, multiple coordinated views (MCV) is a well-established concept that allows for different perspectives on the same data across different interlinked visualizations. Similarly, composite visualizations extend research on MCVs by structuring multiple visualizations in a shared spatial layout [199]. Although much of this research is anchored in 2D [101, 338], recent works have shifted towards applying these concepts to immersive analytics [432].

Research in MCV also highlights the importance of complementarity, formalizing the *rule of complementarity*: "*Use multiple views when different views bring out correlations and/or disparities*" [402]. This focuses solely on the data representation (i.e., output), yet interaction remains an essential aspect regardless of its representation [287]. Given this focus on complementarity views, it seems appropriate to also have complementarity inputs – thus building a *complementary interface*. To address the grand challenge in immersive analytics of "*coping with immersive analytics interaction complexity*" [114], the combination of MR HWDs for spatial capabilities with conventional 2D devices for precision and familiarity offer a promising path forward. Still, such hybrid user interfaces may "*[bring] about new, far more complex transitions which can affect interaction fluidity*" [114].

2.8 Fluid Interaction

Fluid interaction [111] is an interaction design concept in information visualization, aimed at enabling seamless interactions and engaging user experiences[2].

> **Definition: Fluid Interaction**
> "*Fluidity in information visualization is an elusive and intangible concept characterized by smooth, seamless, and powerful interaction; responsive, interactive and rapidly updated graphics; and careful, conscientious, and comprehensive user experiences*" [111].

The concept of fluidity in information visualizations traces back to 1998, where Igarashi et al. describe fluid interaction techniques that "*allow the user to fluidly shift attention from primary to secondary content*" [189]. However, the idea of fluidity re-emerged at a time when touch displays were becoming increasingly viable, shifting interaction design away from traditional WIMP[1] environments toward direct touch-based interactions [111]. Given the limitations of early touch hardware, researchers explored clever combinations of technologies, such as integrating physical tokens with touch displays, to overcome technical constraints and enhance the user experience (see Figure 2.6). Based on such early examples, Elmqvist et al. [111] proposed three properties of fluid interaction:

[2] See also https://hci.uni-konstanz.de/research/fluid-interaction-revisited/

(a) Facet-Streams (b) CollectionDiver

Figure 2.6: Examples of systems that exhibit fluid interaction. (a) Facet-Streams combines tangible tokens on an interactive touch display to materialize filter queries [200]. (b) This concept was later extended in the context of public libraries, using different physical tokens to support distinct filters and using cross-device interaction to visualize the query on a second screen [59, 171].

- **Promotes flow.** Building on the theory of flow [92], the aim is to offer a balanced challenge that keeps users engaged without overwhelming or boring them. Immersive environments can enhance this effect by enabling users to become fully absorbed in their data and tasks[2].
- **Supports direct manipulation.** The goal is to minimize indirection in the interface through direct or embodied interactions. Notably, this can be partially achieved by using a "*layered or spiral approach to learning that permits usage with minimal knowledge*" [111], allowing for immediate use with minimal prior knowledge. In immersive environments, this approach can also help mitigate legacy bias [280] by allowing users to transfer familiar concepts from traditional systems to new interaction paradigms.
- **Minimizes the gulfs of action.** By addressing the gulf of evaluation (i.e., visualizing and perceiving system state) and the gulf of execution (i.e., translating user intentions into actions) [291], the goal is to reduce the effort needed to bridge the gap between user intention and system response. Although immersive environments offer new possibilities for bridging these gulfs, they can also introduce complexity that risks overwhelming users, thereby increasing rather than minimizing the gulfs of action [114].

These goals are intentionally broad to be applied across a wide range of analysis scenarios. As a result, Elmqvist et al. [111] also propose initial design guidelines to promote fluid interaction. In the context of hybrid user interfaces, several of these guidelines are especially relevant and require careful design considerations:

- **Minimize indirection in the interface.** This can be achieved by using direct manipulation and avoiding isolated control panels. In hybrid user interfaces, some degree of indirection may be unavoidable (e.g., for transitioning between devices). Indirection should therefore be used with care in hybrid user interfaces to avoid disrupting the interaction flow.

- **Integrate user interface components in the visual representation.** Often, traditional interface elements such as text fields, sliders, and buttons are still necessary. These components should blend naturally with the visualization to maintain immersion. In hybrid user interfaces, however, offloading such controls to another device can be an effective strategy.
- **Reinforce a clear conceptual model.** The system should present a coherent and consistent view of its state, with all interactions reinforcing the user's mental model. In hybrid interfaces, this includes clearly communicating the roles and relationships of each device, so that the division of functionality is consistent and matches the user's mental model.
- **Avoid explicit mode changes.** Sudden shifts in visual or interaction modalities can disrupt user flow. Where possible, transitions should be smooth and implicit. As such, device transitions in hybrid user interfaces should be as seamless and implicit as possible, minimizing cognitive load and maintaining a continuous user experience.

Although hybrid user interfaces hold great promise for achieving fluid interaction within immersive analytics, these guidelines highlight that careful design is required, especially around designing smooth transitions between devices and minimizing interruptions to user flow. As such, it is essential to investigate how fluid interaction principles can be effectively applied within the context of hybrid user interfaces.

2.9 Chapter Conclusion

This chapter establishes the foundational background by exploring adjacent research areas and positioning hybrid user interfaces in relation to each. While each area offers valuable perspectives, many are either too broad (e.g., cross-device interaction, cross reality environments) or orthogonal in focus (e.g., transitional interfaces) to fully capture the unique characteristics of hybrid user interfaces. Although the concept of complementary interaction provides a useful theoretical framing, its scope still remains too general to address the specific nuances of hybrid systems. However, these adjacent perspectives provide an essential foundation for the continued exploration of hybrid user interfaces.

Similarly, immersive analytics has emerged as a promising application domain that closely aligns with the strengths of hybrid user interfaces. Within this context, the concept of fluid interaction offers valuable guidance for designing seamless and efficient user experiences. Together, these foundations provide the necessary context for both the exemplar use cases presented in Part II and the design principles presented in Part III.

Chapter 3

Hybrid User Interfaces

Chapter Content

This chapter is based on the publication:

Sebastian Hubenschmid*, Marc Satkowski*, Johannes Zagermann*, Julián Méndez*, Niklas Elmqvist, Steven Feiner, Tiare Feuchtner, Jens Emil Sloth Grønbæk, Benjamin Lee, Dieter Schmalstieg, Raimund Dachselt, and Harald Reiterer. "Hybrid User Interfaces: Past, Present, and Future of Complementary Cross-Device Interaction in Mixed Reality." Submitted to: *IEEE Transactions on Visualization and Computer Graphics*. 2025

* Contributed equally

3.1 Chapter Context

Before diving into the design of individual hybrid user interfaces, it is essential to first understand the key defining characteristics and potential design dimensions. Yet, no coherent definition has emerged since its initial conception. As a result, there is a distinct lack of consistent design models and terminologies – fragmenting the research community across overlapping research areas.

Although the adjacent areas presented in Chapter 2 describe various shared aspects, they fail to capture the nuances specific to hybrid user interfaces. For example, the seminal cross-device taxonomy by Brudy et al. [56] provides an overarching model for the research area of hybrid user interfaces. However, their taxonomy only considers established devices (e.g., mobile devices) and homogeneous combinations, leaving topics such as MR environments largely for future work – and thus unsuitable for this thesis. Yet, without firm design principles and in the face of countless possible device combinations, the design space of hybrid user interfaces can appear bewilderingly large. Therefore, this chapter extends the existing cross-device taxonomy by examining complementary cross-device interaction between traditional 2D device technologies (e.g., smartphones, desktops) and novel MR platforms. This chapter provides a brief history of hybrid user interfaces, identifies core characteristics, and establishes a taxonomy based on an extensive literature survey, thereby identifying current usage trends. This taxonomy serves as a foundational framework for the design and analysis of the hybrid user interface exemplars presented in subsequent chapters.

3.2 What is a Hybrid User Interface?

Since 1991, the term "*hybrid user interface*" has continued to evolve. Despite significant advances in device technologies, no coherent delineation has emerged that distinguishes hybrid user interfaces from related research areas. Instead, the concepts of a hybrid user interfaces are spread across many adjacent research areas (see Chapter 2). This fragmentation is further complicated by overlapping terminology: The term cross-device interaction is often used interchangeably with the term "*cross-surface interaction*", while earlier work in that area is often considered as "*multi-device*" or "*multi-display*" systems [56]. Similarly, there are other terms that are related to hybrid user interfaces, but focus on interaction techniques (i.e., "*hybrid interaction*" [36, 217]) or the setting (i.e., "*hybrid virtual environment*" [71, 97, 396]). This lack of common understanding makes it difficult to compare results and share insights.

In the following, we first review prior definitions and the *background* (Section 3.2.1) of hybrid user interfaces, building on previous work by Satkowski and Méndez [346]. Based on this, we discuss how we can *identify* hybrid user interfaces (Section 3.2.2) and derive three *attributes* of hybrid user interfaces that specifically combine MR environments with conventional display technologies (Section 3.2.3) to guide our literature survey.

3.2.1 Background

The term hybrid user interface was coined in the early 1990s by Feiner and Shamash, describing a combination of "*heterogeneous display and interaction device technologies*" [122]. They argue that physical device sizes decreased with the advent of portable computing devices (i.e., laptops). This led to reduced interface real estate while retaining high-resolution input and output. In contrast, immersive technologies such as MR (especially HWDs) – in the early 1990s and still partly today – offer a lower resolution for both input and output space. However, AR has the potential for virtually unlimited interfaces that exceed the capabilities of conventional display technologies. Thus, Feiner and Shamash propose to combine these two technologies by "*taking advantage of the strong points of each*" [122], treating the technologies as complementary instead of competing. They exemplify this concept by presenting a "*hybrid window manager*", combining a high-resolution yet restricted desktop interface with a low-resolution yet virtually unlimited AR head-mounted display interface into one unified application that blurs the boundaries between interfaces.

This initially broad scope opened up a vast design space, but its ambiguity may have limited the adoption of the term within the field, leading to the diffusion of this term as it was applied to an increasing variety of device combinations.

Although one prominent theme was the combination of 2D devices and immersive MR environments, this was not universally shared by all prior works. On the one hand, Butz et al. highlighted that hybrid user interfaces extend to various "*technologies and techniques, including virtual elements such as 3D widgets, and physical objects such as tracked displays and input devices*" [68]. They noted that the resulting global AR space can be shared, which is also discussed by Feiner, as hybrid user interfaces combine all devices "*in a mobile, shared environment*" [118]. Bornik et al. further emphasize the potential combination of MR environments with conventional 2D display technology for hybrid user interfaces, as they "*pair 3D perception and direct 3D interaction with 2D system control and precise 2D interaction*" [43]. This is echoed by Geiger et al., who state that hybrid user interfaces "*combine 2D, 3D, and real object interaction and may use multiple input and output devices and different modalities*" [140]. In contrast, Strawhacker and Bers employ hybrid user interfaces in a broader context without focusing on MR environments, presenting a hybrid user interface that allows "*users [to] switch freely between tangible and graphical input*" [374].

One commonly shared theme is the importance of *complementarity*: For example, Sandor et al. state that "*information [in hybrid user interfaces] can be spread over a variety of different, but complementary, displays*" [343]. Furthermore, in line with Butz et al. [68], they describe that users of hybrid user interfaces can "*interact through a wide range of interaction devices*" [343] – we thus see the potential of hybrid user interfaces not within a random assortment of technologies, but in a principled integration of different standalone "*interaction devices*".

3.2.2 Identifying Hybrid User Interfaces

Research on hybrid user interfaces shares several common themes (e.g., *complementarity*, *heterogeneity*), yet the term remains technology-driven and potentially misunderstood: A large number of prior hybrid user interfaces describe the term as a combination of 2D and 3D technology [43, 110, 169], including the initial system demonstration by Feiner and Shamash [122]. Moreover, the term hybrid user interface was previously also used for different constellations of interaction components (e.g., desktop combined with tangibles [374], mobile devices [37], or conversational interfaces [249]), indicating wider applicability. For the purpose of this thesis, we establish the following definition for hybrid user interfaces:

> **Definition: Hybrid User Interfaces**
> Hybrid User Interfaces are an area of cross-device computing that leverages distinct benefits of heterogeneous interaction components.

In contrast to the preliminary definition in the beginning of this thesis, this definition explicitly concentrate on conceptual device capabilities instead of specific technologies to avoid being limited by current hardware capabilities. To this end, the term *interaction component* is used throughout this thesis to refer to a standalone device set with the input and output capabilities necessary for interacting with a given application (e.g., desktop with mouse and keyboard, VR HWD with controllers). Further, this definition is intentionally vague about possible device combinations (see cross-device interaction [56] for a potential ontology).

While this definition highlights the broad potential design space of hybrid user interfaces, it also intentionally lacks specificity. Therefore, to explore this vast design space and gain concrete insights into commonalities of hybrid user interfaces, *this thesis focuses on the specific hybrid user interface subset combining 2D with MR-enabled interaction components*. This combination captures our interest, as it was initially presented by Feiner and Shamash [122], is most prevalent in prior hybrid user interface literature [346], and also reflects on work presented at the IEEE ISMAR workshop on hybrid user interfaces (2023) [180], constituting the most recent and comprehensive outlet for hybrid user interfaces.

This thesis thereby explicitly excludes the vast research area of *tangible interaction* [192] or nascent research areas (e.g., brain–computer interfaces [83], Internet of Things [342]) from this subset, thereby increasing the specificity of this thesis. While there is great potential in such combinations, they do not align with the primary research objective of this thesis and are, therefore, outside the scope of this thesis. Thus, throughout the remainder of this thesis, the term "*hybrid user interface*" refers to the combination of 2D interaction components with MR environments, allowing for a more focused investigation while avoiding unnecessarily over-specific terminology that might further fragment the research landscape.

3.2.3 Attributes of Hybrid User Interfaces in Mixed Reality

We framed our work on the contemporary and most prevalent subset of hybrid user interfaces – a combination of 2D interaction components with MR environments. We derive three *attributes* based on previous usage of the term to further guide our literature review. Please note that the first two attributes are indicative for the broader area of hybrid user interfaces, while the last attribute specifies the constraints of the investigated subset.

Multiple interaction components in heterogeneous roles. Hybrid user interfaces typically combine multiple interaction components where each fulfills a need that is not adequately addressed by other components. Although each interaction component is self-contained, a hybrid user interface is deliberately spread across multiple components to intentionally "*take advantage of the strong points of each*" [122]. We also see the potential beyond technological aspects and broaden our scope to include combinations of heterogeneous *roles*.

Interaction components are codependent. A hybrid user interface is composed of multiple interaction components, but the real power does not emerge from any individual interaction component, but from the interaction of all of them [405]. A hybrid user interface thus acts as one holistic application from the user's perspective. The deliberate spread of responsibilities across interaction components requires a certain degree of co-dependency between interaction components (i.e., a system may become nonfunctional without all interaction components present). We examine both synchronous (i.e., using interaction components in parallel) and asynchronous usage (i.e., using interaction components in sequence) [43, 187].

Complementing 2D with mixed reality interaction components. A common aspect found within prior hybrid user interface literature is to "*pair 3D perception and direct 3D interaction with 2D system control and precise 2D interaction*" [43]. Given the (at times) conflicting needs of 2D and 3D spaces, combining 2D and MR interaction components can yield a superior result. The remainder of this thesis refers to such interaction components as either *2D interaction component* or *MR interaction component*.

3.3 Review Methodology

We look at research that matches our previously specified attributes to obtain corpus of publications that describe a hybrid user interface combining 2D interaction components with MR interaction components. This section describes the process for identifying, filtering, and analyzing relevant publications, following the PRISMA [304] guidelines. We present the general search strategy (Section 3.3.1), selection process (Section 3.3.2), and data extraction (Section 3.3.3). Moreover, the limitations of our procedure (Section 3.3.4) and possible extensions of the survey (Section 3.3.5) are highlighted. Finally, we describe how this survey can be

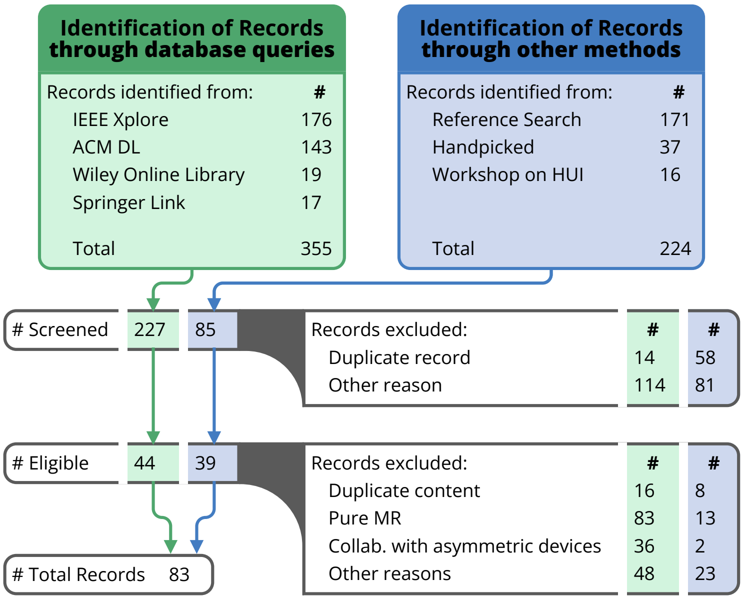


Figure 3.1: Overview of our reviewing process including paper counts, following PRISMA. Records were rejected once a single exclusion criterium was fulfilled, yet, some records potentially fulfilled multiple criteria (e.g., survey papers [9]).

used by other researchers and practitioners (Section 3.3.6). The complete survey corpus and its coding can be found in the accompanying webpage: https://iml-dresden.github.io/huis.

3.3.1 Search Strategy

In accordance with the PRISMA guidelines, we collected the papers using two different identification approaches (see Figure 3.1).

Identification via databases. Our survey focuses on publications that combine 2D interaction components with MR environments. As the term "*hybrid user interface*" is not consistently used in previous work, we used different adjacent terms and their synonyms (see Chapter 2) to build a query that captures a broader range of potential publications that present hybrid user interface applications.

We iteratively refined our keywords and evaluated them on a base corpus. Starting from the initial keyword "*hybrid user interface*", we identified synonyms used in the literature and extended our set of keywords. This resulted in three sets of keywords focused on different aspects that we aimed to capture. First, *Set A* includes keywords that could be used interchangeably with the term hybrid user interface. Next, *Set B* consists of adjacent terms and their synonyms that are not necessarily MR-specific. This is why, with *Set C*, we further narrow down the search to MR-related terms. The complete list of terms in each set is as follows[1]:

[1]Terms were used in singular and plural, as well as with or without hyphens (where appropriate).

Set A: `Hybrid User Interface`, `Complementary Interface`, `Augmented Display`, or `Cross Reality`

Set B: `Hybrid Virtual Environment`, `Hybrid Interaction`, `Cross-Device`, `Cross-Surface`, `Multi-Device`, `Multi-Display`, `Distributed User Interface`, or `Transitional Interface`

Set C: `Augmented Reality`, `Virtual Reality`, `Mixed Reality`, or `Extended Reality`

The query is composed of these three sets and was constructed as follows:

```
Set A or (Set B and Set C)
```

We focused on archival, peer-reviewed publications, including full and short papers in journals and conferences, as well as book chapters, workshop submissions, posters, and works in progress. We searched in common digital libraries for publications in Human–Computer Interaction and Visualization, namely IEEE Xplore, ACM Digital Library, Springer Link, and Wiley Online Library. The searches with the described query in these libraries[2] resulted in 355 publications. The cut-off date for all searches was August 12th, 2024.

Identification via other methods. In addition to the query search, we selected papers through three other methods. First, we selected all papers referring to the original definition of hybrid user interface by Feiner and Shamash [122] through a Google Scholar search using "*Publish or Perish*" [166]. Second, we manually added further publications that we found relevant. Finally, we selected all publications presented at the IEEE ISMAR 2023 Workshop on Hybrid User Interfaces [180]. The three methods resulted in 224 publications (see Figure 3.1). The cut-off date was August 12th, 2024.

3.3.2 Selection Process

After retrieving the initial corpus of papers (n = 579), we preprocessed the different output formats and merged them in one table. Three authors screened the reviewed corpus and filtered out duplicates and papers unrelated to MR. In addition, workshop proposals, complete books, dissertations, or conference proceedings were also removed.

Subsequently, three authors checked the eligibility of the remaining publications (n = 312). For that, the following exclusion criteria were defined:

- The publication is a duplicate (content-wise) of another (e.g., demonstration [329] of system [330]).
- No prototype or system was presented (e.g., discussion of opportunities and challenges [226, 331, 421, 425]).

[2] At the time of performing this query, the Springer library did not support an advanced search. Thus, we filtered the results using a script.

- Only a technological basis is described (e.g., frameworks [184, 342]).
- The publication only presents a summary of systems already in our corpus (e.g., position papers [118, 120]).
- A pure MR system without other interaction components is described (e.g., combining AR and VR environments [18, 81]).
- Other device capabilities are only used as input or output modality (i.e., interaction components are not standalone, e.g., attaching input sensors to an AR HWD [96, 266]).
- The publication only presents a collaboration across asymmetric devices (e.g., one user on a 2D interaction component, another in an MR environment [253, 297]).

During both selection steps, we split the data set between authors and discussed entries if the decision was unclear. This resulted in 83 relevant publications within our corpus (see Figure 3.2).

3.3.3 Data Extraction and Code Book

We first created an initial code book to extract data from the remaining 83 papers in our corpus. Since we see hybrid user interfaces as a sub-category of cross-device interaction, we started with aspects defined by Brudy et al. [56] in their survey. This was extended with common HCI metadata, such as the contribution type defined by Wobbrock and Kientz [410] or the evaluation strategy outlined by Ledo et al. [232]. Furthermore, we recorded each publication's challenges, future work, use case, and devices and terminology used. Once we defined an initial set of codes, we selected 10 papers at random from our literature corpus that three authors reviewed. This allowed us to (1) verify if the categorization works, (2) add missing codes, and (3) establish a common ground between authors. With

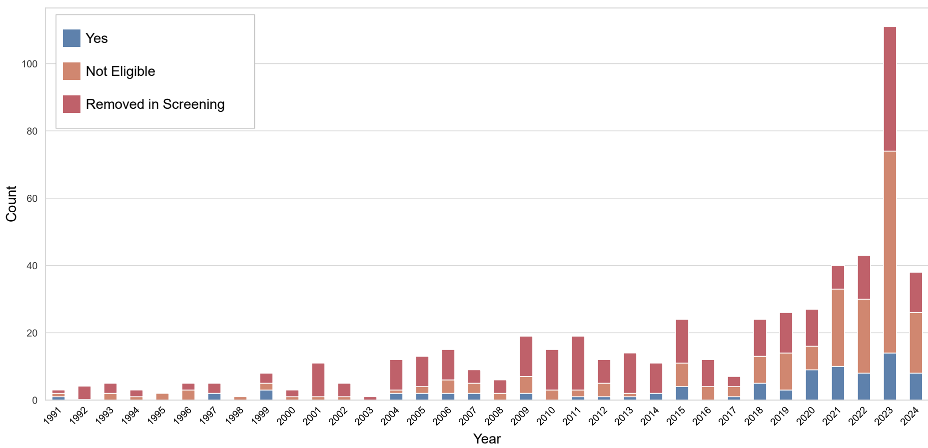


Figure 3.2: Record distribution over the past three decades.

the categories finalized, four authors coded the remaining papers. Two authors independently coded every publication. No author assessed the relevance of their own work. After each paper was coded, the same four authors discussed each publication and combined the two coding entries (e.g., clarifying conflicts). They further clustered challenges, use cases, and terminology used throughout the corpus.

3.3.4 Limitations and Further Considerations

We rigorously designed and conducted our survey. However, we identified limitations that we present in this section.

Terminology Bias. Hybrid user interfaces can be described by a multitude of adjacent terms, making it difficult to create a query that can capture all the systems within this field. Although we aimed to create the best possible query, this still led to a potentially incomplete set of publications. Therefore, we decided to add other sources (see Section 3.3.1) to our survey corpus.

Strict Eligibility Criteria. We searched for a specific type of device combination (i.e., MR-enabling devices and conventional 2D displays), leading to strict eligibility criteria (see Section 3.2.3). To avoid being overly restrictive, we decided to integrate papers that meet our characteristics but leave room for interpretation, especially papers representing potential directions for future research. We discuss these edge cases and their implications in Section 3.4.3. We acknowledge that our corpus is not exhaustive with regard to possible hybrid user interface systems caused by the use of said criteria. Yet, we see our corpus as a representative and substantial set of research on hybrid user interfaces, showcasing trends that can be generalized to the larger set of papers not captured by our query.

3.3.5 Dissemination and Extension

To allow others to benefit from our survey, we have made our literature corpus available using the Indy Survey Tool [90] as a GitHub project that hosts an interactive website for users to explore and filter our survey results, as well as to submit other work to our corpus (see https://imldresden.github.io/huis and Figure 3.3).

3.3.6 How to Use This Survey

With our survey, we aim to reveal common patterns and identify shared dimensions within hybrid user interfaces. Our work can be helpful in three ways: (1) By **classifying past research**, our *taxonomy of key characteristics* provides a general framework that can be used to better describe and understand inherent properties of hybrid user interfaces. With these dimensions, we can share insights between systems with overlapping dimensions and identify related works within our literature corpus that may not share a common terminology. (2) By **identifying**

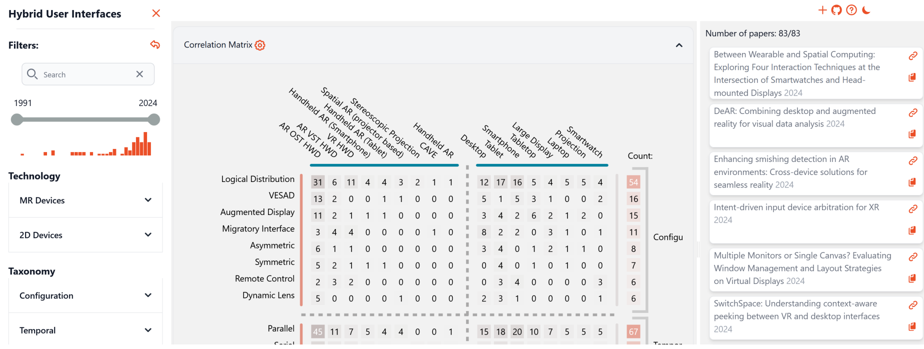


Figure 3.3: Our survey website enables interactive exploration of our corpus.

current trends, our survey shows emerging trends in device combinations, use cases, contribution types, and evaluation strategies. Readers can understand *how* hybrid user interfaces are used, informing the design and evaluation of upcoming hybrid user interfaces. Finally, (3) our survey can be used to **inspire future systems**, serving as a roadmap for the next generation of hybrid user interfaces.

Throughout our reporting (Sections 3.4 and 3.5), we denote the amount of literature in our corpus for each characteristic within a colored box mapped on a gradient with the total paper count (n=83) as the maximum value. Additionally, we provide a curated set of exemplary papers from our corpus for each characteristic, which can be used as a starting point for further reading. Since systems can be attributed to multiple characteristics, dimensions, or other categories, the sum of each category may not match the total number of records within our corpus.

3.4 A Taxonomy of Hybrid User Interfaces

Based on our survey, we establish a *taxonomy of key characteristics* for hybrid user interfaces (Section 3.4.1) that can characterize existing research and inform new research. In addition, we describe *emergent trends and opportunities* (Section 3.4.2) and highlight *edge cases* (Section 3.4.3) of our attributes.

3.4.1 Taxonomy Dimensions

We adopt six cross-device characteristics of Brudy et al. [56] (*Dimensions 1–6*) and introduce two additional dimensions to better represent the design space of hybrid user interfaces. To reflect the nuances of hybrid user interface, we extend and reframe each dimension and explain them accordingly. Although we describe these dimensions as discrete characteristics, they should be seen as continuous spectra, where systems, configurations, and interaction techniques can dynamically span multiple facets.

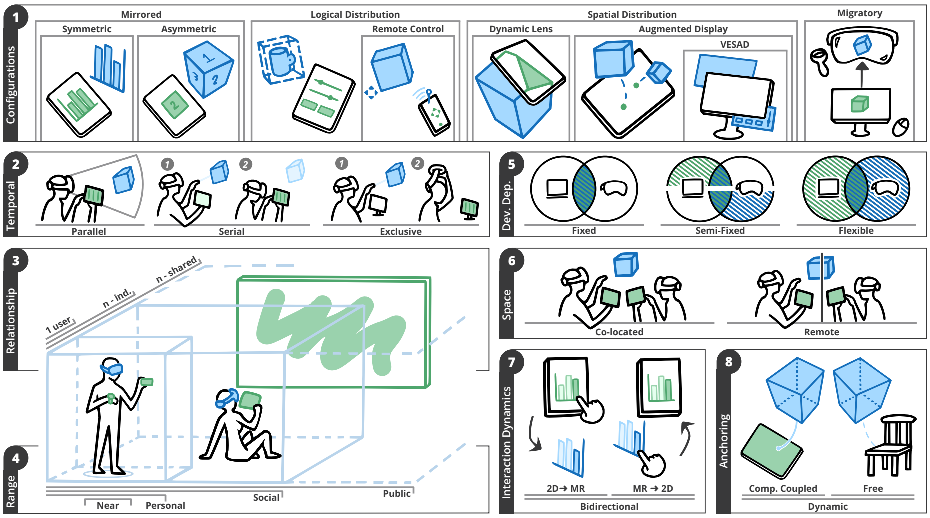


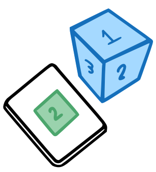
Figure 3.4: We present a taxonomy of key characteristics for hybrid user interfaces with eight dimensions: *configuration*, *temporal*, *relationship*, *range*, *device dependency*, *space*, *interaction dynamics*, and *anchoring*. Interaction components are highlighted in green for content on a 2D screen and blue for content in a mixed reality environment.

Dimension 1: Configuration

The configuration dimension describes how content and control are distributed across interaction components for a single user. Although we based our configurations on the cross-device taxonomy (i.e., *mirrored*, *logical distribution*, *spatial distribution*, *migratory interfaces*), we identified six configurations unique to hybrid user interfaces, namely *asymmetrical mirrored*, *symmetrical mirrored*, *remote control*, *dynamic lens*, *augmented displays*, and *VESADs*.

Mirrored configurations duplicate content in different interaction components. Given the perceptual differences of components in hybrid user interfaces, we distinguish between *symmetric mirror* and *asymmetric mirror* configurations.

- ▷ **Symmetric Mirror** 7 configurations duplicate content between interaction components, with each interaction component displaying the exact same view of information. This can reduce complexity when interacting with virtual content such as 3D sketching [8, 104] or immersive analytics [66, 179]. For example, a hybrid user interface can provide contextual connection based on users' surroundings in MR while providing familiar input on the 2D interaction component.

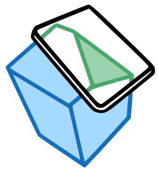


- ▷ **Asymmetric Mirror** 8 configurations show the same view of information but make use of the additional perceptual dimensions offered by MR interaction components: For example, this can be used to display a full 3D model or visualization in the MR environment, which is mirrored to an orthographic front view on the 2D interaction component [191, 357, 380].

Logical Distribution 54 describes configurations where content and control are distributed according to each interaction component's strength, usually involving a mutually exclusive allocation of responsibilities between interaction components. Most systems in our corpus involve some kind of logical distribution, such as offloading text input [34, 154, 431], shared and personal space [195, 337], or general application-control [289, 337, 345] to 2D interaction components, while the MR environment is used to display content in-situ.

▷ **Remote Control** 6 describes a stricter subset of *logical distribution*, where the 2D interaction component provides an alternative (but not exclusive) control over content in the MR environment. This can be useful for providing direct interactions when close by and indirect interaction from farther away (e.g., during 3D sketching [104]) or providing a more ergonomic option [116].

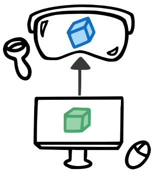
Spatial Distribution describes configurations that deliberately spread content across different spatial locations in the continuous real-world space. Such a distribution can be achieved through different means, described as *dynamic lenses*, *augmented displays*, or *VESADs*. Here, the content is aligned to the 2D interaction component (i.e., *augmented display*, *VESAD*) or reacts to the position of the 2D interaction component (i.e., *dynamic lens*).



▷ **Dynamic Lens** 6 allows 2D interaction components to act as a dynamic peephole [271] into a larger information space. While 2D interaction components provide a constrained view, the MR environment is unrestricted and can use the real environment. This requires the 2D interaction component to be spatially aware, enabling, for example, 3D slicing showing a cross-section of a 3D model on the 2D interaction component [225, 259, 380].

▷ **Augmented Displays** 15 use the unrestricted visual output of an MR environment to extend a 2D interaction component beyond its potential visual and interaction capabilities, acting as a 3D augmentation that is attached to the 2D interaction component [82, 99, 229, 335] . This category was initially defined by Reipschläger et al. [334, 336] as "*seamless combination of high resolution touch and pen enabled displays with head-coupled Augmented Reality*".

▷ **VESADs** 16 ("*Virtually Extended Screen-Aligned Displays*") were initially defined by Normand and McGuffin [292] as a virtual AR screen "*that is centered on, and co-planar with, a smartphone*". The content is strictly aligned to a 2D interaction component as a seamless display extension [122, 183, 229], representing a subset of *augmented displays*. For example, *VESADs* can be useful for offloading menu elements into MR to save screen real estate on a 2D interaction component [49, 292, 336] or provide alternative interaction capabilities [49]. To better represent the diversity of real-world multi-display configurations, we also include configurations that are aligned to 2D interaction components but not necessarily co-planar, such as extending 2D interaction components with angled virtual 2D screens [309, 310, 333].

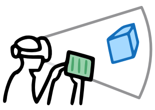


Migratory Interfaces 11 enable users to transfer their content or workflow from one device to another. Although originally considered asynchronous within the cross-device taxonomy [56], hybrid user interfaces can also utilize the MR environment to seamlessly transfer content between 2D interaction components [358] or between a 2D and 3D interaction component [4, 400, 406, 413]. In contrast, their asynchronous usage can take advantage of different environments to best support a holistic workflow (e.g., by switching between VR and desktop environment [3, 74, 186]) while still behaving as one unified, continuous system.

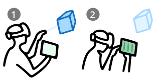
Dimension 2: Temporal

Prior literature [43, 56, 187] classifies hybrid and cross-device systems as *synchronous* or *asynchronous*. Due to the diversity of interaction components within hybrid user interfaces, we adopted the suggestion by Bornik et al. [43] to further differentiate between **parallel** and **serial** usage of interaction components. Fully asynchronous usage of interaction components was further classified as **exclusive** for cases where the usage of one interaction component rules out the usage of another (e.g., due to spatial or time-related restrictions).

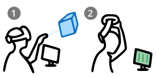
Some papers describe multiple interaction techniques that can be used, for example, in parallel or serial; others only present usages of one distinct aspect. We report the results accordingly: presenting the number of papers describing mixed and distinct temporal usage first and its subset describing only distinct usage second (e.g., 9 / 1).



Parallel 67 / 43 usage indicates that multiple interaction components are used simultaneously, for example, when interacting with one interaction component while observing the output on another interaction component [82], transferring content across devices [179], or extending a 2D interaction component in the MR environment [183, 229, 335].



Serial 36 / 6 usage indicates that multiple interaction components are used one immediately after another, such as selecting objects of interest in the MR environment and then editing them on a 2D interaction component [66]. Here, users have immediate access to all interaction components but can focus on only one interaction component at a time to reduce information overload or divide responsibilities between 2D and 3D interaction [43, 104, 121].



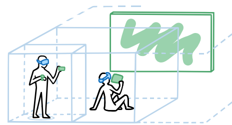
Exclusive 12 / 4 usage describes asynchronous systems where different interaction components cannot be used simultaneously as part of one workflow but have to be used in sequence. This can be useful to bridge the gap between traditional 2D computing environments and VR environments [3, 74, 186, 354].

Dimension 3: Relationship

The *relationship* category denotes the relation between users within one or multiple systems. The vast majority of our corpus describes **single-user** systems 72 :

One user interacts with multiple complementary interaction components (see Figure 3.4: *1 user*). However, we identified several collaborative systems. We abstracted these into **multi-user with individual interaction components** 7 : Each collaborator has their own set of interaction components [337, 354] (see Figure 3.4: *n-ind.*); and **multi-user with shared interaction component** 12 : A display is shared between collaborators as a public space [66, 225, 281, 335] (see Figure 3.4: *n-shared*).

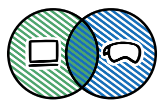
Dimension 4: Range



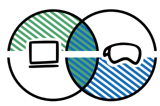
We examined the range of interaction across components as defined in the cross-device taxonomy [56]. Since the MR interaction component is commonly worn on the user's head (e.g., AR HWD) or close to their body (e.g., handheld AR), we use this dimension to describe the scale of the 2D interaction component relative to the user. In addition, the *relationship* can be an indicator of the *range* between interaction components (e.g., *multi-user* can indicate *social* or *public* range). We differentiate between **near** 6 (i.e., interaction component is close to the user's body, e.g., smartwatch [116, 150, 396]), **personal** 65 (i.e., interaction component is in the personal space, e.g., smartphone or tablet [179, 225, 229, 345]), **social** 23 (i.e., interaction component is in a social space accessible to collaborators, e.g., shared display [66, 281, 335]), and **public** 0 scale (i.e., interaction component can be seen and interacted with by arbitrary bystanders).

Dimension 5: Device Dependency

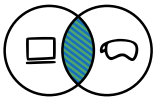
Device dependency describes the autonomy of each interaction component within the overall device ecology. We coded this as **flexible**, **semi-fixed**, and **fixed**:



Flexible 15 device dependency indicates that all interaction components provide basic features – a single interaction component could suffice to interact with a system in a meaningful way. This can be helpful for workflows that can be easily divided properly into specific subtasks for different environments, such as data analysis workflows [186], note taking [413], or sketching [8]. The main value of hybrid user interfaces in this situation comes from the interaction between the components, such as the seamless transition between components.

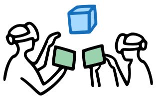


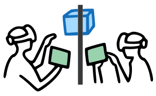
Semi-Fixed 40 device dependency represents systems in which one interaction component is completely independent, while others provide supplementary functionality and are thus reliant on the "*main*" component. This dynamic has been used primarily to extend the capabilities of existing 2D components, such as increasing the available screen space [122, 183, 229] or offering complementary views on existing content [34, 150, 335]. In contrast, a *semi-fixed* dependency can also be used to provide complementary capabilities to MR environments, such as providing a shared public space [195] or extracting content from 2D interaction components [355].



Fixed 27 device dependency describes systems that can only be used meaningfully with all relevant components present. With *fixed* device dependencies, responsibilities are exclusively distributed between components, such as using the 2D interaction component as a haptic surface for touch [44, 66, 82], for contextual interaction within the MR environment [121, 179, 347] or as a spatially aware controller [225, 259].

Dimension 6: Space





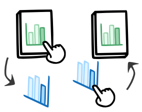
The space dimension describes whether interaction components are **co-located** 83 (i.e., within the same physical space) or **remote** 0. Since it may be difficult to achieve a synchronous hybrid user interface with remote interaction components, all the records we surveyed were exclusively *co-located*. An asynchronous hybrid user interface with *remote* interaction components may be feasible but would likely forfeit potential benefits gained from any combination of complementary devices (cf. [56]). Instead, we see the potential of such remote combinations in collaborative scenarios with *multiple user having individual interaction components* relationships. Potential remote collaboration scenarios might include individual hybrid user interfaces per location and user (e.g., a HWD to visualize data in AR and a handheld device for precise input): Here, the AR visualization could be shared and synced, while the handheld devices allow individual manipulations. However, such scenarios might also require user representations (e.g., to create awareness), a communication channel, and a merge policy for conflicting input.

Dimension 7: Interaction Dynamics

Since systems in our corpus are intentionally spread across multiple interaction components, the interaction dynamics describe how each interaction component can interact with a system. Building on the BISHARE [431] design space, which classified interaction concepts as either *HWD-centric* or *phone-centric*, we identified three kinds of dynamics:

Unidirectional (2D-centric) 35 dynamics, which signifies that input is only possible from the 2D interaction component. Examples include touch input [66, 179, 229], mouse & keyboard input [122, 211, 310], or using 2D interaction component as spatial controller [225, 259] (see Figure 3.4 *2D→MR*).

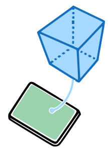
Unidirectional (MR-centric) 4 indicates that it is only possible to use input modalities provided by the MR environment, such as controller [190, 240, 355] or mid-air gestures [195] (see Figure 3.4 *MR→2D*).



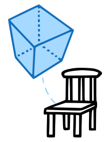
Bidirectional 44 dynamics indicate that all interaction components can interact with the system equally, for example, by switching between 2D and 3D sketching [104], visualizations [289], or transferring content [358, 413].

Dimension 8: Anchoring

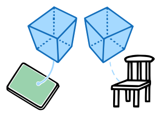
This dimension describes where the content is placed in the MR environment. We extended the design space by Reichherzer et al. [330], which categorizes content as either world-fixed or device-fixed. In addition, previous taxonomies have explored anchoring in more detail (e.g., semantic and spatial coupling in world-fixed content [108], content layout of device-fixed content [72, 325]), which we consider out of scope for this taxonomy. Our corpus is split almost equally between three general anchoring techniques:



Component-coupled 26 anchoring relates content within the MR environment to the 2D interaction components (*device-fixed* [330]). To create the illusion of spatial awareness and proper alignment of 2D interaction components, *component-coupled* usually involves spatial calibration of the 2D interaction component for stationary devices or active tracking for mobile platforms. *Augmented displays* and *VESADs* always involve a *component-coupled* anchor, while other systems require knowledge about the 2D interaction component for transferring content [355].



Free 30 anchoring describes MR content that is independently placed in the world or attached to real objects in the environment (*world-fixed* [330]). This can be used when the 2D interaction component is not directly related to the MR environment, for example, when providing a menu in the 2D interaction component [104, 394] or a simplified but detached view of the MR content [44, 289].



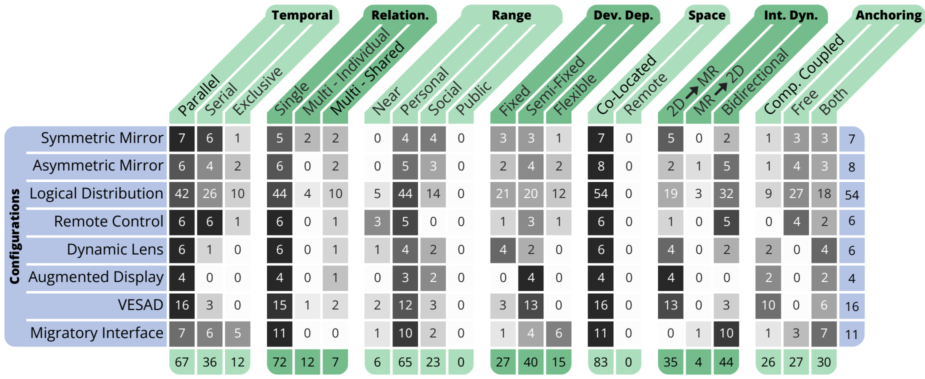
Dynamic 27 anchoring may support both *component-coupled* and *free* anchoring. Previous work has explored this approach in terms of transferring content from a 2D interaction component to the MR environment or vice versa [179, 413, 431] or for cutting through 3D models [225, 259].

3.4.2 Emergent Trends and Opportunities

To discover possible usage patterns of different configurations, we look at the distribution and trends of *dimensions 2–8* across the *configuration* dimension (see Figure 3.5). We can thus identify several trends:

1. *Spatial distribution* configurations (including *augmented display* and *VESADs*) have specific requirements, making them unsuitable for *serial* and *exclusive* temporal usage as well as *flexible* dynamics usage. Their focus on extending a 2D interaction component is indicated by the lack of systems that demonstrate *unidirectional (MR-centric)* interaction dynamics and *free* anchoring.

2. *Migratory* configurations can be used regardless of their *temporal* dimension, but have only been explored so far within *single-user* systems. In this configuration, all interaction components appear to be equally important for interaction, since almost every system in our corpus uses a *bidirectional* interaction dynamics.



Configurations	Temporal			Relation.			Range				Dev. Dep.			Space		Int. Dyn.			Anchoring			
	Parallel	Serial	Exclusive	Single	Multi - Individual	Multi - Shared	Near	Personal	Social	Public	Fixed	Semi-Fixed	Flexible	Co-Located	Remote	2D→MR	MR→2D	Bidirectional	Comp. Coupled	Free	Both	
Symmetric Mirror	7	6	1	5	2	2	0	4	4	0	3	3	1	7	0	5	0	2	1	3	3	7
Asymmetric Mirror	6	4	2	6	0	2	0	5	3	0	2	4	2	8	0	2	1	5	1	4	3	8
Logical Distribution	42	26	10	44	4	10	5	44	14	0	21	20	12	54	0	19	3	32	9	27	18	54
Remote Control	6	6	1	6	0	1	3	5	0	0	1	3	1	6	0	1	0	5	0	4	2	6
Dynamic Lens	6	1	0	6	0	1	1	4	2	0	4	2	0	6	0	4	0	2	2	0	4	6
Augmented Display	4	0	0	4	0	1	0	3	2	0	0	4	0	4	0	4	0	0	2	0	2	4
VESAD	16	3	0	15	1	2	2	12	3	0	3	13	0	16	0	13	0	3	10	0	6	16
Migratory Interface	7	6	5	11	0	0	1	10	2	0	1	4	6	11	0	0	1	10	1	3	7	11
	67	36	12	72	12	7	6	65	23	0	27	40	15	83	0	35	4	44	26	27	30	

Figure 3.5: The co-occurrences of codes between the *configuration* dimension across the dimensions *temporal*, *relationship*, *range*, *device dependency*, *space*, *interaction dynamics*, *anchoring*. Each row is shaded to show the frequency of dimension values for each *configuration*.

3. *Unidirectional (MR-centric)* interaction dynamics is only used in configurations that make use of the extended 3D capabilities of the MR interaction component (i.e., *asymmetrical mirror*, *logical distribution*, and *migratory interface*).

4. Although early work focused mainly on exploring complementary combinations in general *logical distributions*, the increase in hardware sophistication is reflected in an increase in configuration diversity: For example, *symmetric* configurations only appeared around 2016, while *remote control* configurations appeared around 2020. Since hybrid user interfaces address an ever-shifting window of opportunity of contemporary hardware capabilities, the choice of possibilities is determined by the available hardware. We expect that new hardware will lead to novel combinations, creating new hybrid user interface *configurations*.

We can also identify research opportunities by looking at gaps in current usage:

1. Hybrid user interfaces have been exclusively explored in *co-located* spaces. Although this can be partially attributed to the focus on *single-user* systems, several systems already demonstrate the potential of hybrid user interfaces for collaboration. Similarly, we can see a lack of systems within *public* range. In both cases, future research could explore the distinct roles of each interaction component in these settings (e.g., territoriality, establishing shared and private spaces).

2. Only few systems have explored how to interact with 2D interaction components using the MR interaction component (*MR-centric interaction dynamic*). We see the potential in either *public* scenarios (e.g., avoiding hygienic issues) or dynamically enabling remote interaction with 2D interaction components [173]: Here, a *symmetric* configuration may be useful to offer complementary interaction possibilities.

3. We attribute the lack of systems with *near* range to current MR hardware limitations. For example, the limited field of view makes it hard to augment smartwatches. However, *near* devices could provide a complementary interface to MR HWDs by providing at-a-glance information (cf. [41]) or touch interaction.

4. Since the *logical distribution* configuration represents a substantial amount of records in our survey, we see potential to further differentiate this configuration. Although such a fine-grained analysis exceeds the scope of this work, future hybrid user interfaces might help to reveal additional patterns.

3.4.3 Edge Cases

We discovered edge cases that were not unambiguously covered by our exclusion criteria (see Section 3.2.3). We categorize these edge cases into three themes, discuss how they fit into our taxonomy, and sketch potential future directions.

Collaboration with asymmetric devices. One of our exclusion criteria (see Section 3.3.2) was a collaboration with asymmetric devices (i.e., applications where users can interact only with one device in total). However, such systems can be situated within the cross-device taxonomy [56] and even be classified as "*hybrid*", with one user on a tablet and one in AR [144, 253, 297]. We believe that many of our *key characteristics* are not applicable to such systems. Although collaborative systems with asymmetric devices might still provide valuable insights for hybrid user interfaces, we argue that this would dilute our current focus on cross-device (as opposed to "*cross-user*") interaction. Instead, we see them as part of the research field of cross reality environments (see Section 2.4).

Reproducing reality and virtual reality. To avoid limiting our survey to currently available technologies, we also considered systems that simulate or reproduce reality, such as VST HWDs. Such device combinations are commonly used to overcome the limitations of current OST AR HWDs (e.g., increased field of view [292], avoiding different focal planes [154, 407]), evaluate systems in scenarios that are difficult to reproduce in reality [202] (e.g., supermarkets [107]), or even simulate hardware capabilities that were not feasible at the time of publication (e.g., (transparent) tablets [225, 380]). Although this can negate some of the benefits of hybrid user interfaces (e.g., high-resolution displays of current mobile devices are limited by the clarity of VST HWDs), their conceptual application design contains valuable insights for the complementary use of interaction components, regardless of the technology used. By extension, we consider combinations that employ a VR environment if they demonstrate complementary use of interaction components according to our *attributes*, such as using a tablet in VR [104, 375]. In contrast, we excluded papers that did not fulfill our *attributes*, such as ones that do not establish a mutual dependency between the virtual environment and the simulated device [107].

Tangible interaction without visual output. Mobile devices combined with AR HWDs have great potential for tangible interaction techniques that do not necessarily rely on the device's visual output [372]. This can be useful to extend the interaction design space. However, as they do not fulfill our *attributes* (i.e., standalone 2D interaction component), we excluded records that do not rely on the mobile device's screen. We see such systems in another subset of hybrid user interfaces, at the intersection between the broader field of hybrid user interfaces and *tangible interaction* [192], which map digital interactions to dedicated real-world objects.

3.5 How Are Mixed Reality Hybrid User Interfaces Used?

In this section, we take a closer look at how hybrid user interfaces are used throughout our literature corpus. We highlight findings of our corpus in terms of previously used *terminology* (Section 3.5.1), *use cases* (Section 3.5.2), *devices and combinations* (Section 3.5.3), *contribution types* (Section 3.5.4), and *evaluation strategies* (Section 3.5.5). Finally, we discuss our survey results and provide a *summary of insights* (Section 3.5.6).

3.5.1 Terminology

We extracted the terminology used by each publication to better understand the terms used to describe hybrid user interfaces. Overall, we found that the author's terminology touched the following areas: **Hybrid User Interface** 22, referring to the term introduced by Feiner and Shamash [122]; **Hybrid ‹*other*›** 16, referring to areas such as hybrid computing environment, hybrid display system, or hybrid setup; **Cross-Device** 17 and **Cross Reality Interaction** 8, relating to the adjacent research areas; **Multi-Device Interaction** 10, describing a setup related to cross-device interaction; **Augmented Displays** 4, predominately used by Reipschläger et al. [334]; **Transitional User Interface** 2, used in relation to cross reality collaboration or hybrid user interfaces. In addition, several records used a variety of **other terms** 8, such as *compound environment*, *display combination*, or *multimodal interaction*. Lastly, the rest had **no explicit terminology** 11.

3.5.2 Use Cases

We analyzed usage scenarios as described by the papers in our corpus and further clustered them into common use cases. The most prominent use case was the area of **visual analytics** 26, ranging from abstract visualization for immersive analytics [76] (e.g., [66, 82, 179, 229, 335, 345]), dashboards [394], user study analysis [186, 289], and scientific visualizations [4, 43, 259]. This is unsurprising, since concepts such as multiple-coordinated views are widely established and lend themselves well to configurations such as the *asymmetric mirror*. This also partially overlaps with the **medical** 4 domain, which uses hybrid user interfaces

		1991	1997	1999	2004	2005	2006	2007	2009	2011	2012	2013	2014	2015	2017	2018	2019	2020	2021	2022	2023	2024	
MR-enabling Comp.	AR OST HWD	1	2	2	1	2	0	2	0	0	0	0	0	2	1	4	1	6	9	6	7	4	50
	AR VST HWD	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	2	0	0	6	3	13
	VR HWD	0	0	0	1	0	0	0	0	0	0	1	1	0	0	1	2	1	1	2	1	2	13
	Smartphone	0	0	0	0	0	0	0	0	0	0	0	0	2	0	1	0	2	0	0	1	0	6
	Tablet	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	3	0	5
2D Components	Tablet	0	1	2	0	0	2	0	0	0	0	0	1	3	0	1	1	4	3	3	0	1	22
	Smartphone	0	0	0	0	0	0	0	0	0	0	1	0	3	0	0	0	4	6	0	4	3	21
	Desktop	1	1	1	0	2	0	0	1	0	1	0	0	1	0	1	1	1	1	4	3	5	24
	Tabletop	0	0	0	1	1	0	1	0	1	0	0	0	0	0	3	1	1	0	1	0	0	10

MR-enabling Components	Tablet	Smartphone	Smartwatch	Desktop	Laptop	Projection	Large Display	Tabletop	
AR OST HWD	13	14	1	14	4	3	5	8	50
AR VST HWD	3	4	4	2	1	0	2	2	13
VR HWD	3	3	1	5	1	1	0	1	13
Smartphone	2	2	0	0	1	0	0	0	6
Tablet	0	0	0	1	1	0	2	0	5
Spatial AR	0	0	0	1	2	3	0	0	4
Stereoscopic Projection	2	0	0	0	0	0	0	0	2
CAVE	0	0	0	2	0	0	0	0	2
	22	21	5	24	7	6	8	10	

Figure 3.6: (Left) Distribution of used device types for each year within our literature corpus (years without records were omitted). Here, the complete table is shaded based on the maximum value (i.e., 9) found in it. (Right) Amount of device combinations within our corpus across a selected set of MR and 2D interaction components. Rows are shaded to show the frequency of 2D and MR component combinations.

for 3D examination [4, 43, 259] and surgery [190]. Another popular use case is **3D modeling** 11 [333, 334, 375], in which we include 3D sketching [8, 104]. Several records present the area of **development and authoring** 7, including development toolkits [142, 289, 330] or programming-related tasks [34, 336, 343, 396]. We attribute several records to general **productivity** 11 tools, such as window management [122], extending desktop configurations [211, 309, 310], file transfer [358], or general user interface improvements [49, 205, 217]. We also found several records in domains such as **gaming** 4 [281, 386], **entertainment** 4 (e.g., music [113, 219], television [15, 207]), **collaboration** 6 [68, 131, 337, 354], or **text entry** 5 [20, 154, 225]. Lastly, the remainder had **study-specific** 6 or **other** 11 use cases.

3.5.3 Devices and Combinations

As publications in our corpus combine multiple interaction components (i.e., devices) in a complementary way, we recorded device technologies and their most common combinations (see Figure 3.6). For records that did not have specific terminology to describe their hardware (e.g., [380]), we used current terms based on the dimensions of the device.

In terms of MR interaction components, **AR HWDs** 63 were the most common device technologies, with a large part of systems using **OST AR HWDs** 50 and others using **VST AR HWDs** 13. For **handheld AR** 11, we differentiate between **handheld AR on smartphones** 6 and **handheld AR on tablets** 5. Other types of AR systems were less common, such as **projector-based spatial AR** 4, **stereoscopic projections** 2, and **CAVEs** 2. Lastly, since we also included VR environments, our corpus contained systems with **VR HWDs** 13.

For 2D interaction components, we observed a similar spread between available device technologies: Several systems used mobile platforms 55 such as **smartwatches** 5, **smartphones** 21, **tablets** 22, and **laptops** 7. Stationary interaction components 48 were almost as common, including **desktops** 24, **projectors** 6, **large displays** 8 (e.g., wall-displays), and **tabletops** 10.

Looking at the device combinations (see Figure 3.6 (left)) over the years, we can observe an increase in device variation as more new form factors become available. This proliferation has resulted in a variety of device combinations (see Figure 3.6 (right)), with a focus on HWDs, although other MR devices were also used. Most often, 2D interaction components were represented by handheld devices (e.g., smartphones), or desktop interaction components. The usage of **AR OST HWDs** was preferred for the described 2D interaction components. For **smartwatches**, there is a preference for **AR VST HWDs** 4, probably because their field of view is more suitable for this form factor (cf. [292]).

3.5.4 Contribution Types

We classified papers in our corpus following the definition of Wobbrock and Kientz [410] (see Figure 3.7). As some papers provide more than one contribution (e.g., an artifact that was used as an apparatus in an empirical user study), we distinguish between primary and secondary contributions and report the results accordingly (i.e., 15 / 4).

Artifact 56 / 8 contributions manifest new knowledge in a design-driven approach creating new systems, tools, and techniques. See Section 3.5.3 for further descriptions of, for example, device combinations used to create artifacts.

Empirical 20 / 9 contributions provide new knowledge in an evaluation-driven approach based on user studies. See also Section 3.5.5 for descriptions of evaluation strategies.

Theory 6 / 5 contributions improve existing concepts, creating frameworks. We consider thorough descriptions of design spaces as theoretical contributions.

Method 1 / 1 contributions create new knowledge that informs how researchers carry out their work.

Two papers in our corpus [229, 431] are each categorized with a single primary contribution and two equivalent secondary contributions: Both present a design space as their main contribution (theory), followed by a system (artifact) used as the apparatus of a user study (empirical).

No papers were classified as datasets, surveys, or opinion contributions (see Figure 3.7). This is not surprising, as we focused on actual systems, which resulted in most artifact and empirical contributions. Similarly, some of the papers that could fulfill the criteria of other types of contribution were not considered, as they typically did not meet the eligibility criteria of our survey. Five papers contributed a theoretical contribution (i.e., a design space [229, 431]), showing that design spaces can be a way of exploring this nascent research area without necessarily implementing an extensive system or study apparatus.

3.5.5 Evaluation Strategies

For papers that provided a primary or secondary empirical contribution, we coded their evaluation strategies as proposed by Ledo et al. [232] (see Figure 3.7). Most papers featured one empirical contribution, others combined multiple user studies with different strategies.

The most common evaluation strategies within our corpus were empirical user studies that examine how users interact with a system 55. In their cross-device taxonomy, Brudy et al. [56] further split this evaluation strategy into *qualitative and quantitative usage* and *informative (observational and elicitation)* evaluations: While **usage** 37 focuses on the usability and usefulness of the system and how it is appropriated [56], **informative** 18 evaluations involve studies that precede and inform the development of a system, involving users in the design process [56]. **Demonstrations** 16 are used to describe how systems are employed in an actual use case scenario but do not necessarily involve a real system implementation. In contrast, **technical performance** 4 evaluations (cf. *technical evaluation* [56]) focus on benchmarking an implemented system in terms of its technical capabilities. Lastly, several systems did not include any kind of evaluation 18. Similarly, we did not observe any kind of **heuristics** 0 evaluation (i.e., using guidelines to analyze usability), which could be attributed to the lack of appropriate guidelines for hybrid user interfaces.

3.5.6 Summary of Insights

We present four main insights from our literature survey.

Term Fragmentation. Our corpus shows that terminology in prior research is mostly evenly split between *hybrid user interfaces* 22 and *cross-device interaction* 17, while the emergent area of *cross reality* 8 is gaining traction. A large number of eligible records 35 in this survey were found in adjacent research areas, barely mentioning such *hybrid* device combinations and thus making it difficult to search for and identify relevant prior work. This fragmentation is also reflected in their specificity, ranging from rather broad (e.g., *cross-device interaction* 17, *multi-device interaction* 10) to very narrow (e.g., *hybrid ‹other›* 16) descriptions. While broad terms encapsulate a breadth of unrelated systems, overly specific denominators may lead to fragmentation and could impede the understanding of the field.

A similar fragmentation can be seen in the use of terminology to describe design dimensions. Although the cross-device taxonomy [56] provides an appropriate framework, we adapted many dimensions to better capture the unique design dimensions of hybrid user interfaces. Many useful terms that apply to hybrid user interfaces are hidden within artifact contributions (e.g., anchoring [330], dependency [431], configurations such as augmented display [334, 336], and VESADs [292]).

Years	Contribution							Evaluation						
	Artifact	Empirical	Theory	Method	Dataset	Opinion	Survey	Informative	Usage	Demonstration	Technical Evaluation	Heuristic	None	
1991	1	0	0	0	0	0	0	0	0	0	0	0	1	1
1997	2	0	0	0	0	0	0	0	0	1	0	0	1	2
1999	3	0	0	0	0	0	0	0	0	1	0	0	2	3
2004	2	0	0	0	0	0	0	0	2	0	0	0	0	2
2005	2	0	0	0	0	0	0	0	1	0	0	0	1	2
2006	2	0	0	0	0	0	0	0	1	0	1	0	0	2
2007	1	1	0	0	0	0	0	0	2	0	0	0	0	2
2009	2	0	0	0	0	0	0	1	0	1	0	0	0	2
2011	1	0	0	0	0	0	0	0	0	0	0	0	1	1
2012	1	0	0	0	0	0	0	0	1	0	0	0	0	1
2013	1	0	0	0	0	0	0	0	1	0	0	0	0	1
2014	2	0	0	0	0	0	0	0	1	0	1	0	0	2
2015	4	0	0	0	0	0	0	0	2	0	0	0	2	4
2017	1	0	0	0	0	0	0	0	0	1	0	0	0	1
2018	5	0	0	0	0	0	0	0	3	0	0	0	2	5
2019	2	0	1	0	0	0	0	1	2	0	0	0	1	3
2020	7	0	2	0	0	0	0	1	3	3	1	0	2	9
2021	5	4	1	0	0	0	0	2	6	3	0	0	1	10
2022	4	4	0	0	0	0	0	4	5	1	0	0	2	8
2023	4	8	1	1	0	0	0	6	5	2	0	0	1	14
2024	4	3	1	0	0	0	0	3	2	3	1	0	1	8
	56	20	6	1	0	0	0	18	37	16	4	0	18	

Figure 3.7: Distribution of primary contribution types and evaluation strategies over time within our corpus (years without records were omitted). Each cell within a row (i.e., year) is shaded based on how many papers presented this contribution or evaluation type in the respective year, showing the relative development over time (i.e., in a column).

Parallel usage is predominant. We found that *parallel* usage of multiple interaction components is by far the most predominant design choice. We think the reasons are twofold. (1) *Parallel* usage offers greater design possibilities. In contrast, *exclusive* (and to some extent *serial*) usage considers one interaction component at a time, thus limiting the design potential. (2) As the time between using different interaction components *exclusively* increases, it becomes harder to see the whole system as one coherent interface – and harder to classify. Systems with *exclusive* usage might be better described as *cross reality*, which is concerned with the general usage of multiple systems in different "*realities*".

Optical see-through is the prevalent hardware choice, despite its drawbacks. AR HWDs 63 were the most common MR interaction component. Although current VST HWDs 13 offer a wider field of view for digital content [292], the use of OST HWDs 50 was much more prevalent. We attribute this to (1) the unrestricted real-world field of view of OST HWDs greatly facilitating interaction with 2D interaction components; (2) using OST HWDs further emphasizes the complementary nature of hybrid user interfaces, as the addition of a secondary interaction component offsets the drawbacks of the AR HWD; and (3) VST HWDs having only recently matured enough to be used in conjunction with other devices (e.g., in terms of text legibility due to limited pass-through resolution). We also observed a steady increase in empirical contributions over the past years, indicating that the hardware is now mature enough to conduct studies that are not confounded by hardware restrictions.

Lack of collaborative systems. Our corpus shows a distinct lack of multi-user systems 19, especially within public spaces 0. Although there is a great deal of work in the field of computer-supported cooperative work (CSCW) for collaboration across asymmetric devices – which we intentionally excluded from our survey – we see great potential in the use of hybrid user interfaces in collaborative scenarios.

3.6 Chapter Conclusion

This chapter establishes a consistent terminology and conceptual understanding of hybrid user interfaces that combine 2D devices with MR-enabling devices, providing the foundation for the remainder of this thesis. Based on the historical usage of the term "*hybrid user interface*", we derived a working definition and identifying attributes for the structured literature survey. This, in turn, allowed us to establish a taxonomy of key characteristics of hybrid user interfaces that captures their diversity across a wide range of application domains.

The taxonomy provides an overarching framework for contextualizing the exemplars presented in Part II, providing a common foundation through which the exemplars can be connected. Although exploring the entire design space far exceeds the scope of this thesis, these exemplars offer in-depth investigations of key aspects. The exemplars thus serve as a foundation for deriving initial design principles to support fluid interaction in immersive analytics.

Additionally, this chapter highlighted emerging trends and challenges in the design of hybrid user interfaces. Our survey shows that immersive analytics is one of the most prominent and promising application areas, stressing the need for appropriate design principles. While these findings remain exploratory at this stage, Part III will integrate insights from this survey and the presented exemplars into a broader synthesis of design insights and research implications.

Part II

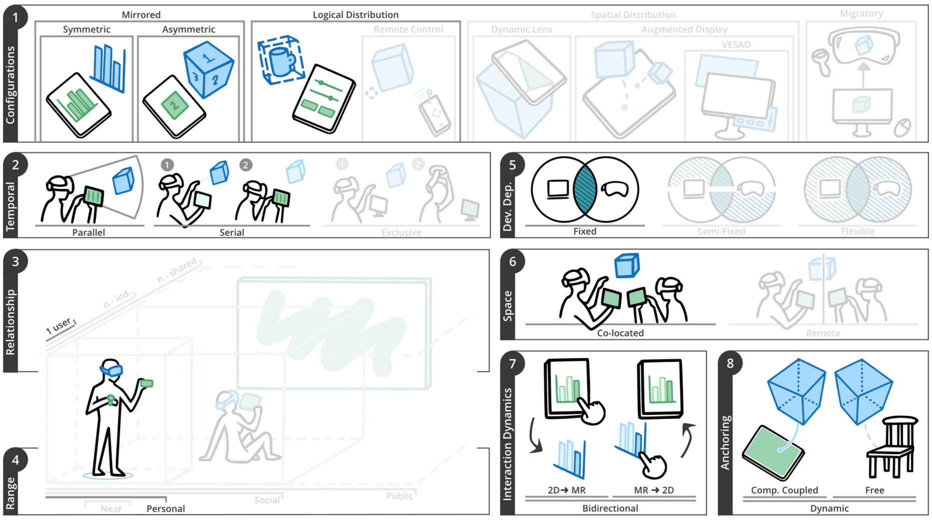
EXEMPLARS

Chapter 4

Exemplar: STREAM

Chapter Content

The following exemplar can be classified as:

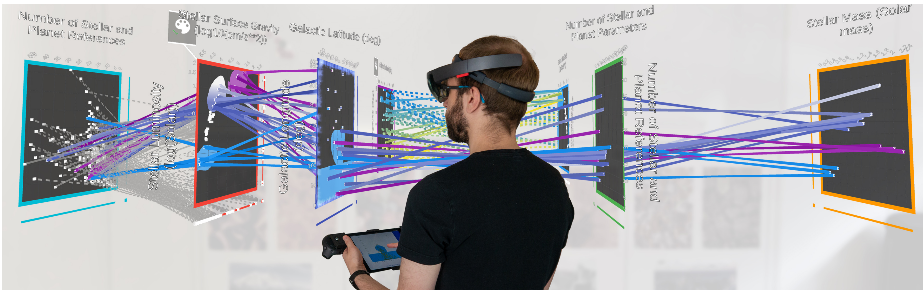


This chapter is based on the publications:

Sebastian Hubenschmid, Johannes Zagermann, Simon Butscher, and Harald Reiterer. "STREAM: Exploring the Combination of Spatially-Aware Tablets with Augmented Reality Head-Mounted Displays for Immersive Analytics." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '21. New York, NY, USA: Association for Computing Machinery, 2021, pp. 1–14. ISBN: 978-1-4503-8096-6. DOI: 10.1145/3411764.3445298

Katja Vock, **Sebastian Hubenschmid**, Johannes Zagermann, Simon Butscher, and Harald Reiterer. "IDIAR: Augmented Reality Dashboards to Supervise Mobile Intervention Studies." In: *Mensch und Computer 2021*. MuC '21. New York, NY: ACM, 2021, pp. 248–259. ISBN: 978-1-4503-8645-6. DOI: 10.1145/3473856.3473876

Note: The primary work described in this chapter was originally submitted as my Master's thesis at the University of Konstanz, titled "*MIDAiR – Multimodal Interaction for Visual Data Analysis in Augmented Reality*". Since this work represents a key example of the potential of hybrid user interface, this work was later published, with refined analysis and revised design insights. It served as my entry into the field of hybrid user interfaces that ultimately resulted in the research objectives of this thesis and is thus also included here as the introductory exemplar.



Supplemental Video

Figure 4.1: STREAM combines spatially-aware tablets with augmented reality head-worn device for visual data analysis. Users can interact with 3D visualizations through a multimodal interaction concept, allowing for fluid interaction with the visualizations.

4.1 Chapter Context

Research on hybrid user interfaces often emphasizes leveraging the strengths of one device to compensate for the limitations of another. This often involves focusing on specific input or output modalities, such as using a secondary device solely for text input [153, 154] or display quality [19]. In contrast, this chapter takes a holistic approach by exploring a hybrid user interface that utilizes the wide range of input modalities available when combining a tablet with an AR HWD. Given the hardware limitations of the HoloLens 1 at the time, we instead leverage the potential of contemporary tablets to track themselves in space.

The resulting STREAM (***S**patially-aware **T**ablets combined with Augmented **Re**ality Head-Worn Devices for **Im**mersive Analytics*) prototype focuses specifically on *multimodal* interaction with 3D visualizations (see Figure 4.1). We use an established 3D parallel coordinates visualization that consists of individual linked 2D scatter plots (cf. [66, 88, 89]), which is well-suited for demonstrating 2D interaction (e.g., configuring scatter plots) as well as 3D interaction (e.g., visualization layout), contributing to **RO2: Task Allocation**. Here, the spatially-aware tablet offers familiar touch interaction with the visualization itself, while the tablet's spatial-awareness and the HWD's head-gaze and egocentric navigation can be used for interacting within the 3D scene. To bridge the gap between tablet and AR environment, we developed a novel eyes-free interaction concept for fluid interaction [111], contributing to **RO3: Interaction Techniques**: This allows users to interact with the tablet while observing the immediate effects in the AR visualization, without incurring the cost of switching between different output modalities [151]. Furthermore, our prototype supports a seamless transition between the tablet's display and the AR environment by merging the AR visualization with the tablet's display, contributing to **RO1: Transitioning Between Devices**.

While STREAM was previously submitted as my Master's thesis and later revised for publication, it served as my (and thus, this thesis') prelude into hybrid user interfaces and ultimately resulted in the research objectives of this thesis. As a result, it is the only exemplar in this thesis that addresses all research objectives.

4.1.1 Research Questions

Since our prototype contains many unknown design elements with little prior research, we aimed to assess the feasibility of our concept and uncover initial usability problems. Thus, we evaluated our prototype in a user study to investigate three research questions:

> **RQ1.1 Use of Spatially-Aware Tablet.**
> How can a spatially-aware tablet complement interaction within an AR environment?
>
> **RQ1.2 Multimodal Interaction.**
> In what ways can a multimodal hybrid user interface facilitate fluid interaction in immersive analytics scenarios?
>
> **RQ1.3 System Usability.**
> Do the interaction techniques implemented in STREAM support effective and usable interaction for hybrid user interfaces?

Although hybrid user interfaces inherently involve multimodal interaction due to combining different devices, this chapter specifically explores how to fully leverage and incorporate available input modalities in one system, which enables unique interaction design opportunities.

4.1.2 Multimodal Interaction

There is a large body of research demonstrating the general benefits of multimodal interactions, such as better accessibility [301], better flexibility [230], and better task performance [302] as users can choose the best input modality for a given task, as well as more reliability, as users may fall back to alternative input modalities [300]. With regard to information visualizations, multimodal interaction can offer unique interaction opportunities [238] and make visualizations accessible to a broader audience [237]. Research has shown the benefits of touch interaction in immersive environments [73], such as greater reliability for fine-grained interaction, and explored the combination of touch with other input modalities, such as natural language input when interacting with a network visualization [340, 370], proxemic interaction by rearranging tablets to interactively build and manipulate visualizations [228], pen and speech for interacting with different visualization types [369], and tangible interaction with a spatially-aware tablet for performing 3D selections [35].

4.1.3 Eyes-Free Interaction

However, hybrid user interfaces may require users to switch between different output modalities (i.e., mobile device and AR HWD during *serial* temporal usage), which can be costly [151]. To address this, eyes-free interaction – allowing users to interact with a device (i.e., tablet) with minimal intrusion on their visual attention [419] – could leverage the benefits of the tablet as input modality without incuring transitioning costs. Although eyes-free interaction has been employed on a variety of devices (e.g., phones [25, 252, 294], wearables [52, 312, 412]) and has shown to be beneficial when interacting with large data visualizations via a smartwatch [173], we found no prior work investigating eyes-free interaction within hybrid user interfaces.

4.2 STREAM

We created STREAM (***S**patially-aware **T**ablets combined with Augmented **Rea**lity Head-**M**ounted Displays*) as a proof of concept to investigate the combination of spatially-aware tablets with AR HWDs for multimodal interaction with 3D data visualizations. We chose an established 3D parallel coordinates visualization that is easy to understand (cf. [66, 88, 89]), allowing us to study our interaction concepts without confronting users with both the complexity of a novel visualization and novel interaction concepts. Although our interaction concept relates to the unique challenges afforded by this particular 3D visualization, we are confident that our findings can be applied to other 3D visualizations as well (e.g., heatmaps). The following sections describe the *visualization*, *device responsibilities* of our tablet and AR HWDs, *interaction design*, and *prototype implementation*. Our prototype is also freely available as open source project[1].

4.2.1 Visualization

STREAM uses a 3D parallel coordinates visualization consisting of linked 2D scatter plots (see Figure 4.2), allowing users to interactively explore a multidimensional data set. Scatter plots can be individually placed anywhere within the AR environment and links can be established between any two scatter plots (barring circular connections). Each link represents a set of lines, and each line represents a multidimensional record from the data set that is shot through the scatter plots and manipulated by the same (in terms of position, color, visibility). Links in STREAM are directional (as indicated by particle effects beneath each link), allowing users to progressively reduce or combine subsets of the data, akin to a *stream* of data. To preserve line connections, missing values are placed in a dedicated area beneath each axis and are further highlighted with a dashed line.

Scatter plots manipulate the data (and thus, lines) by defining their position and color, but can also filter out items for subsequent connections. For color and filtering, users can toggle different *attributes* for each scatter plot and define *selections* within each scatter plot. Each *selection* has either one of eight predefined solid colors, or one of four predefined color gradients. Furthermore, each scatter plot has three *attributes*: *filter*, *colorize*, and *sort*. The *filter attribute* determines if a scatter plot's *selections* remove data items not contained within a *selection*

[1] https://github.com/hcigroupkonstanz/STREAM

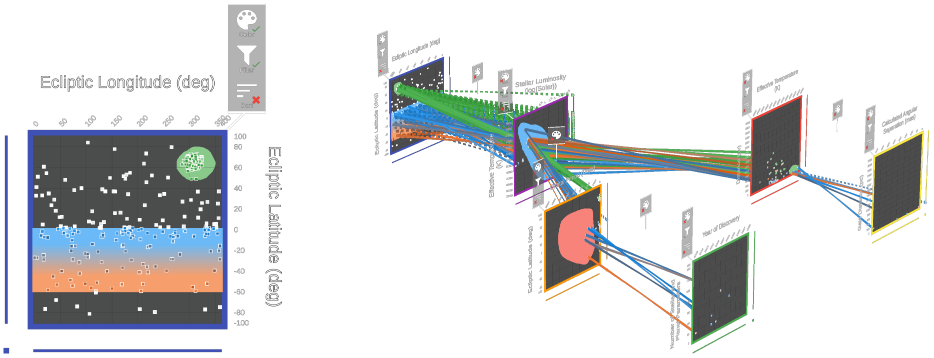


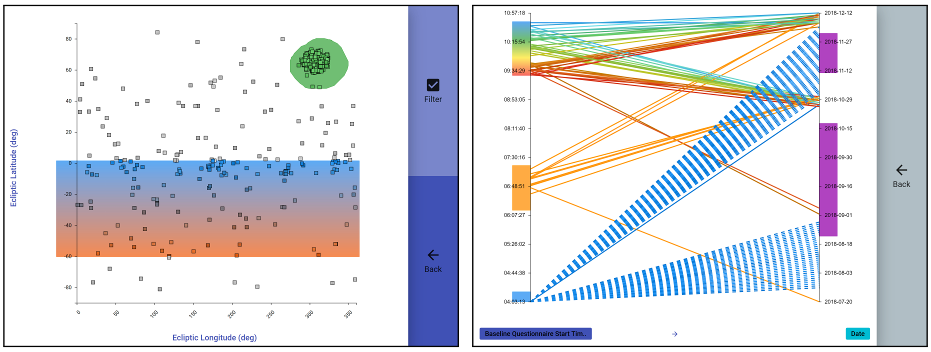
Figure 4.2: The visualization in STREAM is composed of linked 2D scatter plots. (a) A scatter plot contains separate areas beneath each axis for missing values. Data *selections* can filter or colorize the data, depending on the scatter plot's *attributes*, which are displayed in an info panel (top right). (b) Scatter plots can be linked together. Links are directional, meaning that the data *flows* from one scatter plot to the next (here: left to right), allowing users to progressively filter the data.

from subsequently connected scatter plots; the *color attribute* colorizes data items based on their current *selection* for *all* connected scatter plots (i.e., only one scatter plot can have the color *attribute* enabled), allowing for a linking & brushing approach [208]; lastly, the *sort attribute* discards the scatter plot's X-axis in favor of sorting the data items by their Y-axis values. Similarly, each *link* can also have the *color attribute*: When active, the data is colored based on the relative differences between the two connected scatter plots (e.g., green for increasing, red for decreasing values).

4.2.2 Device Responsibilities

To make full use of the AR HWD and the spatially-aware tablet, each component of our 3D visualization (i.e., scatter plot, link) has a representation in the 3D AR scene, and a 2D counterpart suitable for viewing on the tablet (see Figure 4.3). The 2D visualization always matches the user's perspective on the 3D visualization (e.g., 2D scatter plot is flipped horizontally when users looks at 3D scatter plot from behind). Both representations are synchronized in real time, thus offering fluid interaction across devices that leverages the unique benefits of each device (cf. [66]).

The *AR HWD* is suited for viewing and interacting with 3D visualizations, thanks to its stereoscopic output and egocentric navigation (cf. [222]). We therefore employ many of its available input modalities (i.e., head-gaze, egocentric navigation) for tasks that require 3D input (e.g., positioning of scatter plots). While the by default available mid-air gestures are also suited for 3D input, we chose to forgo mid-air gestures in favor of concentrating on the tablet interaction. In addition, our choices were also influenced by the technical restrictions of available



(a) (b)

Figure 4.3: 2D visualization views on the tablet. The back button allows users to return to the main menu. (a) When a scatter plot is selected, users can draw *selections* directly in the scatter plot, or click on the label beneath each axis to change dimensions. (b) When a link is selected, users see a simple parallel coordinates similar to looking at the selected link from the side. Here, users can create vertical *selections* on each side (i.e., scatter plot), or change the Y-dimension of the two connected scatter plots.

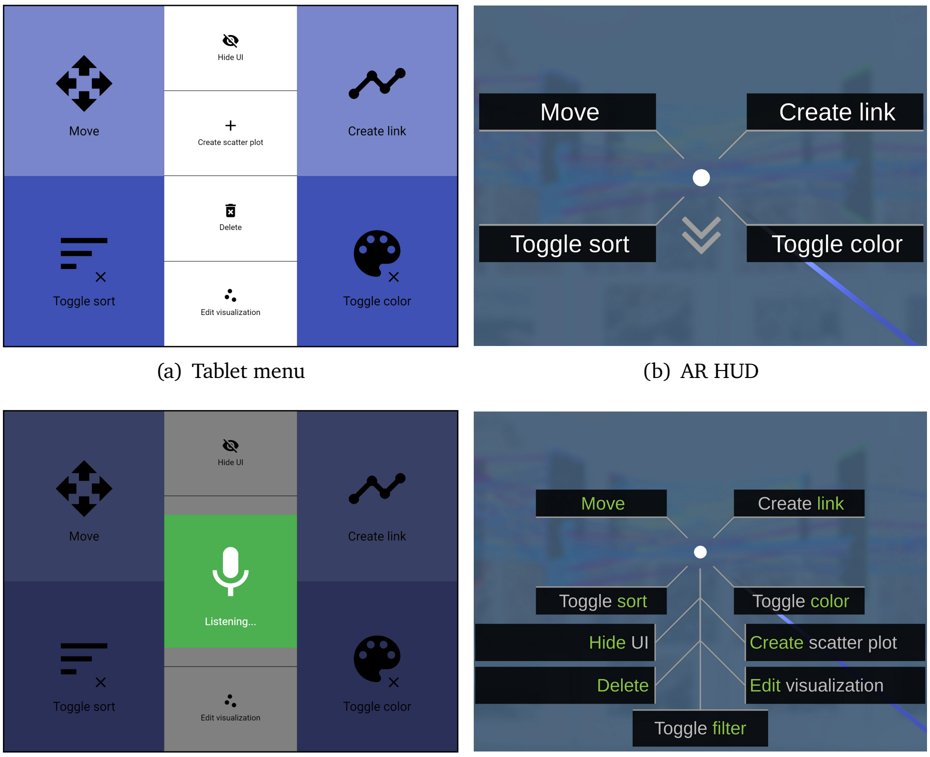
hardware (i.e., Microsoft HoloLens 1). Although we used head-gaze in this specific scenario, our concepts are responsive and work with both head-gaze and eye-gaze, depending on what is available.

In contrast to the AR HWD, the *spatially-aware tablet* excels in viewing and interacting with 2D information, thanks to its high-resolution display and touch-based interactions. It is therefore suitable for 2D interaction (e.g., creating *selections*), but can also assist in tasks that require 3D input, thanks to its spatial awareness. The 2D display can also be beneficial when investigating relative differences, as the simplified 2D representation (see Figure 4.3(b)) removes any perspective distortion that can occur in a 3D scene.

4.2.3 Interaction Design

The following paragraphs illustrate our interaction design to control the previously described visualization.

Selection. STREAM uses a selection-based interaction approach, meaning that users may only interact with one object (i.e., scatter plot or link) of the visualization at a time. To better distinguish between the objects, all scatter plots receive one of eight colors upon creation, as indicated by a colored frame. The current selection is highlighted by particle effects in AR and by the tablet's matching background color scheme (see Figure 4.4(a)) – or gray, if a link is selected. Furthermore, a line in the user's field of view points towards the currently selected object (see Figure 4.4(b)).



(a) Tablet menu

(b) AR HUD

(c) Tablet menu during active voice commands

(d) AR HUD during active voice commands

Figure 4.4: (a) The tablet menu contains up to four large buttons that can be clicked without looking at the tablet. Additional menu actions with increased interaction cost are placed in the middle. The tablet's background color matches the color of the selected object. (b) Users see the prominent items for eyes-free interaction on their AR HUD around their cursor, allowing them to click on the appropriate button on the tablet without looking down. A colored line points towards the currently selected object. (c+d) While voice commands are active, the AR HUD shows all available commands with their activation keyword highlighted in green.

The current selection can be changed through the user's head-gaze, as indicated by an AR cursor. This allows users to quickly select objects from far away. To mitigate the Midas touch problem, we use a long dwell time (3 s) until the selection is confirmed. An instantly-visible loading indicator will appear around the AR cursor and turn green once the selection is complete. Similar to Jacob [193], the selection process is further enriched by the tablet: During selection, the tablet will instantly display the color of the selected object. In addition, users can tap anywhere on the tablet to skip the dwell time (*gaze and commit*).

Eyes-free interaction. We employ a touch-based menu (see Figure 4.4(a)) for general system control actions (e.g., creating and deleting objects). To allow users to observe their actions in the AR environment and avoid the cost of display

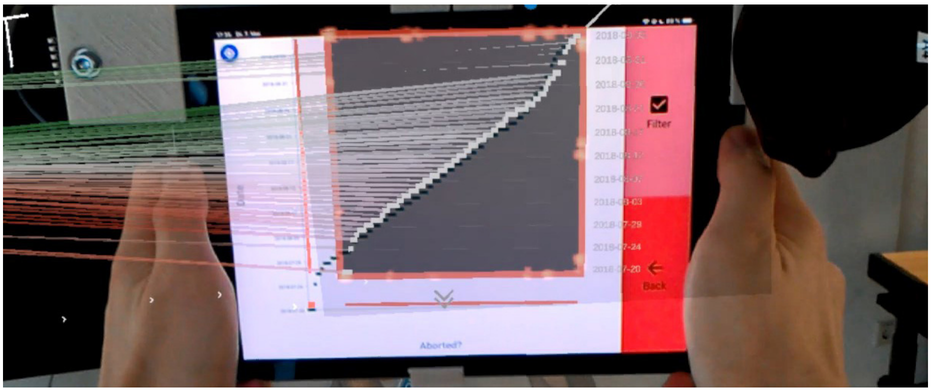


Figure 4.5: The *tablet lens* is activate while holding the tablet vertically. The scatter plot rotates to match the tablet's 2D view of the scatter plot.

switching [151], we employ a novel eyes-free interaction concept [419] consisting of two components: (1) The tablet menu is divided into four large prominent areas, one for each corner: Assuming that the user holds the tablet in both hands, each corner can be touched with the user's thumb without having to look at the tablet. (2) A head-up display (HUD) mirrors the actions of the tablet menu's four prominent areas (see Figure 4.4(b)), so that users do not have to memorize where each action is located on the tablet. Although this approach is restricted to at most four actions, future work could explore and compare alternative design solutions (e.g., scrollable menu, marking menus [294], incorporate touch input beyond the front touchscreen [231]). Additional actions are available in the middle of the menu, which require a display switch and thus have an increased interaction cost. We make use of this increased interaction cost for actions that require the user's attention (e.g., deleting scatter plot), or require the user to look at the tablet for further action (e.g., viewing 2D visualization).

Symbolic interaction. In contrast to eyes-free interaction where users focus on the AR environment, STREAM also offers symbolic interaction when using the tablet as output modality. This is mainly used for scatter plot manipulation when viewing a 2D visualization (see Figure 4.3), allowing users to directly create *selections* by drawing on the scatter plot visualization (e.g., encircling clusters) or drawing directly on an axis (e.g., selecting a range of values). This could also be extended to support 3D selections (cf. [35]). Furthermore, the symbolic interaction can be useful for assigning axis dimensions, where users can scroll through a list or use the on-screen keyboard to search for specific dimensions.

Tablet lens. To decrease to cost of switching displays between the AR environment and tablet for *symbolic interaction*, we make use of the tablet's spatial awareness as a kind of 2D *lens* into the AR environment: While holding the tablet vertically, the tablet automatically displays the 2D visualization of the targeted AR object, bridging the gap between tablet and AR environment. If a scatter plot is selected, the AR scatter plot automatically rotates to match the tablet's position (see Figure 4.5)

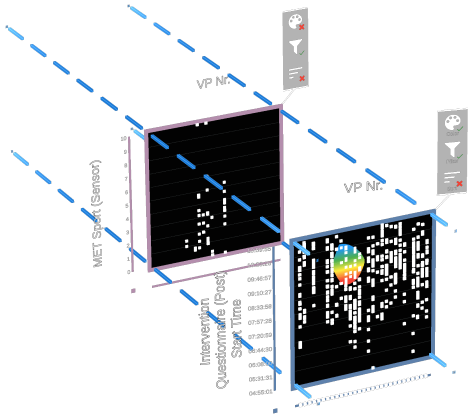


Figure 4.6: Scatter plots can be aligned to other nearby scatter plots. Dashed blue lines indicate the alignment range to the user.

This allows users to quickly make changes using *symbolic interaction* without losing the context of the AR environment. After activating the *tablet lens*, users can also hold the tablet in a more comfortable (i.e., angled) position to interact with the 2D visualization view. Once the tablet is put back into a level (i.e., horizontal) position, the *tablet lens* is deactivated and the main menu is shown.

Voice input. In addition to the eyes-free interaction, STREAM also offers voice commands as alternative input modality for system control. Although complex voice commands can be beneficial (cf. [369, 370]), they exceed the scope of this work. Similarly, many actions cannot be expressed as voice commands (e.g., drawing 2D *selections*), and therefore have no counterpart. To prevent accidental voice commands in potential collaborative settings, STREAM uses a trigger for voice command activation: Users need to hold two fingers on their tablet to activate voice commands. Once the system is listening, the AR HUD shows a list of all available actions with highlighted trigger word (see Figure 4.4(d)).

Scatter plot placement. Users can place individual scatter plots anywhere within the environment using multimodal interaction: The user's head-gaze determines the general position of the scatter plot, the distance can be adjusted through a drag gesture on the tablet, and the scatter plot's rotation is synchronized with the tablet's rotation. The placement can be confirmed by tapping on the tablet or via voice command. To facilitate the comparison of relative differences between scatter plots, users can also align scatter plots to each other: When a scatter plot is moved near an existing scatter plot, it aligns itself to the nearby scatter plot (see Figure 4.6). Dashed blue lines indicate the origin and range of the alignment – by moving the scatter plot away from these lines, the alignment is canceled. While the alignment is active, users can also hold a finger on the tablet to extend the alignment infinitely. Once the finger is lifted, the scatter plot is automatically placed at its current position.

Figure 4.7: For the spatially-aware tablet, we attached Vive trackers to an iPad with the help of a custom 3D printed mount. Two trackers provide stable tracking, regardless of the tablet's orientation. The tablet's screen demonstrates the *focus mode* and thus only shows a single button to minimize distractions for the user.

Linking. STREAM uses the *head-gaze and commit* input model to establish links between scatter plots. When a scatter plot is selected, users can create a new link via the main menu, starting from the selected scatter plot; a preview of the new link then follows the user's head-gaze in AR and snaps to valid scatter plots. Lastly, the connection can be canceled or confirmed either by touching the tablet, or via voice command. The links can also be inverted via the *system control* menu.

Focus mode. Because the AR interaction hints may be distracting when analyzing the data, STREAM offers a *focus mode* in the main menu which disables the *eyes-free interaction* by hiding both the AR HUD and removes any interaction elements from the tablet, save for a small central button to turn off the *focus mode* (see Figure 4.7). While in *focus mode*, users can still use voice commands for *system control* actions or use the *tablet lens* to quickly manipulate a scatter plot.

Proxemic interaction. STREAM also offers subtle, implicit proxemic interactions [17]: Small text only appear when a user is close; icons and text such as dimension labels will automatically rotate towards users in vicinity; and links will disappear if the user is standing inside them. We decided against any explicit proxemic interactions, as they may be hard to interpret for users [12, 228].

4.2.4 Prototype Implementation

We chose the Microsoft HoloLens 1 as AR HWD for STREAM, as it allows users to move around freely and allows users to quickly glance at their tablet. To compensate for the limited augmented field of view, the visualization size (20 cm) is a trade-off between fitting well within the user's sight, yet large enough to guarantee readability. For the tablet we chose an Apple iPad Pro 2017 (2048 × 1536 pixels, 9.7″, 437 g) with a custom 3D printed frame for mounting two HTC Vive Trackers 2018[2] (see Figure 4.7), resulting in a total tablet weight of 715 g.

[2]Due to infrared interference between the HoloLens 1 and the Valve Lighthouses 1.0 causing tracking issues, we used the Valve Lighthouse 2.0.

The software uses a client/server structure using TCP and WebSockets, allowing for real-time synchronization between all clients. We use a dedicated Windows 10 machine to process the Vive Tracker data to make the tablets spatially aware. To unify the different coordinate systems (e.g., HoloLens, Vive Trackers), we place a Vive Tracker on a visual marker and calibrate each client once during start-up (cf. [14]). We used Unity 2019.1 for developing the HoloLens application and processing Vive Tracker data, while the spatially-aware tablet runs as native web application. Our server was written in TypeScript and runs in a multithreaded Node.js v12 runtime. Due to limited hardware capabilities of our mobile devices (i.e., HoloLens, iPad), the server is responsible for data processing. Similarly, we employ GPU instancing to render all data items in AR using custom shaders that receive position and color data via textures. This also allows us to smoothly animate data items using compute shaders with little impact to performance.

4.3 User Study

We used STREAM to investigate in a user study how users can utilize the provided multimodal interaction capabilities. We thus focus on three research questions: (1) the *use of the spatially-aware tablet*; (2) the *multimodal interaction*; and (3) the general *system usability*. To narrow the scope of this study, we chose a single-user scenario with guided tasks that do not require preexisting knowledge of visual data analysis. Additionally, we intentionally chose non-experts to shift the evaluation focus towards our interaction concepts. We collected qualitative and quantitative data from each participant to gain insights into the specific usage of STREAM.

4.3.1 Participants

We recruited 8 participants (4 female, 4 male) from different backgrounds (e.g., economics, natural sciences, psychology, history) aged between 21–27 ($M = 24.13$, $SD = 2$) from the local university. No participant had any disability hindering their physical movement, no participant had any form of color blindness, and all participants were right-handed. We asked participants to rate their experience in different topics on a scale from 0–5 (no experience – very experienced). Participants had mixed prior knowledge concerning data visualizations and data analysis: 4 had moderate experience (3–4) with visualizations (with 2 participants having prior experience with 3D data visualizations), and 4 participants were moderately experienced (3–4) with data analysis tools (e.g., R, Microsoft Excel). Combined, only 3 participants had no experience in either data visualization or data analysis. Prior usage of AR and VR was similarly mixed: 4 participants have not used any AR application prior to this study, while 3 participants have not used any VR application. In total, 3 participants had no prior experience in either VR or AR applications. Although all participants used a smartphone on a daily basis, only one participant used a tablet on a daily basis.

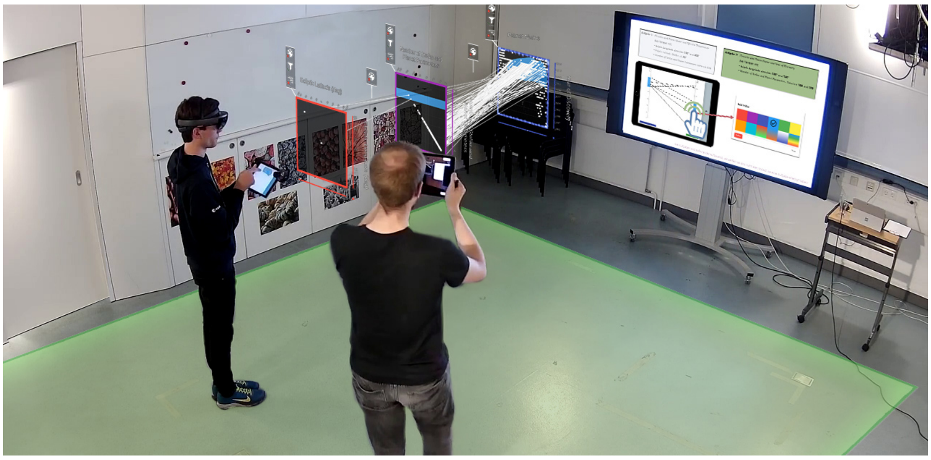


Figure 4.8: Our study setup provided users with ample space to move around (green, not visible during study). A large display was used to guide users through the first two tasks. The experimenter (right) could follow the participant's progress in AR through the *spectator viewer*.

4.3.2 Apparatus

The study took place in a spacious room with a walkable workspace of approximately 3 m × 2.5 m (see Figure 4.8). A large 84″ monitor was used to convey information about the application, and later display the participant's tasks during the actual study. In addition, a desk outside of the participant's workspace was used for filling out questionnaires and the final interview. During the study, participants wore a Microsoft HoloLens 1 and carried an Apple iPad Pro (9.7″) equipped with two HTC Vive Trackers (see Figure 4.7). In addition, the STREAM prototype was adjusted to focus on this study's tasks: Once less than ten data points are visible on a scatter plot, the tablet shows each data point's exact values in its 2D visualization; features that did not pertain to the tasks were removed (i.e., *focus mode*, filter and sort *attribute*), resulting in a simplified user interface. We also implemented a *spectator viewer* to allow the experimenter to view the participant's AR environment through an Apple iPad Pro (9.7″) (see Figure 4.8).

4.3.3 Tasks and Data Set

Due to our narrow focus on evaluating STREAM's interaction concepts with non-experts, we opted for an artificial workflow that force participants to interact with STREAM's features, with tasks that resemble the workflow of similar systems. We imported the NASA Exoplanet data set[3], as its high-dimensionality is a good fit for our visualization. To ease the burden on our participants and hardware, we reduced the data set to 500 randomly picked planets (from approx. 4000) and picked 22 dimensions (from 356).

[3] https://exoplanetarchive.ipac.caltech.edu/, last accessed 2025-05-12.

Each task required participants to create new scatter plots, choose appropriate dimensions and selections, and link scatter plots together to find the planet that matches the task's description. We used four tasks, all with a similar structure: "Find the *Planet Name* of the planet with the properties: *Ecliptic Longitude:* Between 100° and 200°; […]" Participants received step-by-step instructions to solve the first two tasks, and on-demand support for the latter two tasks. Although no time limit was imposed on participants, we chose tasks that were solvable in less than 30 minutes total, as prolonged wear of the HoloLens can cause discomfort [73].

4.3.4 Procedure

Participants were first welcomed and provided with introductory documents, containing information about the purpose of the user study and its procedure, a consent form, and a demographic questionnaire. The experimenter then explained the STREAM prototype and its visualizations to the participants, along with its interaction concepts (e.g., eyes-free interaction). Afterwards, participants were instructed on how to properly put on the HWD to ensure a comfortable wear and good visibility. After the application was started remotely by the experimenter and calibrated, participants received step-by-step instructions for the first two tasks by showing the participants visual instructions on a large display and monitoring their progress through STREAM's *spectator viewer* on a tablet. Once participants successfully completed the initial two tasks, they were given two similar tasks to solve on their own (without time limit). Afterwards, the experimenter asked participants to fill out a user experience questionnaire [352] and conducted a semi-structured interview. Participants then received monetary compensation for their time. In total, study duration ranged between 40–90 minutes ($M = 67$ min, $SD = 16$ min), with participants spending between 24–50 minutes ($M = 31.14$ min, $SD = 9.6$ min) using STREAM: Some participants went straight for the goal, while others playfully engaged with the application to solve the problem. One participant aborted the study after the second task due to simulator sickness, but still took part in the questionnaire and interview.

4.3.5 Data Collection

We collected audio data from a centrally placed microphone and video data from two opposing ceiling-mounted cameras for verification of participant behavior. Furthermore, the screen of both the participant's and the experimenter's tablet were recorded. We did not record any video from the HoloLens due to significant performance degradation. Instead, all application and interaction data was logged on our server (e.g., user position within the room, touch input coordinates), allowing for a complete reconstruction of the study. This data was supplemented by a demographic questionnaire before the study and the user experience questionnaire after task completion. To gain further qualitative insights, we asked participants to use the *Thinking-Aloud technique* during task completion and conducted a semi-structured interview afterwards.

4.4 Findings

We organized our observations (using the critical incidents technique), logging data, and user feedback using affinity diagramming to extract common themes. We structure these themes based on our research questions of (1) the *use of the spatially-aware tablet*, (2) the *multimodal interaction*, and (3) the *system usability*.

4.4.1 Use of Spatially-Aware Tablet

We specifically investigated the use of the tablet's *spatial awareness* and how participants used the *eyes-free interaction* concept. In addition, participants expressed concerns about *dropping the tablet* during use.

Spatial awareness. The tablet's spatial awareness was used sparingly when moving scatter plots: After placing the first scatter plot, the tablet's spatial awareness was quickly forgotten in favor of aligning scatter plots to each other. Participants also thought that the tablet had to match scatter plot exactly (cf. high degree of compatibility [27, 61]), and thus tried to hold the tablet vertically to rotate the scatter plot. Although one participant generally appreciated the rotation feature, the participant found the tablet uncomfortable to hold in certain positions (e.g., when rotating by 90°).

In contrast, the spatial awareness did play a significant role as spatial trigger (i.e., activating *tablet lens*): All participants often rotated their tablet briefly into a vertical position to switch to the tablet's 2D visualization view, thus circumventing a menu action (out of a total of 235 actions to activate the 2D visualization view, 127 were performed via touch and 108 via spatial trigger). Three participants even used this spatial trigger more often than the corresponding menu action. However, participants quickly put the tablet down into a position where the 2D visualization was still active, yet more comfortable to hold (see Figure 4.9).

Eyes-free interaction. Every participant reported that they understood the eyes-free interaction concept (i.e., interacting with the tablet without looking at it),

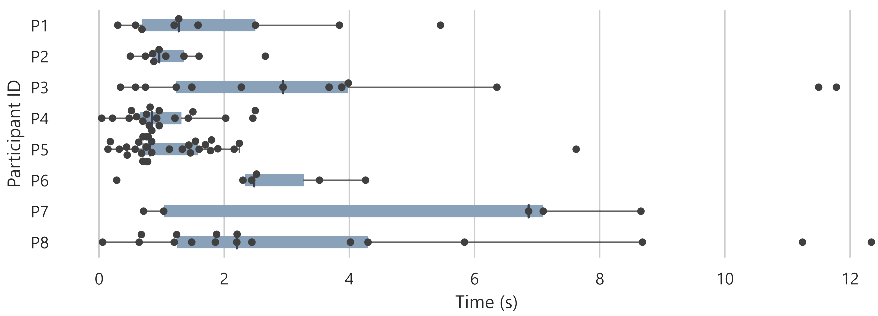


Figure 4.9: Boxplot of how long each participant held their tablet vertically during each use of the *tablet lens*. Vertical lines show quartiles, individual points show duration of each tablet lens activation.

although two participants admitted that they still looked at the tablet out of habit. Still, most participants appreciated the eyes-free interaction, especially due to the large buttons: "*Consciously I never had to look at the tablet [...] I thought the two big buttons on the left and right side were really intuitive and practical*" – P1. Yet, even with HUD indicators and large buttons, all participants did not feel confident enough to use the eyes-free interaction right away: "*I did actually use [the eyes-free interaction] once I knew [...] where each button is located*" – P7.

Fear of dropping tablet. Participants generally appreciated having something tangible to manipulate objects from a distance: "*That you can link this with your tablet, that's like a kind of remote control, that's really good*" – P8. However, four participants expressed concerns about accidentally dropping the tablet, especially when holding the tablet in one hand (e.g., while drawing *selections* or using the on-screen keyboard). Three participants found the tablet heavy when holding it in one hand, yet upon further inquiry they indicated that this fear was not related to the tablet's extra weight.

4.4.2 Multimodal Interaction

For the use of multimodal interaction, we investigated four topics: (1) whether our *choice of input modalities* for a given tasks were justified; (2) if and how *voice commands* were used; (3) if the multimodal *selection* of 3D objects posed any issues; and (4) how participants used the *AR environment*.

Choice of input modalities. Participants generally liked that they had the option of using several different input modalities: "*I also liked that you can speak, that you can do that with the [HoloLens], and with the tablet and all that*" – P6, and "*I liked [the different input methods], I think they were somewhat intuitive after a while*" – P5. Two participants also suggested alternatives, such as mid-air gestures for linking scatter plots. Due to the concerns of dropping the tablet, one participant suggested using smartphones or VR controllers for general interaction, but did prefer the tablet for interacting with the 2D scatter plots (e.g., when creating *selections*). Two participants also tried to place a new scatter plot in-between two connected scatter plots, expecting that the existing link would automatically adapt to include the new scatter plot. This may warrant further investigation into utilizing proxemic interaction for more natural input (cf. [12, 228]).

Voice commands. While all participants had to perform at least one voice command as part of the introductory task, only one participant continued to use voice commands as part of their workflow (out of 635 commands where both options were available, only 36 were performed with voice commands). Still, three participants appreciated that they at least had an alternative option available, but five participants stated that they strongly preferred touch over voice: "*I'm more of a haptic person, I want to grab things with my hands*" – P3 and "*I think touch is somehow more intuitive*" – P5. However, four participants were open to the idea of using more complex voice commands (e.g., as used in [73, 370]), especially if it saves time. Still, showing the keywords as part of the HUD proved beneficial: "*I liked that the [hints] were there [...] for a voice command that I didn't have present, I liked that I could look at it and there it is*" – P1.

Selection. Generally, participants found the selection through head-gaze easy to use, especially in simple visualizations with few selectable objects. While selecting scatter plots was straight-forward, selecting a link between two scatter plots caused problems for two participants, as their hitbox (i.e., selection target) was not apparent to the user. However, once an object was selected, it was immediately obvious to all participants due to both the particle effects in the AR environment as well as the tablet's matching background color: "*What I liked, that it became instantly clear that you're now on the green scatter plot, because your tablet's background was now green [...] I really liked that*" – P7.

To prevent the Midas touch problem, we employed a long dwell time with immediately-visible loading indicator. As a result, all participants reported that they never selected anything by accident, though two participants did find the immediate change of the tablet's background color distracting. Our logging data confirms that if an object was being selected, users almost always either completed the selection, or canceled the selection very early: 80% of canceled selection attempts lasted less than 0.8 seconds ($M = 0.48$ s, $SD = 0.59$ s). Yet, some participant did sometimes look away moments before the selection was completed, thereby canceling the selection accidentally. Two participants therefore suggested feedback (e.g., auditory) in addition to the existing visual feedback (i.e., indicator turning green) when the selection has been completed.

Lastly, we investigated if participants were able to skip the selection's dwell time by tapping on the tablet. Even though all participants had to skip the selection at least once as part of the tutorial, only three participants used it consistently during subsequent tasks. However, these three participants used this skipping method more frequently than waiting for the dwell time to complete: "*That was very cool, that I could tap on [the tablet] and then it directly finished loading, I liked that*" – P7.

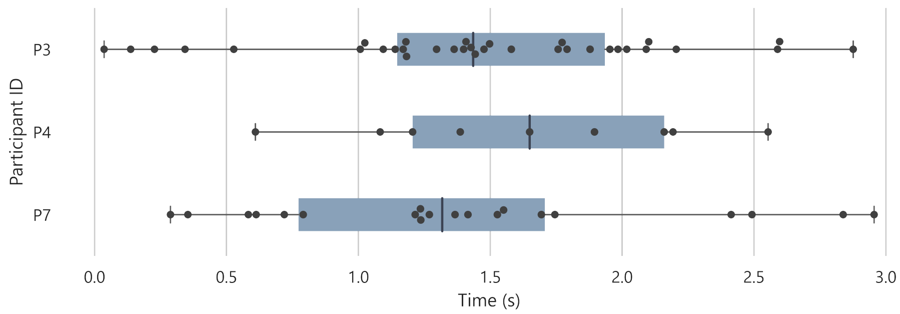


Figure 4.10: Boxplot showing when participants manually skipped selection, filtered for participants that skiped at least once. Vertical lines show quartiles, individual points show when a tap occurred.

Our logging data shows that manual skips occurred on average approximately 1.5 s after the selection started (M = 1.43 s, SD = 0.7 s, see Figure 4.10).

AR environment. Seven participants liked the use of AR, as they felt safe when moving around and felt an emotional connection to a workspace: "*You still have the feeling of somehow sitting in a workplace*" – P7. One participant suggested aligning virtual objects to real objects (e.g., walls), or put them into predefined anchors. Two participants also felt that a large, sterile room was necessary, as their own workspace was too small and too cluttered.

Participants generally made use of the spacious room to move around: Movement was sometimes explicit (e.g., participants stepping closer to zoom in) and sometimes implicit (e.g., slightly moving to select an object). Participants with more complex visualization (i.e., visualizations with more links and scatter plots) moved around more than users with less objects, suggesting that head-gaze alone may not be sufficient for selection in complex visualizations. The option to move around was, however, appreciated: "*I like that you can move around, because otherwise you just sit there. […] You can actively work with it*" – P3.

4.4.3 System Usability

Here we focus on the general system usability and on topics such as the *user experience*, the interaction with the *visualization*, and the use of space through *placement and alignment of scatter plots*.

User experience. Participants were positive about using the system and felt that it had "*much potential*" and that it was "*interesting*", "*cool*", but also "*complex*": "*What I liked, it reacted very quickly and was always there wherever I brought it with me*" – P6, and "*It was so much fun to go through these tasks with it*" – P1. This is reflected in the user experience questionnaire (ranging from -3 to 3), receiving high scores in the hedonic qualities (stimulation: M = 2.3, SD = 0.33; novelty: M = 2.13, SD = 0.42) and scoring well in attractiveness (M = 1.78, SD = 0.36).

Due to our artificial workflow, the task related quality aspects received mostly lower scores (perspicuity: M = 0.88, SD = 0.92; efficiency: M = 1.25, SD = 0.08; dependability: M = 1.13, SD = 0.29). Additionally, all participants required some time to get used to the system: "*As soon as you performed each action twice it actually was very intuitive and you could get a feel for it, I thought that was great*" – P7. This habituation period may be partially due to the "*many new terms*" some participants had to learn which were "*initially overwhelming*", but also due to the novelty of an AR HWD which most participants did not experience before and a novel interaction concept (i.e., eyes-free interaction with spatially-aware tablet). Two participants therefore valued the initial introduction (i.e., presentation and guided tasks), but wished for an even more interactive tutorial.

Visualization. The visualization was quickly understood by seven participants: "*First it looked very complex, but as soon as I understood the filter feature […]*

I thought it was actually very understandable" – P4. In our scenario, the 3D visualization served as a helpful overview, with three participants mentioning that it helped them keep track of each step and visually see how the data is reduced.

Participants also immediately understood the connection between the 2D visualization and its AR counterpart. While the tablet's size was appreciated when interacting with the 2D scatter plot, five participants did expect common touch gestures (e.g., *pinch-to-zoom*) to be available, which were not included due to an overlap with the voice command activation trigger: "*Intuitively I thought you could zoom in somewhere on the tablet*" – P4.

Scatter plot placement and alignment. As part of the tutorial, participants had to place the first scatter plot at a position of their liking, and subsequent scatter plots in a straight line to get familiar with the alignment feature. Participants instantly grasped the multimodal interaction concept (i.e., using head-gaze, egocentric navigation, and touch gestures) to position scatter plots. Although the alignment was not necessary for the study's tasks (as there was no comparison of relative differences), the alignment was very well-received: "*Especially [the alignment] was really good, because [...], once I opened a second layer, I could align the scatter plots with the ones next to them, and then it's clearly arranged*" – P1, and "*I was very, very happy about the alignment function*" – P7.

4.5 Insights and Implications

In this section, we reflect on our implementation and present *design insights* (D 1.1–D 1.6) to guide the development of spatially-aware touch devices in AR, as well as *research implications* (I 1.1–I 1.7) for topics that need further investigation. We structure our reflections based on our research questions of (1) the use of the *spatially-aware tablet*, (2) the *multimodal interaction*, and (3) the *system usability*.

4.5.1 Spatially-Aware Tablet

Although our interaction design and study tasks did not use the tablet's spatial awareness to its full potential, our results are still in line with previous research: (1) the tablet's physical rotation was not well-understood, uncomfortable, and overshadowed by the alignment feature (cf. [159]); and (2) our tablet lens, though not entirely used as we intended, was quickly adapted into the participant's workflow as spatial trigger (cf. [27]) **(D 1.1)**. Prior research has investigated a multitude of different spatial actions (e.g., tablet flip [375], holding a device at different angles for different actions [294]). Here, a study could investigate the feasibility of different spatial actions in terms of accuracy, fatigue, and speed **(I 1.1)**. Similarly, we chose a tablet as it provided a good trade-off between display size for symbolic interaction and mobility for egocentric navigation. This allowed us to interact with the 3D environment on a familiar 2D interface, which can also be applied to more general use cases (e.g., fine-grained control of object properties, cf. [380]). Since our design choice was informed by prior work on mobile devices in MR environments (e.g., [35, 375, 380]), we did not compare our approach against alternative devices (e.g., VR controllers, mid-air gestures) or device sizes. Further studies are necessary to compare different approaches for interacting with 2D visualizations on different devices regarding accuracy, fatigue, and task completion time **(I 1.2)**.

Furthermore, our eyes-free interaction concept allowed users to concentrate on the AR environment, while also providing some of the benefits of touch interaction, such as haptic feedback **(D 1.2)**. However, our eyes-free interaction technique has inherent design limitations, as it occupies most of display space for the large buttons and is limited to at most four different actions (one for each corner). Because our eyes-free interaction design represents only one of many possible alternatives, further studies are necessary to compare alternatives, such as touch gestures (e.g., [61, 173]) for different actions or touch gestures to control radial menus (e.g., [26]) **(I 1.3)**. Despite these limitations, the eyes-free interaction can be a good fit for general hybrid user interfaces, as it allows users to benefit from a physical interface (e.g., tablet) while concentrating on the AR environment.

Design Insights – Spatially-Aware Tablet

- **D 1.1** Employ spatial triggers (e.g., tablet flip) as explicit input
- **D 1.2** Use eyes-free interaction with mobile devices (e.g., tablets) for AR HWDs

Research Implications – Spatially-Aware Tablet

- **I 1.1** Comparison of different explicit spatial actions with tablets
- **I 1.2** Comparison of AR input devices and sizes
- **I 1.3** Comparison of eyes-free interaction techniques

4.5.2 Multimodal Interaction

AR devices offer many opportunities to employ multimodal interaction, which can be beneficial for offering alternative input methods (in our case, voice commands) and combining the advantages of different modalities to unlock more degrees of freedom (e.g., placing objects using head-gaze, touch gestures, and egocentric navigation). Although voice commands were underused in our case, using the tablet as eyes-free trigger to activate voice commands proved to be useful and can be beneficial for collaborative scenarios (i.e., prevention of accidental voice commands). In addition, AR HWDs allow for unintrusive, always-visible, and context-aware hints of which voice commands are available to the user **(D 1.3)** – which, in our case, extended the already available HUD for eyes-free interaction.

For the selection of different objects within the 3D visualization, we chose head-gaze combined with egocentric navigation, as this allowed for the selection of distant objects. While this worked well when only few objects were available for selection, users had to increasingly use egocentric navigation for more complex visualizations. This can be mitigated for example by selecting objects via touching the corner of our spatially-aware tablet (cf. [375]), but forgoes the advantages of head-gaze. Research could explore multimodal combinations that could, for example, allow users to select an occluded object from far away **(I 1.4)**.

Prior work (e.g., [193]) also advises for the use of dwell time to prevent the Midas touch effect. This works well in AR, as current AR devices employ an AR cursor to signify the exact position of the user's head gaze. This AR cursor can be useful to add contextual, unintrusive information, such as a loading indicator that informs the user of the current selection **(D 1.4)**. However, this indicator may give a false sense of security, causing users to look away before the selection has been completed; here, delaying the feedback of a completed selection by a few milliseconds can be beneficial.

During selection, the tablet performed two roles: (1) The tablet showed the color of the currently selected object, which was beneficial for some, but also distracting for others. (2) When selecting an object, users could tap on their tablet to instantly skip the dwell time. This skipping proved to be very beneficial, making the selection process much more responsive **(D 1.5)**.

Lastly, we also identified many opportunities for implicit proxemic interaction that can be helpful to make text within the 3D scene more readable (e.g., by rotating text towards the user). We decided against explicit proxemic actions on the basis of prior work, which found that proxemic zones may be hard to convey (cf. [12, 228]). However, these zones or interaction opportunities can be easily conveyed with AR, depending on the current context: For example, our alignment lines show the range in which the alignment is active. Further research could therefore explore the feasibility of different cues for proxemic interaction in AR **(I 1.5)**.

Design Insights – Multimodal Interaction

- **D 1.3** Display available voice commands in AR HUD
- **D 1.4** Use an immediately-visible loading indicator for selection in AR via head-gaze
- **D 1.5** Add explicit action to instantly skip dwell time

Research Implications – Multimodal Interaction

- **I 1.4** Investigate multimodal selection in complex 3D scenes
- **I 1.5** Explore the feasibility of visual cues for proxemic interaction in AR

4.5.3 System Usability

To narrow our focus on the use of spatially-aware tablets in AR, we chose a 3D visualization that was easily decomposed into 2D components for interaction with the tablet. This allowed us to split our tasks into 2D interaction (e.g., drawing data *selections*) and 3D interaction (e.g., scatter plot position), thus leveraging the strengths of multimodal interaction and especially touch-based interaction. Due to hardware limitations, we made the visualization size a trade-off between good legibility and fitting within the user's augmented field of view. However, this required a large room to make full use of the visualization. Future iterations of AR HWDs may become commonplace in the user's workspace; future research could therefore investigate how to make effective use of this limited space for organizing and interacting with visualizations in the user's workspace **(I 1.6)**. Although we could make use of navigation techniques (cf. [61]), this space could also offer new interaction opportunities, for example by using predefined anchors in the user's workspace.

Due to our choice of visualization, we also added an alignment feature, allowing users to compare data between two scatter plots without introducing errors due to position differences. While this feature was not necessary in our user study, it appealed to all of our participants **(D 1.6)**. This alignment feature has the potential of reducing clutter within an AR environment, even outside of the context of immersive analytics – an important yet often neglected topic.

Lastly, while we chose a single-user scenario to explore the feasibility of our interaction concept, we still added the *spectator viewer*, allowing us to provide guidance to users in AR. One participant expressed the potential for this kind of asymmetric collaboration: "*I found it really cool that you could look at what I was doing with my HoloLens by simply coming to me with a tablet, and that you did not have to also put on a HoloLens or boot up something else; but that you could simply look into it 'quick and dirty', I thought that was cool. Because assuming you just discovered something really great and you want to show it to someone else, then they can just look via the tablet*" – P7. Recent works have demonstrated the benefits of symmetric collaboration in MR environments for visual data analysis (cf. [66, 235, 335]), but the *spectator viewer* of our STREAM prototype also indicates the promise of asymmetric mixed device collaboration **(I 1.7)**.

Design Insights – System Usability

D 1.6 Support 3D object alignment in mixed reality environments

Research Implications – System Usability

I 1.6 Investigate effective use of space for AR visualizations (e.g., in small office workplace)

I 1.7 Investigation of asymmetric co-located collaboration (e.g., HWD and handheld AR)

4.6 Limitations and Future Work

To uncover initial usability problems and investigate the general feasibility of our interaction concept, we intentionally used an artificial workflow and non-expert users in this study. Thus, our results show the possible potential of using spatially-aware devices for interacting with a 3D data visualization in AR, but given the large amount of design parameters, further studies are necessary to investigate the advantages compared to other systems.

Many of the uncovered issues stem from the early state of our prototype: Most participants worried about dropping the tablet, which may be partially attributed to the heavy weight introduced by the spatial trackers. Since our results show that the spatial awareness was not used as much as we expected, we can now investigate more limited, yet lightweight solutions (e.g., internal sensors). Alternatively, we may compare our tablet-based approach against a smartphone-based approach to see if the handiness of the smartphone can offset the extra screen space of the tablet.

We also intentionally reduced the duration of this study to accommodate our users as the HWD can be very uncomfortable over longer periods. Since all of our participants needed some time to get accustomed to our system, this left little time to actually test out the system after getting used to its features. Given that future hardware iterations (e.g., Microsoft HoloLens 2) are more comfortable to wear, future studies should account for this habituation period by offering more extensive interactive tutorials.

4.7 Follow-up Research Trajectories

Our follow-up work on the IDIAR (**i**nteractive **d**ashboard **i**n **a**ugmented **r**eality) prototype similarly studies multimodal interaction in hybrid user interfaces, focusing on the specific use case of tracking mobile health (mHealth) [365] data using interactive dashboards (see Figure 4.11).

The design of IDIAR was guided by the needs of our collaborators in the SMARTACT[4] project (e.g., [124, 414]): Here, mobile intervention studies employ mobile devices to observe participants' behavior change over several weeks. Researchers regularly monitor these highly dimensional data streams to ensure data quality and prevent data loss (e.g., missing engagement or malfunctions [409]). The multitude of problem sources hampers possible automated detection of such irregularities. To this end, interactive dashboards can leverage human capabilities for the identification of patterns and irregularities [123, 287]. Still, visualizing the large amount of data stemming from high-dimensional data and large cohorts of participants (e.g., 20–40 participants) might raise the need for extra screen space.

IDIAR addresses these issues by providing interactive dashboards in immersive AR, allowing researchers to quickly identify irregularities within ongoing studies and intervene by sending messages to participants. The AR environment provides

[4] https://uni-konstanz.de/smartact, last accessed 2025-05-12.

Supplemental Video

Figure 4.11: Interactive dashboards in augmented reality (IDIAR) visualize data from an ongoing mobile intervention study, enabling researchers to quickly identify irregularities. Four dashboards provide users with information about an ongoing mobile intervention study. A multimodal interaction approach combining smartphone-based touch, head gaze, and voice input allows for familiar operation.

a larger and more flexible workspace in terms of scale and organization of dashboards, while still keeping the users' current work practices. IDIAR leverages a multimodal interaction approach combining smartphone-based touch, head gaze, and voice input.

We studied IDIAR in a user study with 15 domain experts from our SMARTACT[4] project. Although the user study and its findings are once again not part of this thesis, our analysis shows that the multimodal interaction helps in error recovery, while the familiar smartphone interaction allowed participants to understand the system quickly and provided a sense of immediacy, which was especially useful for writing text. Participants also appreciated the use of space afforded by both the AR environment and the organization of dashboards.

In summary, both STREAM and IDIAR highlight the value of exploring multimodal interaction in holistic application scenarios for immersive analytics. While this multimodal interaction approach was generally appreciated, input modalities that extended beyond conventional device capabilities offered limited utility. For example: Although STREAM allowed users to rotate scatterplots by physically rotating the tablet, this feature saw little use. Here, recent follow-up work has demonstrated more effective uses of such spatially-aware tablets, for instance by employing them as see-through lenses [225] in a *dynamic lens* configuration.

In contrast, the eyes-free interaction concept, integrated into both STREAM and IDIAR, remains underexplored in research thus far. Although early feedback highlights its utility for *serial* temporal usage, its effectiveness depends on careful design and may be hindered by user familiarity. As hybrid user interfaces become more common, there may be increasing value in fully leveraging available input modalities, even if the corresponding output of one device remains unused.

4.8 Chapter Conclusion

The “*STREAM*” exemplar investigates the use of multimodal interaction using a hybrid user interface for immersive analytics. Our eyes-free interaction technique enables seamless interaction between the tablet and the AR environment, leveraging the tablet’s form factor to place prominent actions into each corner that can be operated thanks to an AR head-up display. In addition, our interaction concepts allow users to bridge the gap between tablet and AR environment, enabling fluid interaction between the 2D and 3D representation of our visualization. Lastly, STREAM demonstrates the utility of multimodal interaction in hybrid user interfaces, allowing users to employ the appropriate input modality for a given task. STREAM therefore contributes towards all three overarching research objectives. While STREAM’s interaction techniques and findings are dependent on the specific context described in this chapter, Part III will provide a broader reflection and generalization of these insights.

Chapter 5

Exemplar: ARound the Smartphone

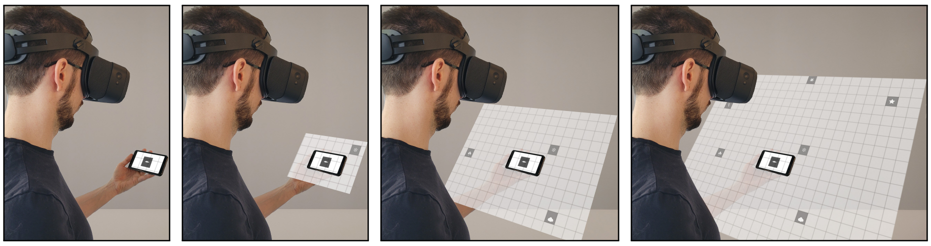
Chapter Content

The following exemplar can be classified as:

This chapter is based on the publications:

Sebastian Hubenschmid, Johannes Zagermann, Daniel Leicht, Harald Reiterer, and Tiare Feuchtner. "ARound the Smartphone: Investigating the Effects of Virtually-Extended Display Size on Spatial Memory." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '23. Hamburg Germany: ACM, 2023, pp. 1–15. ISBN: 978-1-4503-9421-5. DOI: 10.1145/3544548.3581438

Jonathan Wieland, Hyunsung Cho, **Sebastian Hubenschmid**, Akihiro Kiuchi, Harald Reiterer, and David Lindlbauer. "Push2AR: Enhancing Mobile List Interactions Using Augmented Reality." In: *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2024, pp. 671–680. ISBN: 979-8-3315-1647-5. DOI: 10.1109/ISMAR62088.2024.00082



(a) NO-AR (b) SMALL-AR (c) MEDIUM-AR (d) LARGE-AR

Supplemental Video

Figure 5.1: Exploring virtually-extended displays using a video see-through augmented reality head-worn display, comparing: (a) the NO-AR condition is used as baseline; (b) the SMALL-AR condition is comparable to a tablet; (c) MEDIUM-AR is similar to a desktop monitor; and (d) the LARGE-AR condition is equivalent to a television monitor.

5.1 Chapter Context

The previous exemplar highlights the potential of hybrid user interfaces in immersive analytics scenarios. Yet, in the absence of a suitable baseline, it remains difficult to quantify their actual benefits. With this chapter, we thus build upon the established concept of hybrid user interfaces by revisiting the original prototype introduced by Feiner and Shamash [122], which proposes virtually extending a physical display using an AR HWD (i.e., a VESAD configuration). Although this concept was introduced over 30 years ago, it continues to be relevant for contemporary hardware. However, the feasibility and quantitative benefits of such VESAD configurations have never been evaluated, despite the availability of an apparent baseline (i.e., no virtual screen extension). In contrast, prior work on physical screens has investigated the effects of different screen sizes on spatial memory, workload, and user experience [321, 423]. This chapter therefore replicates and extends prior studies in the context of a VESAD configuration that virtually extends the screen of a smartphone with an AR HWD (see Figure 5.1), thereby investigating the unique benefits and challenges of a hybrid user interface compared to more conventional setups (e.g., single displays [321]).

5.1.1 Research Questions

To better understand the relation between physical and virtual screens in hybrid user interfaces, this chapter investigates the impact of virtually-extended screen sizes through the complementary use of an AR HWD and a smartphone. We conducted a controlled laboratory experiment with 24 participants testing the effect of different sizes of virtually-extended displays on spatial memory, workload (e.g., ergonomics, cognitive load), and user experience, thereby contributing to the overarching research goal **RO1: Transitioning Between Devices**. This leads to the following three research questions:

RQ2.1 Spatial Memory.
How does screen size affect users' navigation behavior and performance in a *spatial memory* task?

RQ2.2 Workload.
Does a virtual screen extension cause an increased *workload* in terms of cognitive load and ergonomics due to increased context switching between the real and virtual display?

RQ2.3 User Experience.
In what way is *user experience* influenced by screen size?

Although we expect to find many similarities to prior studies on spatial memory (e.g., [321, 423]), hybrid user interfaces also face unique challenges that must be considered. For example: (1) By separating the display into a real and virtual display, users may have to split their visual attention [326], counteracting potential benefits gained from virtually increasing the display size [150]; (2) larger screen sizes may require more head-movement, which may affect ergonomics [70]; and (3) larger screen sizes cannot be fully kept in view or visually processed effectively, which could affect cognitive load. All of these may prove detrimental to spatial memory.

5.1.2 Spatial Memory

Yet, spatial memory is an important aspect of human cognition that has been well-studied in relation to HCI [350], especially in the context of reducing cognitive effort for navigation and search tasks [5, 317]. In this context, prior work has investigated different input (e.g., peephole navigation [206, 283, 305], body movement [132, 214, 320], mouse and touch [201, 382]) and output modalities (e.g., audio cues [136], tab interfaces [139]), display sizes [321, 423], visualization and memorization techniques (e.g., fisheye lenses [364], providing an overview [174, 198], focus+context [67, 293], storytelling [135]), as well as the use of landmarks such as gridlines [241], body parts [33], anchors and background images [390], in graphical user interfaces [389] and 3D environments [138, 247, 286].

Specifically, prior work indicates that – compared to indirect mouse input – direct touch interaction can improve memorization accuracy [382], spatial memory, and navigation performance [201]. Zagermann et al. [423] showed that embodied interaction can increase spatial memory when compared to indirect touch (e.g., via trackpad) and direct touch interaction, but at the cost of user experience and efficiency. This is especially relevant in the context of peephole navigation [271], as smartphones can be used for both static (i.e., touch) and dynamic (i.e., spatial movement, e.g., [305]) peephole navigation. In terms of peephole size, a study by Rädle et al. [321] shows that an increased peephole size can positively affect learning speed, navigation speed, and task load – albeit with diminishing returns. For AR, we also need to consider the virtual field of view. In this regard, a study by Caluya et al. [70] shows that a smaller virtual field of view has no significant impact on spatial memory, but can increase head movement.

Although an increased screen real-estate can be beneficial, prior work in multi-display environments also hints at several potential issues that may counter the benefits. One issue is the split attention effect [141], as users have to split their visual attention between multiple displays, resulting in overall worse performance [326]. In this regard, Rashid, Nacenta, and Quigley [325] provide an overview of different factors influencing attention switches in multi-display environments, such as display contiguity and angular coverage. In addition, a study by Nacenta, Mandryk, and Gutwin [288] shows that a physical gap between displays can significantly reduce performance. Yet, unlike prior methods of expanding screens, VESADs leverage AR to seamlessly extend the smartphone screen, thus eliminating any "*displayless space*" [288] and potentially avoiding the attention split between two displays. Still, Grubert et al. [150] and Eiberger et al. [106] observed significant overhead when switching between display output of an AR HWD and smartphone due to different focal planes, while Normand and McGuffin [292] observed no such overhead with video see-through HWDs. To the best of our knowledge, there is no prior research investigating these challenges for VESADs and their size.

In summary, prior work highlights the relation between peephole size and spatial memory. We expect that bigger peepholes (i.e., larger screens) perform better in terms of spatial memory, workload, and user experience (cf. [321, 423]), but may also negatively impact workload (i.e., head movement) as screen sizes exceed the HWD's field of view (cf. [70]). Here, a 2D environment in AR with direct touch interaction (e.g., via smartphone), gridlines, and visual anchors can strike a good balance in terms of clutter [84], efficiency [84, 423], and user experience [423].

5.2 Experiment

We aim to investigate whether reported findings on spatial memory can be transferred to the use case of hybrid user interfaces. On the one hand, larger display screens have been shown to be beneficial for spatial memory [321], especially when using touch interaction [423]. On the other hand, the split attention effect may negatively impact performance [72, 325].

Based on prior work, we therefore expect that bigger peepholes (i.e., larger screens) perform better in terms of spatial memory, workload, and user experience (cf. [321, 423]), but may also negatively impact workload (i.e., head movement) as screen sizes exceed the HWD's field of view (cf. [70]). We therefore also expect to see diminishing returns, as participants no longer profit from the increased screen size beyond a certain threshold (e.g., tablet-sized [321]). A 2D environment in AR with direct touch interaction (e.g., via smartphone), gridlines, and visual anchors can strike a good balance in terms of clutter [84], efficiency [84, 423], and user experience [423].

We conducted a controlled laboratory experiment using different virtual screen sizes to investigate the impact of virtually-extended display sizes on spatial memory, workload, and user experience. In addition, we also investigated possible differences in the split attention effect between different virtual screen sizes.

5.2.1 Conditions

We differentiate between three extension sizes (similar to Rädle et al. [321]) to mimic existing devices. In addition, we added a condition without any virtually-extended display as baseline condition. Lastly, participants wore an AR HWD in all conditions to ensure better comparability of conditions. In order of smallest to largest display size, we compared the following display sizes (see Figures 5.1 and 5.2):

- **NO-AR:** This condition serves as baseline with no virtual extension, using a display size of 5.5″.
- **SMALL-AR:** This condition mimics the size of a tablet, using a total display size of 11″.
- **MEDIUM-AR:** This condition is similar to current desktop monitors with a display size of 23″.
- **LARGE-AR:** The largest condition has approximately the size of a television monitor at 43″.

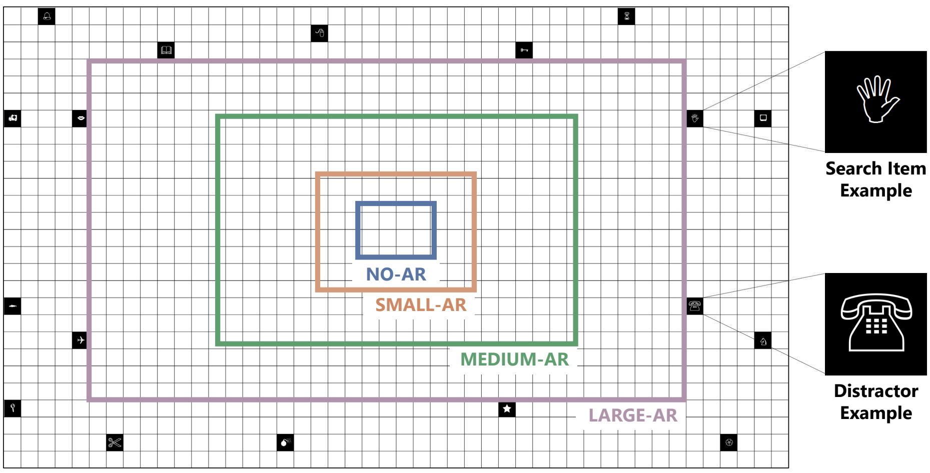


Figure 5.2: Exemplary 2D grid map used during the *navigation phase* containing search icons and distractor icons. A space in the middle of the map with the size of LARGE-AR was intentionally left blank to hide any icons in the starting position of each condition. Colored borders indicate the size and starting positions of each condition.

5.2.2 Tasks

To keep our results comparable to prior studies on spatial memory, we employed an established task [201, 214, 241, 283, 285, 286, 320, 423] which consists of a *navigation phase* and an *object location recall phase*. The task makes use of a 2D grid map (46 columns × 27 rows, see Figure 5.2) with an approximate real world size of 124 cm × 73 cm. While the visible area differed for each condition, the map size stayed consistent. All conditions and maps had a visible aspect ratio of 16:9.

Navigation phase. For the *navigation phase*, participants had to put on an AR HWD and were provided with a smartphone that was extended with a VESAD (depending on the condition, see Figure 5.1). Participants had to search for and navigate to a symbol on the grid map using touch panning gestures on the smartphone to move the map (i.e., using static peephole navigation [271]). To increase ecological validity, the navigation behaved similarly to off-the-shelf map applications by emulating physical inertia and stopping when the participant touched the phone again. For each search trial, the application started in its default position in the middle of the grid map (see Figure 5.2). The current search icon was shown as a semi-transparent symbol, which remained statically in the middle of the smartphone screen. The task was automatically completed once participants navigated to the icon and placed it approximately in the middle of the smartphone (i.e., once the map symbol's center touched the semi-transparent search symbol). In addition, we included a total of five item sets: an item set for the training task, showing letters of the alphabet and four distinct item sets (see Figure 5.2) to avoid learning effects between conditions. The locations of all icons were randomized on each map (i.e., in every condition) to prevent learning effects across conditions. Yet, we ensured that the length of navigation paths remained comparable across conditions and that there were no differences with regard to the complexity or theme of the icons (cf. [423]).

Participants were tasked to find a series of 6 different icons (with 4 repetitions each), which were distributed on the 2D grid map that also included similarly-looking symbols that served as distractors (cf. [423], see Figure 5.2). As shown in Figure 5.2, no icons were visible in the starting positions of all conditions and icons were not placed within an area equivalent to the size of LARGE-AR. With this, we ensured that, for example, participants do not visually scan the map in larger conditions before actually interacting with it, which might impair the comparability of conditions (e.g., navigation path lengths and task completion times). Each condition therefore consisted of 24 search trials, resulting in 96 trials per participant and a total of 2304 trials over all participants.

Object location recall phase. For the *object location recall phase*, participants sat in front of a desktop PC with a mouse, reducing potential influences of motoric or kinesthetic memory to increase the internal validity of the spatial memory measurement, as a common practice for studying spatial memory (cf. [201, 423]).

Here, participants were first presented with an empty grid map. In this phase, the entire map was visible on the screen and no navigation was possible. Instead, participants had to place the icons from the *navigation phase* in their prior location by clicking on the corresponding position on the map using the mouse. The current icon was shown at the top of the screen. The icon order was based on the search order from the *navigation phase*.

5.2.3 Measurements

We used quantitative and qualitative metrics to address our research questions.

Spatial Memory. To measure the impact on spatial memory, we measure the path length, task completion time, and navigation speed during the *navigation phase* as well as the icon placement recall accuracy during the *object location recall phase*. For better comparability across conditions, we use the normalized path length, which is calculated as a ratio between the participant's actual path and the shortest possible path using Euclidean distance. Here, we omitted the first repetition from our analysis due to the initial randomness during the first navigation [201]. The task completion time was logged as duration in seconds between the start of a repetition until the icon was found and placed in the middle of the smartphone. Since navigation speed can be derived from the path length and task completion time, we complemented this measurement by recording the maximum navigation speed to investigate whether larger display screen sizes allow participants to flick through the map more quickly. Lastly, the icon placement accuracy was measured in the Euclidean distance in pixels between the icon's actual position and the position where participants placed the icon during the *object location recall phase*.

Workload. To evaluate the objective workload, we measure the pupil size [105], the total amount of gaze movement, total degree of head movement during navigation, and the subjective task load. The pupil size was measured using a built-in eye-tracker which logged a relative value between 0–1 according to the pupil size range detected by the AR HWD. Here, a larger pupil size is seen as an indicator for increased task load [105]. The total amount of gaze movement and total degree of head movement were both calculated in the quaternion distance (i.e., angles) between each data point and divided by the duration until the next data point. We also used the raw NASA TLX [164] which allows us to measure subjectively perceived task load. Lastly, we traced the participants' gaze onto the virtually-extended screen to measure how much time participants spent looking at which screen (i.e., smartphone or virtual screen).

User Experience. We used the user experience questionnaire (UEQ) [352] to gain more insights into the attractiveness, hedonic qualities, and pragmatic qualities of each condition. We complemented these results with a semi-structured interview at the end of the study session to gather qualitative insights into participants' preferences.

5.2.4 Apparatus

For all navigation tasks, we employed a Varjo XR3 as video see-through AR HWD due to its high digital field of view (155° horizontal field of view; 90 Hz refresh rate; 12 megapixel video pass-through per eye, 100 Hz eye-tracker) attached to a state-of-the-art computer (Intel i9 9900K, Nvidia RTX 3090). We intentionally decided against an optical see-through HWD to avoid potential issues with different focal planes [106, 150, 292]. The information landscape was overlayed on top of a Google Pixel XL (5.5″, 2560 × 1440 pixel, Android 10). The smartphone was cut out from the digital overlay, allowing participants to still fully see their hands and the smartphone's display and its content. We also reduced the transparency of the entire map to 40 % so that participants were still able to make out their physical surroundings. The AR HWD was tracked using four Valve Base Stations placed in every corner of the room, while the smartphone was tracked with a fiducial marker mounted to the smartphone and tracked via the HWD's front-facing cameras. We used a stabilization algorithm to avoid inaccurate smartphone tracking when participants were not looking at the smartphone (e.g., for LARGE-AR). The object location recall task was performed using a desktop PC on a 4K 27″ monitor.

The software for all devices was implemented in Unity 2021.1 and is available as open source project[1]. The applications communicated through a client-server architecture using TCP. We connected all devices via 5 GHz Wifi or ethernet to reduce latency – ensuring that there was no perceivable latency between smartphone display and AR overlay.

5.2.5 Participants

We recruited 24 participants (10 female, 14 male) aged 21–36 (M = 24.4, SD = 3.1) from the local university. Participants were recruited through flyers that advertised an AR study about memory games. We recruited participants who were fluent in the local language to avoid potential differences in the linguistic meaning of different icons (cf. [201, 241, 423]). 22 participants were undergraduate students from different fields (e.g., computer science, social studies, history, biology, life science, law), 1 participant was a PhD student, and 1 participant was administrative staff. Although participants were mostly experienced in the use of smartphones (M = 4.375, SD = .824, on a Likert scale from 1 (inexperienced) to 5 (experienced)) and all participants owned a smartphone (n = 24), experience with AR applications was mixed (M = 2.75, SD = 1.327, on a Likert scale from 1 (inexperienced) to 5 (experienced)). All participants had normal (n = 12) or corrected to normal (n = 12) vision.

5.2.6 Procedure

Participants first signed a consent form, completed a demographic questionnaire, and received an introductory presentation about the task and guidelines for wear-

[1] https://github.com/hcigroupkonstanz/ARound-the-Smartphone

ing the HWD to ensure correct eye-tracking calibration. We assigned each participant a different order of conditions using full counterbalancing to avoid any learning effects. In each condition, participants started by putting on the AR HWD and received a smartphone. During all *navigation phases*, participants remained seated at a table and held the smartphone in landscape orientation. Participants started in the *navigation phase* where they first solved a training task until they felt comfortable with the system. Next, participants solved 4 repetitions of finding 6 different icons. After the *navigation phase*, participants took off the AR HWD to use a desktop system, where they completed the *object location recall phase* using mouse input. At the end of each condition, participants filled out a raw NASA TLX [164] and a UEQ [352]. We concluded each session with a semi-structured interview about topics such as memorization strategies, subjective preferences, and preferred display sizes. The study duration ranged between 40–70 minutes, and all participants received monetary compensation for their time. We also awarded an additional monetary reward to the fastest participant to further encourage participants to perform the tasks as quickly as possible. We followed all necessary ethical and sanitary guidelines provided by the local university.

5.3 Results

In this section, we present our results based on our three research questions of *spatial memory*, *workload*, and *user experience*. Since a Shapiro-Wilk test revealed that our data did not follow a normal distribution, we analyzed the data with a non-parametric approach. We used a Friedman test followed by a pairwise Wilcoxon test with Bonferroni correction as post-hoc analysis to test for statistical significance, where appropriate. We indicate the medians (*Mdn*) and standard deviations (*SD*) using subscripts $_{NO}$ for NO-AR, $_{S}$ for SMALL-AR, $_{M}$ for MEDIUM-AR, and $_{L}$ for LARGE-AR to improve readability. We assume $\alpha = .05$ for statistical significance. For pairwise comparisons, we adjusted significance values by the Bonferroni correction for multiple tests. The user study data is available in a data repository[2]. To improve readability, we report results of statistical analyses as tables in the appendix.

5.3.1 Spatial Memory

We measure spatial memory based on the *navigation path length*, *task completion time*, and *recall accuracy*. As a related measure, we also investigated the *navigation speed* during the *navigation phase*. Our findings are summarized in Figure 5.3.

Navigation Path Length

We found statistically significant differences in each repetition when comparing normalized navigation path length across conditions. Pairwise post-hoc com-

[2]DOI: 10.18419/darus-3326

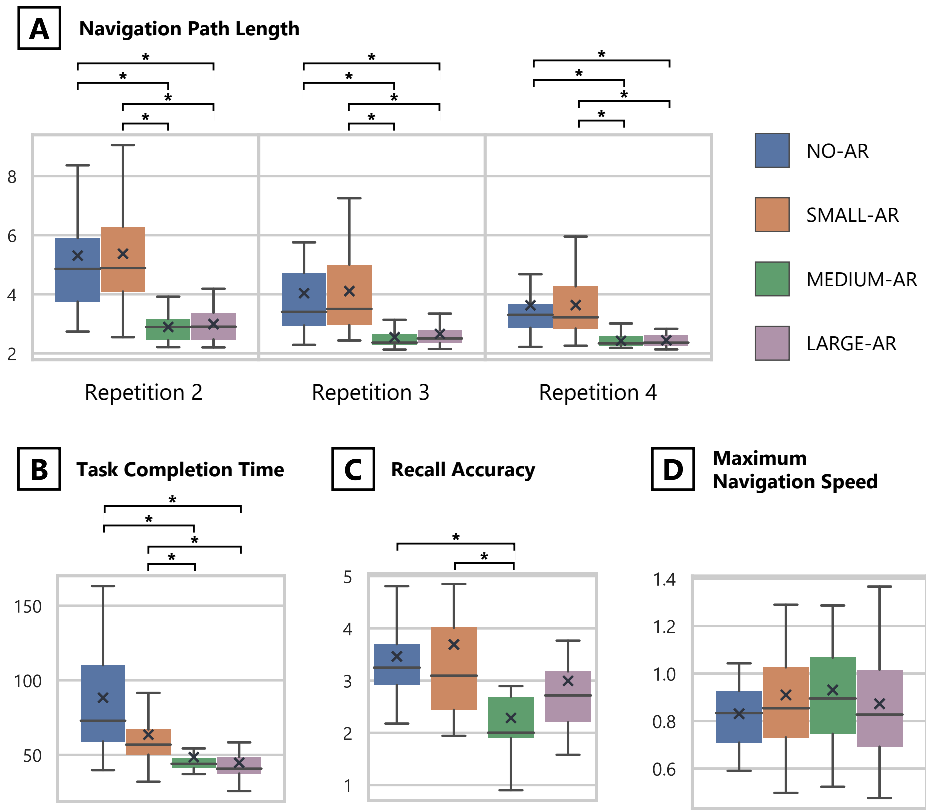


Figure 5.3: Summary of measures for *spatial memory*. (A) Boxplots of pairwise comparison of normalized navigation path length during the *navigation phase* (optimal normalized path length is 1). We omitted Repetition 1 due to the randomness of initial navigation trials [201]. (B) Boxplots showing average task completion time in seconds during the *navigation phase*. (C) Boxplots showing participants' accuracy of the *object location recall phase* in average Euclidean distance in blocks between the actual and the recalled position of the sign. (D) Maximum navigation speed in cm/s.

parisons showed that LARGE-AR (Mdn_L = 2.629, SD_L = .365) and MEDIUM-AR (Mdn_M = 2.532, SD_M = .339) had consistently shorter navigation path lengths than NO-AR (Mdn_{NO} = 4.168, SD_{NO} = 1.317) and SMALL-AR (Mdn_S = 3.92, SD_S = 1.274) throughout all repetitions (see Tables 5.1 and 5.2 and Figure 5.3 (A)); pairwise comparisons between NO-AR and SMALL-AR or MEDIUM-AR and LARGE-AR did not show any statistically significant differences. Comparing normalized navigation path lengths between repetitions within each condition showed a decrease in path lengths: Here, the overall test showed statistically significant differences for all conditions with individual differences for each condition, when comparing the first with following repetitions (see Tables 5.3 and 5.4). Only MEDIUM-AR showed a significant difference between Repetition 1 and 2.

Table 5.1: Results of the Friedman's test for navigation path lengths between repetitions, as visualized in Figure 5.3 (A). Statistically significant entries are marked with a star*.

Repetition	Result
Repetition 2	$\chi^2(3) = 44.10$, $p < .001$*
Repetition 3	$\chi^2(3) = 44.50$, $p < .001$*
Repetition 4	$\chi^2(3) = 53.55$, $p < .001$*

Table 5.2: Results of the pairwise comparison of navigation path lengths between repetitions, as visualized in Figure 5.3 (A). Statistically significant entries are marked with a star*.

Comparison	Repetition 2	Repetition 3	Repetition 4
NO-AR ↔ MEDIUM-AR	$z = 1.750$, $p < .001$*	$z = 1.750$, $p < .001$*	$z = 1.958$, $p < .001$*
NO-AR ↔ LARGE-AR	$z = 1.750$, $p < .001$*	$z = 1.583$, $p < .001$*	$z = 1.708$, $p < .001$*
SMALL-AR ↔ MEDIUM-AR	$z = 1.750$, $p < .001$*	$z = 1.917$, $p < .001$*	$z = 2.125$, $p < .001$*
SMALL-AR ↔ LARGE-AR	$z = 1.750$, $p < .001$*	$z = 1.750$, $p < .001$*	$z = 1.875$, $p < .001$*

Task Completion Time

We found statistically significant differences ($\chi^2(3) = 48.05$, $p < .001$) when comparing task completion times across conditions (see Figure 5.3 (C)). A pairwise post-hoc comparison reveals that NO-AR ($Mdn_{NO} = 72.883$, $SD_{NO} = 41.538$) is significantly longer than both MEDIUM-AR ($Mdn_M = 44.01$, $SD_M = 16.569$, $z = 1.875$, $p < .001$) and LARGE-AR ($Mdn_L = 40.782$, $SD_L = 13.11$, $z = 2.333$, $p < .001$). Similarly, SMALL-AR ($Mdn_S = 56.884$, $SD_S = 23.817$) is also significantly longer than MEDIUM-AR ($z = 1.542$, $p < .001$) and LARGE-AR ($z = 1.083$, $p = .022$). No statistically significant differences were found between NO-AR and SMALL-AR, or MEDIUM-AR and LARGE-AR.

Recall Accuracy

We found statistically significant differences in the object recall accuracy during the *object location recall phase* ($\chi^2(3) = 14.121$, $p = .003$, see Figure 5.3 (B)). A

Table 5.3: Friedman's test for learning effects between repetitions for each condition. Statistically significant entries are marked with a star*.

Condition	Result
NO-AR	$\chi^2(3) = 39.35$, $p < .001$*
SMALL-AR	$\chi^2(3) = 40.35$, $p < .001$*
MEDIUM-AR	$\chi^2(3) = 42.65$, $p < .001$*
LARGE-AR	$\chi^2(3) = 42.95$, $p < .001$*

Table 5.4: Pairwise comparisons of learning effects between repetitions for each condition. Statistically significant entries are marked with a star*.

Condition	Repetition 1–2	Repetition 1–3
NO-AR	$z = 0.708$, $p = .344$	$z = 1.667$, $p < .001$*
SMALL-AR	$z = 0.500$, $p = 1$	$z = 1.542$, $p < .001$*
MEDIUM-AR	$z = 1.167$, $p = .010$*	$z = 1.958$, $p < .001$*
LARGE-AR	$z = 0.958$, $p = .061$	$z = 1.750$, $p < .001$*

pairwise post-hoc comparison shows that NO-AR ($Mdn_{NO} = 3.248$, $SD_{NO} = 1.026$, $z = 1.12$, $p = .015$) and SMALL-AR ($Mdn_S = 3.095$, $SD_S = 1.804$, $z = 1.188$, $p = .009$) are significantly less accurate than MEDIUM-AR ($Mdn_M = 2.003$, $SD_M = .873$). While MEDIUM-AR performed best on average, no significant differences could be found between LARGE-AR ($Mdn_L = 2.715$, $SD_L = 1.199$) and MEDIUM-AR.

Navigation Speed

For navigation time, we divided *normalized navigation path length* by the *task completion time*. Although both measures already showed statistically significant differences between conditions, we still include the results here to substantiate our findings. Since we found statistically significant differences in navigation speed ($\chi^2(3) = 49.35$, $p < .001$), we performed a pairwise post-hoc comparison. The comparison shows that NO-AR ($Mdn_{NO} = .196$, $SD_{NO} = .039$) had a significantly slower navigation speed than MEDIUM-AR ($Mdn_M = .263$, $SD_M = .06$, $z = -2.417$, $p < .001$) and LARGE-AR ($Mdn_L = .24$, $SD_L = .051$, $z = -1.625$, $p < .001$). Likewise, SMALL-AR ($Mdn_S = .212$, $SD_S = .038$) also was significantly slower than MEDIUM-AR ($z = -1.792$, $p < .001$) and LARGE-AR ($z = -1$, $p = .044$). To further complement these findings, we also measured the maximum navigation speed during each condition (see Figure 5.3 (D)). Although MEDIUM-AR performed best on average, we found no statistically significant differences ($\chi^2(3) = 3.052$, $p = .384$).

5.3.2 Workload

To better understand the effects of display sizes on workload-related measures such as ergonomics and cognitive load, we measured the amount of *head movement*, the participant's *pupil size*, the amount of *eye-gaze movement* during all conditions, the participant's *visual attention*, and the *subjective task load* (see Figure 5.4 and Figure 5.5 (A)). For eye-tracking measures, we omitted data from 7 participants due to insufficient tracking quality. For movement-based data, we omitted data from 3 participants due to technical issues with our prototype.

Head Movement

We measured the amount of head movement using the rotational data of the AR HWD (see Figure 5.4 (B)). Here, we found statistically significant differences in the amount of head rotation between conditions ($\chi^2(3) = 35.682$, $p < .001$). A pairwise post-hoc comparison reveal that LARGE-AR ($Mdn_L = 4.932$, $SD_L = 5.47$) is significantly higher than both NO-AR ($Mdn_{NO} = .176$, $SD_{NO} = 1.889$, $z = -2.353$, $p < .001$) and SMALL-AR ($Mdn_S = .278$, $SD_S = 2.388$, $z = -2.118$, $p < .001$), while MEDIUM-AR ($Mdn_M = 1.812$, $SD_M = 3.85$) is significantly higher than NO-AR ($z = -1.294$, $p = .021$).

Pupil Size

We also analyzed participants' pupil size as an objective indicator for mental demand [105]. We compared the relative pupil size between conditions which revealed statistically significant differences ($\chi^2(3) = 15.568$, $p = .001$). A pairwise post-hoc comparison shows participants had a significantly larger pupil size in NO-AR ($Mdn_{NO} = .072$, $SD_{NO} = .012$) than MEDIUM-AR ($Mdn_M = .069$, $SD_M = .009$, $z = 1.211$, $p = .023$) and LARGE-AR ($Mdn_L = .069$, $SD_{NO} = .01$, $z = 1.579$, $p = .001$). No statistically significant differences were found for SMALL-AR compared to all other conditions ($Mdn_S = .07$, $SD_S = .01$).

Eye-Gaze Movement

We found statistically significant differences in the amount of eye-gaze movement between conditions ($\chi^2(3) = 31.518$, $p < .001$, see Figure 5.4 (C)). A pairwise post-hoc analysis shows that SMALL-AR ($Mdn_S = .398$, $SD_S = .596$, $z = -1.353$, $p = .013$), MEDIUM-AR ($Mdn_M = .391$, $SD_M = .551$, $z = -1.471$, $p = .005$), and LARGE-AR ($Mdn_L = .686$, $SD_L = .553$, $z = -2.471$, $p < .001$) have a larger amount of eye-gaze movement when compared against NO-AR ($Mdn_{NO} = .041$, $SD_{NO} = .0595$). No significant differences between the AR conditions were found.

Visual Attention

We investigated the amount of time participants focused on the AR extension (see Figure 5.4 (D)) and visualize gaze behavior (see Figure 5.4 (E)). We thereby omit the NO-AR condition as participants were fully focused on the smartphone. For duration of visual focus, we found significant differences in how much time participants spent looking at the smartphone screen or the AR extension depending on the screen size ($\chi^2(2) = 20.235$, $p < .001$). A pairwise post-hoc analysis shows that participants looked significantly less at the AR extension during SMALL-AR ($Mdn_S = .33$, $SD_S = .185$) when compared to MEDIUM-AR ($Mdn_M = .453$, $SD_M = .205$, $z = -.941$, $p = .018$) and LARGE-AR ($Mdn_L = .697$, $SD_L = .196$, $z = -1.529$, $p < .001$). This gaze behavior also becomes apparent from the heatmap visualization (see Figure 5.4 (E)), where clusters of darker color indicate that especially for SMALL-AR participants looked mostly at lower central portion of the smartphone screen, with occasional glances at the virtual extension around the smartphone. Even in MEDIUM-AR and LARGE-AR, where participants overall spent more time looking at the virtual extension, the participants' gaze was mostly centered on or just around the smartphone. For LARGE-AR, participants' gaze appears to traveled across the extended display up to the MEDIUM-AR size (percent of total gaze time in MEDIUM-AR area, excluding inner sizes: $Mdn = 32.828\%$, $SD = 9.397\%$), but quickly falls off towards the edge (percent of total gaze time in LARGE-AR area, excluding inner sizes: $Mdn = 7.336\%$, $SD = 6.897\%$). Similarly, gazes in MEDIUM-AR also fall off towards the edges, but the overall virtual display area of MEDIUM-AR appears more evenly used.

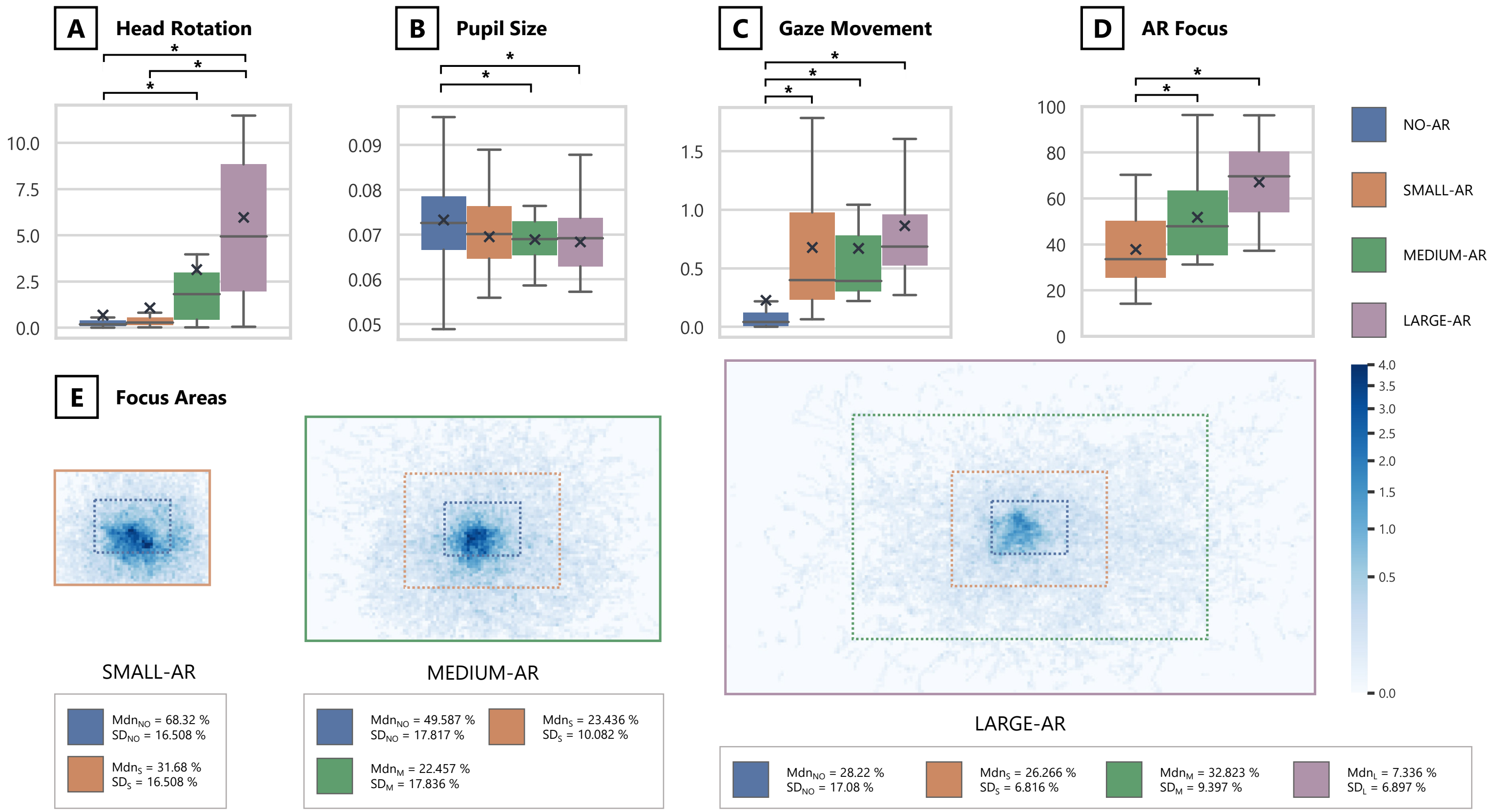


Figure 5.4: Summary *workload* measures. (A) Boxplots showing the average head rotation in °/s. (B) Boxplots showing the average pupil size during all conditions. Here, pupil sizes ranges are a relative value between 0 and 1 as calculated by the AR HWD. (C) Boxplots showing the average gaze movement in °/10ms. (D) Boxplots showing the average time in percent that participants spent looking at the AR screen extension. (E) Heatmaps showing how long participants focused on what part of the VESAD. Heatmaps indicate fixation duration per position related to total gaze time, whereby results were scaled with a power-law function ($\gamma = 0.5$) to increase the visibility of lower values. We marked corresponding virtual extensions sizes in each heatmap to better contextualize the values. We also added statistics below each heatmap indicating how much percent of a participant's focus each display size received (excluding inner conditions).

Subjective Task Load

We used the NASA TLX [164] to measure task load after each *navigation phase* (see Figure 5.5 (A) and Tables 5.5 and 5.6). We found statistical significant differences in all subscales: *mental demand*, *physical demand*, *temporal demand*, *performance*, *effort*, and *frustration* as well as the *overall scores*. In a pairwise post-hoc analysis, we found a statistically significant improvement of SMALL-AR compared to NO-AR in *effort*. A comparison of NO-AR and MEDIUM-AR shows statistically significant improvements for MEDIUM-AR across all subscales. When comparing NO-AR to LARGE-AR, we found significant improvements for LARGE-AR in *mental demand*, *temporal demand*, and *frustration*. We also found statistically significant improvements from SMALL-AR to MEDIUM-AR in *mental demand* and *frustration*. Repeating this pairwise comparison for the overall score, results show that MEDIUM-AR and LARGE-AR have a significantly lower workload than NO-AR, while MEDIUM-AR also has a significantly lower workload than SMALL-AR. On average, MEDIUM-AR performed better than LARGE-AR in all subscales, but no statistically significant differences were found.

Table 5.5: Pairwise comparison of the raw NASA TLX, as visualized in Figure 5.5 (A). Statistically significant entries are marked with a star*.

Scale	NO-AR ↔ SMALL-AR	NO-AR ↔ MEDIUM-AR
Mental Demand	$z = 0.333$, $p = 1$	$z = 1.521$, $p < .001^*$
Physical Demand	$z = 0.833$, $p = .152$	$z = 1.229$, $p = .006^*$
Temporal Demand	$z = 0.958$, $p = .061$	$z = 1.479$, $p < .001^*$
Performance	$z = 0.313$, $p = 1$	$z = 1.042$, $p = .031^*$
Effort	$z = 1.000$, $p = .044^*$	$z = 1.542$, $p < .001^*$
Frustration	$z = 0.688$, $p = .390$	$z = 2.000$, $p < .001^*$
Overall	$z = 0.750$, $p = .265$	$z = 2.229$, $p < .001^*$

Scale	SMALL-AR ↔ MEDIUM-AR	NO-AR ↔ LARGE-AR
Mental Demand	$z = 1.188$, $p = .009^*$	$z = 1.063$, $p = .026^*$
Physical Demand	$z = 0.396$, $p = 1$	$z = 0.354$, $p = 1$
Temporal Demand	$z = 0.521$, $p = .974$	$z = 1.646$, $p < .001^*$
Performance	$z = 0.729$, $p = .302$	$z = 0.813$, $p = .175$
Effort	$z = 0.542$, $p = .877$	$z = 1.125$, $p = .015^*$
Frustration	$z = 1.313$, $p = .003^*$	$z = 1.396$, $p = .001^*$
Overall	$z = 1.479$, $p < .001^*$	$z = 1.354$, $p = .002^*$

Table 5.6: Results of the Friedman's test for the raw NASA TLX, as visualized in Figure 5.5 (A). Statistically significant entries are marked with a star*.

Scale	Result
Mental Demand	$\chi^2(3) = 21.618$, $p < .001^*$
Physical Demand	$\chi^2(3) = 14.397$, $p = .002^*$
Temporal Demand	$\chi^2(3) = 25.07$, $p < .001^*$
Performance	$\chi^2(3) = 10.235$, $p = .017^*$
Effort	$\chi^2(3) = 20.127$, $p < .001^*$
Frustration	$\chi^2(3) = 35.226$, $p < .001^*$
Overall	$\chi^2(3) = 38.949$, $p < .001^*$

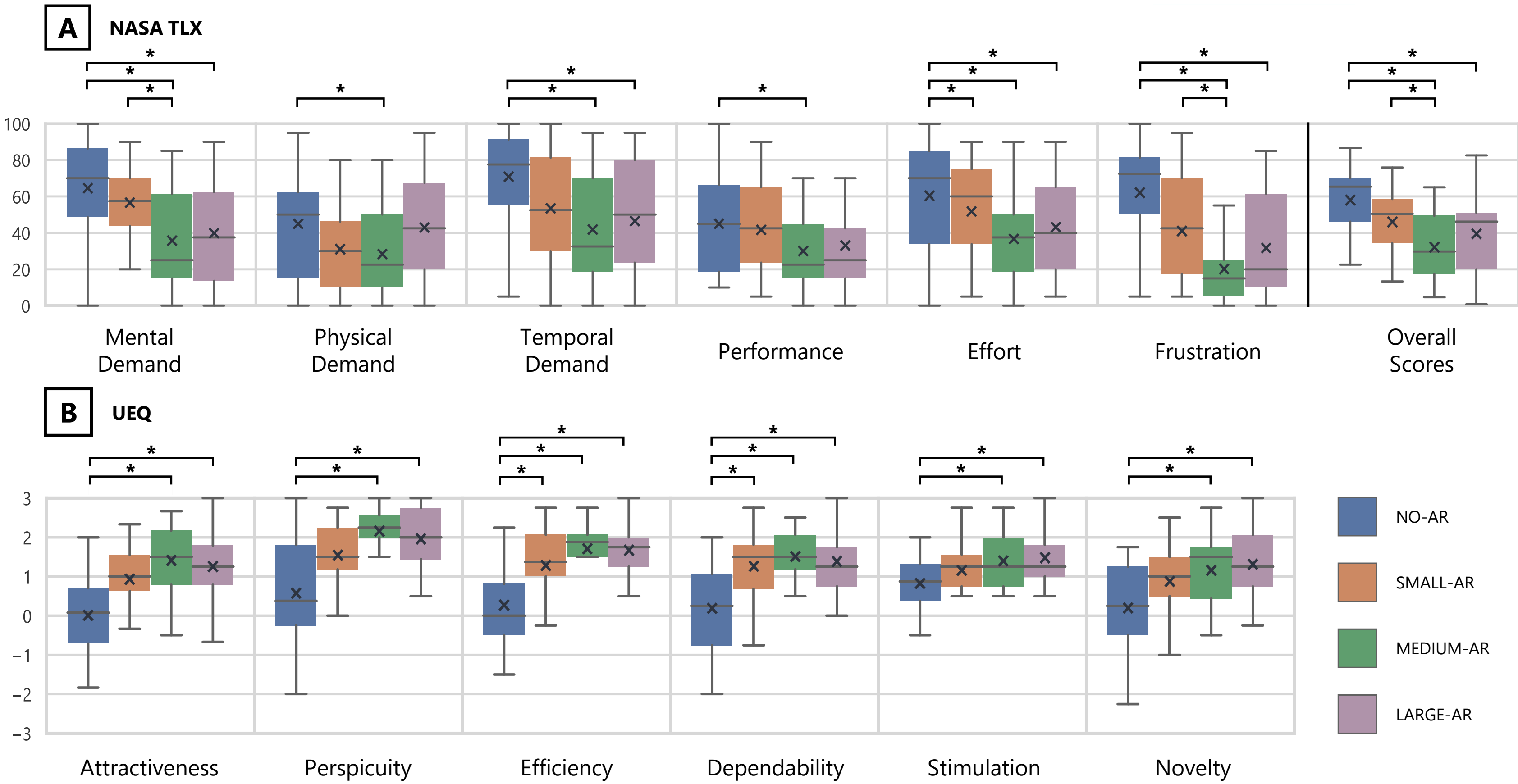


Figure 5.5: Measures for *workload* (A) and *user experience* (B). (A) Boxplots showing the results of the NASA TLX questionnaire for all conditions. (B) Boxplots of the results from the UEQ for all conditions.

Table 5.7: Results of the Friedman's test of the user experience questionnaire, as visualized in Figure 5.5 (B). Statistically significant entries are marked with a star*.

Scale	Friedman's Test
Attractiveness	$\chi^2(3) = 25.146$, $p < .001^*$
Perspicuity	$\chi^2(3) = 24.614$, $p < .001^*$
Efficiency	$\chi^2(3) = 36.500$, $p < .001^*$
Dependability	$\chi^2(3) = 30.434$, $p < .001^*$
Stimulation	$\chi^2(3) = 18.067$, $p < .001^*$
Novelty	$\chi^2(3) = 21.790$, $p < .001^*$

Table 5.8: Results of the pairwise comparison of the user experience questionnaire, as visualized in Figure 5.5 (B). Statistically significant entries are marked with a star*.

Scale	NO-AR ↔ SMALL-AR	NO-AR ↔ MEDIUM-AR	NO-AR ↔ LARGE-AR
Attractiveness	$z = -0.854$, $p = .131$	$z = -1.667$, $p < .001^*$	$z = -1.479$, $p < .001^*$
Perspicuity	$z = -0.896$, $p = .097$	$z = -1.646$, $p = .001^*$	$z = -1.375$, $p < .001^*$
Efficiency	$z = -1.250$, $p = .005^*$	$z = -1.896$, $p < .001^*$	$z = -1.854$, $p < .001^*$
Dependability	$z = -1.500$, $p < .001^*$	$z = -1.812$, $p < .001^*$	$z = -1.521$, $p < .001^*$
Stimulation	$z = -0.646$, $p = .499$	$z = -1.062$, $p = .026^*$	$z = -1.458$, $p = .001^*$
Novelty	$z = -0.667$, $p = .442$	$z = -1.146$, $p = .013^*$	$z = -1.604$, $p < .001^*$

5.3.3 User Experience

We measure user experience based on the *user experience questionnaire* and *subjective preferences* gained from a semi-structured interview. Figure 5.4 (B) shows an overview of our statistical findings.

User Experience Questionnaire

We employed a UEQ [352] after each *navigation phase* (see Figure 5.5 (B) and Tables 5.7 and 5.8). We discovered statistically significant differences in all scales: *attractiveness*, *perspicuity*, *efficiency*, *dependability*, *stimulation*, and *novelty*. A pairwise post-hoc analysis reveals that SMALL-AR was ranked better than NO-AR in *efficiency* and *dependability*. MEDIUM-AR was ranked better than NO-AR in all scales; similarly, LARGE-AR was ranked better than NO-AR in all scales. Although MEDIUM-AR ranked, on average, best in all scales except *stimulation* and *novelty*, no statistically significant differences compared to LARGE-AR could be found.

Subjective Preferences

We asked participants about their most and least favorite condition (multiple choices were allowed) and the reasoning behind this choice.

Regarding the most favored condition, participants were split between MEDIUM-AR ($n = 14/24$) and LARGE-AR ($n = 11/24$). Furthermore, 4 participants chose the SMALL-AR extension as their favorite. While some participants ($n = 8/24$) stated they liked the LARGE-AR condition for the extensive display size that provided a better overview, other participants found LARGE-AR too big ($n = 5/24$), too overloaded ($n = 1/24$), and disliked the head movement associated with the LARGE-AR condition ($n = 3/24$). For LARGE-AR, participants noted that "*I liked the television size [...] because when I scan the whole map to see where each symbol is, then the biggest size helps the most*" – P19 and that "*you could really take advantage of the headset by really looking around to see where the different items are*" – P13. In contrast, other participants argued that "*the TV monitor was almost too big, it was hard to keep everything in sight*" – P11. Participants ($n = 2/24$) also felt that they were most familiar with the MEDIUM-AR condition, as it resembled a typical desktop monitor in size. Lastly, one participant indicated some possible motion sickness issues due to increased amount of eye gaze movement in the LARGE-AR condition and therefore rated the SMALL-AR condition higher in comfort: "*[...] because I had to move my eyes less. [...] Once I have to move my eyes too much, I get motion sick*" – P2.

Regarding the least favored condition, participants almost unanimously ($n = 23/24$) chose NO-AR. One participant only disliked the SMALL-AR condition as it provided no real advantage to NO-AR. Another participant disliked both the NO-AR condition and the LARGE-AR condition because they felt that the LARGE condition made it harder to remember the icons, as a result of seeing the icons almost instantly. Participants expressed that they felt lost on the map ($n = 13/24$) due to seeing too little of the map ($n = 10/24$), especially when no other icon was visible. Participants also felt that the small display size forced them to search too much ($n = 5/24$) and see little to no relations with other objects ($n = 3/24$).

5.4 Discussion

In this section, we discuss the results of our laboratory experiment in the context of findings from prior work. We structure our discussion based on our research questions concerning *spatial memory*, *workload*, and *user experience*.

5.4.1 Spatial Memory

In terms of spatial memory, MEDIUM-AR and LARGE-AR significantly improved navigation path length and task completion time, thereby also improving overall navigation speed. While MEDIUM-AR also clearly resulted in a significantly higher learning effect and recall accuracy, we did not find any similar significant effects for LARGE-AR. Thus, our findings suggest that – similar to findings from Rädle et al. [321] and Zagermann et al. [423] – there is a "*sweet spot*" for display size in terms of spatial memory and that a larger available display size improves navigation performance. Prior findings by Gao et al. [137] and Uddin, Gutwin, and Cockburn [390] indicate that additional landmarks (e.g., seeing more icons in larger conditions) can facilitate the formation of spatial memory, which is in line with our qualitative findings. Despite this effect and unlike Rädle et al. [321], however, participants actually performed slightly worse again beyond the MEDIUM-AR size. Surprisingly, SMALL-AR also slightly decreased spatial memory despite being bigger than NO-AR – further deviating from prior findings [138, 321, 390].

This implies that AR extensions have an implicit cost associated with splitting the display into a real and virtual screen.

Contrary to our expectations, maximum navigation speed did not show a significant difference between the conditions. We expected the maximum navigation to increase with display size, as participants could be more confident in quickly flicking across the information space (as they are "*scrolling into the unknown*") and scanning the available area for the symbol. Instead, our results show a relatively consistent maximum navigation speed, indicating that the time required to visually scan the map and the time required to navigate to a new segment were roughly consistent across all conditions.

In summary, a larger display size generally contributes to a better task completion time. Spatial memory, however, can actually decrease for small virtually-extended screens. Here, MEDIUM-AR presents a "*sweet-spot*" for spatial memory, after which spatial memory starts to slightly degrade again.

5.4.2 Workload

In line with findings by Caluya et al. [70], our results show that head rotation increases with larger display sizes (i.e., insufficient virtual field of view for the given content is compensated by increased head movement). The LARGE-AR size comes at the cost of significantly more head rotation than both NO-AR and SMALL-AR. Although MEDIUM-AR fits comfortably within the participant's field, our data also shows a significant increase in head movement and rotation compared to SMALL-AR and NO-AR. Conversely, the cognitive load significantly decreased for the MEDIUM-AR and LARGE-AR conditions, which indicates a trade-off between ergonomics in terms of head rotation and cognitive load and is in line with prior findings by Rädle et al. [321].

Similarly, our data also shows that virtually-extended display size correlates with how much participants spent looking at the virtual screen. However, our analysis reveals that most of the participants' gaze is still focused on or around the smartphone. While the increased display space is well-used in the SMALL-AR condition, adding more display space beyond the MEDIUM-AR sees barely any use. Given the significant increase in head rotation for LARGE-AR, the extra space only caused additional physical load with little to no additional benefit.

To our surprise, gaze movement did not significantly increase with display size, but rather whether or not an virtually-extended screen was used. Since we ensured that both (real and virtual) screens were on the same focal plane (i.e., using a video see-through HWD, cf. [106, 150, 292]) and that there was no visible gap between the displays (cf. [288]), we expected gaze movement to correlate with display size. Here, further research is necessary to explore the actual underlying causes. For example, the smartphone bezel might provide a physical frame of reference, thereby further splitting the screen into two distinct displays, contributing to an increase in context switching (cf. [141, 325]). Alternatively, the smartphone may provide a "*sweet spot*" in terms of angular coverage [325], thus fitting well within the fovea-wide field of view. Another reason might be due to our participants' familiarity with a smartphone's physical affordances: By introducing a virtual screen, we added an unfamiliar affordance, thus leading to a higher cognitive load without much added benefit for small extensions: "*[During SMALL-AR], I was still focused on the smartphone. It took me a while until I looked at the [VESAD] again, it took me a while to convince myself that I can peek across the border*" – P18.

In summary, while larger virtually-extended displays are worse for ergonomics, the increased display space was well-used until MEDIUM-AR. Yet, there is no benefit in increasing the size beyond MEDIUM-AR (cf. [321]). In contrast, a small virtually-extended display causes disproportionately high cognitive workload.

5.4.3 User Experience

Our results show a clear subjective preference for any AR display extension over the NO-AR condition across attractiveness, pragmatic qualities, and hedonic qualities. Here, both MEDIUM-AR and LARGE-AR were consistently rated significantly higher than the NO-AR condition, which was confirmed in our semi-structured interviews: "*It's not too much to overwhelm you with information, but it's also not too small so that you have to search too much. […] It's like, if your monitor is too big you start to lose track of your cursor*" – P9.

Although MEDIUM-AR is ranked slightly better than LARGE-AR (cf. [423]), the difference might be due to a legacy bias [314]: Most of our participants are likely used to working on a notebook or desktop monitor, thus explaining their preference towards MEDIUM-AR over LARGE-AR. In contrast, other participants might use a television-sized monitor or multi-monitor environment in their everyday life, thus preferring LARGE-AR over MEDIUM-AR.

In summary, participants consistently favored larger AR extensions such as MEDIUM-AR and LARGE-AR. The subjective size preference between MEDIUM-AR and LARGE-AR might depend on the use case and participants' day-to-day experiences.

5.5 Limitations and Future Work

Due to the narrow focus on comparing virtually-extended display sizes against a smartphone baseline condition, this study has the following limitations.

We intentionally limited the overall map size to accommodate both NO-AR and LARGE-AR. While LARGE-AR may see further improvements with a larger information space (e.g., with regard to task completion time), this would negatively affect the NO-AR condition. Future studies could exclude the NO-AR condition to better study the effects of larger virtually-extended display. In addition, we assigned each condition dedicated icons and icon locations to achieve full counterbalancing of our conditions. Although we ensured that icons were equally placed between different maps and our results are mostly in line with prior work (e.g., [321, 423]), our results may be influenced by the layout of each map. We also intentionally switched to a desktop interface during the *object location recall phase*. While this reduced the ecological validity, it allowed us to better compare our results with prior studies (e.g., [321, 423]) and isolate spatial memory from other confounding influences (e.g., muscle memory).

Another limitation may be given by the use of a video see-through AR HWD. We intentionally decided against an optical see-through AR HWD to avoid confounding factors with respect to different focal planes (cf. [106, 150, 292]). However, video see-through HWDs are more cumbersome than optical see-through HWDs and have a reduced real-world field of view, which may negatively impact larger display sizes. In this regard, prior research already indicated that a restricted virtual field of view does not negatively impact spatial memory [70]. As AR HWDs continue to improve, further studies are necessary to investigate the impact of device ergonomics on larger virtually-extended display sizes. In line with this, future research could investigate whether a fully virtually simulated AR environment (e.g., combining a virtual reality HWD and a physical prop) could be used to increase the internal validity of the measurements – albeit at the cost of ecological validity.

Since this study compared different virtually-extended screen sizes with a fixed input modality as a static smartphone-sized peephole [271], there are many aspects left unexplored. For instance, future work could explore the effects of different physical screen sizes (e.g., smartwatches, tablets) and their relation to virtually-extended screens. Furthermore, future work could also investigate dynamic peephole navigation as input modality by tracking the smartphone in space (e.g., see [305]): Replacing touch interaction with physically moving the handheld device to explore the information space could further improve spatial memory (cf. [423]).

5.6 Insights and Implications

In this section, we synthesize our findings from our laboratory experiment (see Sections 5.2 and 5.3) and our discussion (see Section 5.4) to provide design insights (D 2.1–D 2.4) for the design of VESADs and research implications (I 2.1–I 2.4) to inspire future work. Key implications are summarized in call-out boxes at the end of this section.

While our results generally show that any virtually-extended display size can have significant benefits in terms of task completion time and user experience **(D 2.1)**, our results consistently favor MEDIUM-AR over LARGE-AR. Looking at our workload results, we conclude that MEDIUM-AR presents the best trade-off in terms of performance (e.g., task completion time, navigation path length) and increased task load (e.g., subjective and objective physical demand) **(D 2.2)**. However, a much finer granularity in the comparison of display sizes is necessary to find the "*tipping point*" that presents the ideal virtually-extended screen size **(I 2.1)**.

For LARGE-AR, the advantages gained from increasing the display sizes begin to diminish (cf. [321]), while the workload (e.g., ergonomics) starts to outweigh the benefits **(D 2.3)**. In this regard, we could fully realize the potential of AR to bend the information space around the user (i.e., similar to off-the-shelf ultra-wide monitors or CAVE [91] systems). Here, we could investigate if bending the virtual screen shows any benefits for virtually-extended display sizes (e.g., comparing CAVE-like system, ultra-wide monitor, and straight display) **(I 2.2)**.

For SMALL-AR, our objective results contradict the subjective user preferences: While users appreciate even the smallest virtually-extended display size (e.g., subjectively increased efficiency), our objective results indicate adverse effects in terms of spatial memory (e.g., navigation path length, recall accuracy). In terms of spatial memory and workload, we therefore conclude that a small virtually-extended screen is worse than providing no virtually-extended screen **(D 2.4)**. Although our results indicate clues (e.g., increased eye gaze movement) why SMALL-AR performed worse, further research is necessary to find the underlying cause **(I 2.3)**. However, prior works (e.g., [102, 229, 292, 333, 431]) already show several promising scenarios of utilizing the small space next to a display (e.g., for pushing UI elements out of the screen). As we did not study the interaction with the AR content *per se* (e.g., via mid-air input [292]), we suggest to systematically investigate use cases and scenarios for interaction with out-of-screen UI elements (comparable to the size of SMALL-AR) regarding the effect on a user's workload as the interaction might alleviate shortcomings related to spatial memory **(I 2.4)**.

Design Insights

- **D 2.1** Any virtually-extended screen size is beneficial in terms of user experience and task completion time for navigation tasks.
- **D 2.2** Virtually extending a smartphone to the size of a desktop monitor presents a "*sweet spot*" in terms of spatial memory, workload, and user experience.
- **D 2.3** Virtual extensions that are larger than a desktop monitor are detrimental to spatial memory and ergonomics with little additional benefit.
- **D 2.4** Small display extension negatively impact spatial memory and workload.

Research Implications

- **I 2.1** Find "*tipping point*" of virtually-extended display size that represents optimal trade-off between ergonomics and performance.
- **I 2.2** Compare different levels of bending the virtually-extended display around the user (cf. [137, 138, 246]).
- **I 2.3** Investigate effects of context switch between real and virtual displays (cf. [325]).
- **I 2.4** Investigate if interaction with out-of-screen UI elements can alleviate shortcomings of small display extensions (cf. [292]).

Supplemental Video

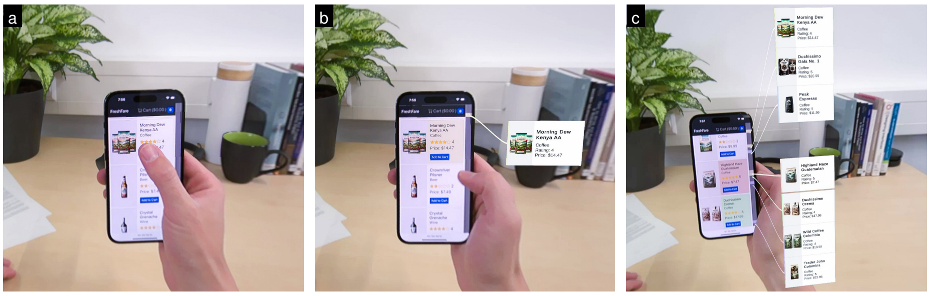


Figure 5.6: *Push2AR* is a novel interaction concept that enhances scroll list interaction on phones by pushing list items to the AR space. (a+b) The user bookmarks an item by swiping right on it to slide it into the AR space. (c) The position of bookmarked items in AR is synchronized with their respective position on the smartphone.

5.7 Follow-up Research Trajectories

The findings of this study demonstrate the benefits of virtually extending a smartphone screen with an AR HWD. While some factors require further investigation (see Section 5.6), the insights gained offer practical guidance for real-world applications. Building on this foundation, our recent work on *Push2AR* [407], applies these findings to a concrete use case targeting mobile AR glasses (see Figure 4.6).

Push2AR translates the experimental insights from *ARound the Smartphone* into a more ecologically valid setting, focusing on the ubiquitous task of browsing and comparing items in mobile list interfaces Traditionally, mobile list interactions, while functional, are constrained by the dimensions of the screen, often requiring excessive scrolling and switching between tabs to access and compare information. This traditional method can introduce visual clutter and increase users' cognitive load as they attempt to comprehend and interact with large amounts of data on a small display [162].

Push2AR enhances this experience by enabling seamless interaction between phone and AR space: A user is scrolling through a list of items while online shopping (Figure 5.6 a). Once they find an item to bookmark, they simply "*push*" the item to the AR space by sliding the finger from the item on the phone to the right (Figure 5.6 b). Positions of these AR items are synchronized with the scroll positions of respective items on screen. Using *Push2AR*, users extract and bookmark items of interest to make comparisons among large amounts of data (Figure 5.6 c). *Push2AR* further provides interactive *scroll bar indicators* that map the pushed items to their location on the phone. They serve as a bridge between the mobile and AR interaction spaces, allowing users to quickly jump to an item on the phone by tapping on its scroll bar indicator. By allowing users to interact with large data sets through such familiar gestures, the *Push2AR* concept demonstrates how to leverage the strengths of both AR and smartphone usability.

We implemented our interaction concept as an open-source web extension[3], showcasing how *Push2AR* can be applied to a wide range of real-life webpages. We further evaluated our approach in a user study with 16 participants and compared it against a smartphone-only web-browsing baseline to answer the research question on how *Push2AR* affects interaction efficiency, task load, and satisfaction. Although the experimental comparison and findings are not part of this thesis, users reported lower task load and higher levels of satisfaction when navigating through dense information using the AR space, but exhibited higher task completion times. Our findings show that *Push2AR* significantly reduces the number of page switches when browsing through lists online, leading to less frustration and, thus, improved user experience.

Both *ARound the Smartphone* and *Push2AR* illustrate the potential of hybrid user interfaces to improve elementary interactions, such as virtually extending a screen for navigation or offloading user interface elements, thereby improving user experience. Yet, these represent just two interaction techniques that make use of the distinct strengths of each device in VESAD configurations. Follow-up work has, for example, replicated our findings for virtually extending a smartwatch [196], explored the effects of distance between smartphone and AR content on text legibility [19], or applied such screen extensions to multiple view layouts for visual analysis tasks in cross reality environments [93]. By seamlessly expanding and integrating with existing display space, VESADs therefore offer promising directions for supporting fluid interaction.

5.8 Chapter Conclusion

The "*ARound the Smartphone*" exemplar examines the use of an AR HWD to virtually extend the display size of a smartphone. Here, we study the effect of different display sizes on spatial memory, workload, and user experience, thus informing how users transition between devices, contributing towards **RO1: Transitioning Between Devices**. Our findings highlight the practical benefits of using a VESAD configuration in a hybrid user interface, but also highlight potential challenges. Although we found that all virtually-extended screen sizes were beneficial to user experience and task completion time, the small display extension shows that there is significant overhead when introducing a virtual screen extension: If the virtual display extension is too small, the disadvantages of splitting the screen into a real and virtual screen outweigh the benefits of an increased screen size; if the extension is too large, device ergonomics start to supersede any benefit gained from extending the screen size. While this study does not directly address immersive analytics, it offers insights into effective device transitions in hybrid user interfaces and highlights design considerations for future applications, thus providing a foundational understanding of the objective benefits of hybrid user interfaces.

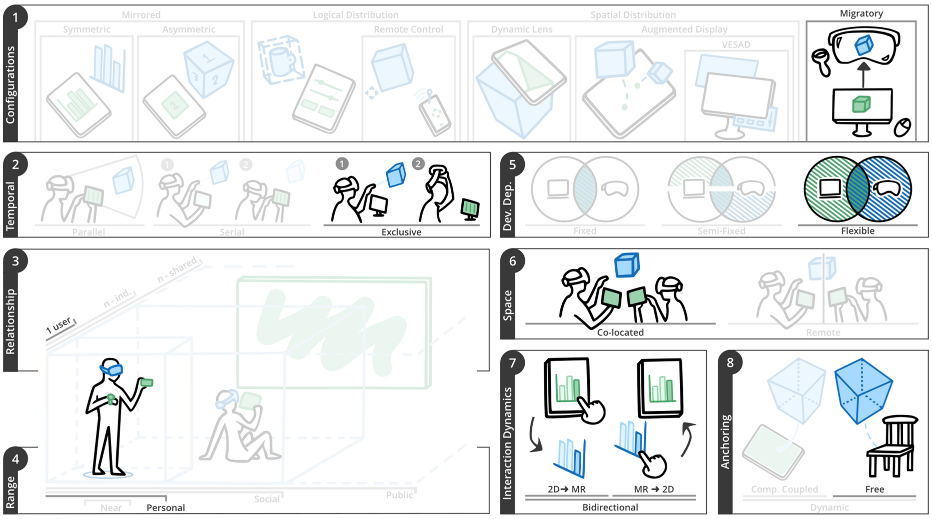
[3] https://github.com/hcigroupkonstanz/Push2AR

Chapter 6

Exemplar: ReLive

Chapter Content

The following exemplar can be classified as:



This chapter is based on the publications:

Sebastian Hubenschmid*, Jonathan Wieland*, Daniel Immanuel Fink*, Andrea Batch, Johannes Zagermann, Niklas Elmqvist, and Harald Reiterer. "ReLive: Bridging In-Situ and Ex-Situ Visual Analytics for Analyzing Mixed Reality User Studies." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '22. New Orleans LA USA: ACM, 2022, pp. 1–20. ISBN: 978-1-4503-9157-3. DOI: 10.1145/3491102.3517550

Sebastian Hubenschmid*, Johannes Zagermann*, Daniel Fink*, Jonathan Wieland*, Tiare Feuchtner*, and Harald Reiterer*. "Towards Asynchronous Hybrid User Interfaces for Cross-Reality Interaction." In: *ISS'21 Workshop Proceedings: "Transitional Interfaces in Mixed and Cross-Reality: A New Frontier?"* 2021. DOI: 10.18148/kops/352-2-84mm0sggczq02

Rasmus Lunding, **Sebastian Hubenschmid**, Tiare Feuchtner, and Kaj Grønbæk. "ARTHUR: Authoring Human–Robot Collaboration Processes with Augmented Reality Using Hybrid User Interfaces." In: *Virtual Reality* 29.2 (2025), p. 73. ISSN: 1434-9957. DOI: 10.1007/s10055-025-01149-6

* Contributed equally

Supplemental Video

Figure 6.1: The ReLive mixed-immersion tool. ReLive combines an immersive analytics virtual reality view (left) with a non-immersive visual analytics desktop view (right) for analyzing mixed reality studies. The virtual reality view allows users to analyze prior studies in-situ, while the desktop facilitates an ex-situ analysis of aggregated data.

6.1 Chapter Context

While hybrid user interfaces can enhance interaction in mobile scenarios (as demonstrated in Chapters 4 and 5), much of the analytical work is still performed on a desktop: Here, integration with established applications and access to powerful 2D data manipulations and statistical analyses remain essential. Yet, using multiple devices *synchronously* (i.e., *parallel* or *serial* usage) can be cumbersome. Users may want to spend prolonged time on the desktop (e.g., for statistical analyses) without wearing a HWD due to comfort. What about *asynchronous* use?

Consider Apple's Handoff[1] feature: users start writing an email on one device and can finish writing that email on another device. Handoff transfers the application state (i.e., the content of the mail) from one device to another. This works well for homogeneous devices supporting similar modalities such as physical or virtual keyboard input: the content of the mail is stored in the cloud and shown in different, yet homogeneous views. While we can employ a fluidly responsive design to support screens of different sizes for providing access to the same content, sharing this content across realities on the Virtuality Continuum [276] (e.g., from a smartphone to a VR environment) may be more complex. Here, we have to consider additional aspects, such as representing the same content in both 2D and 3D, or switching between entirely different input modalities. To ensure that the user can effectively switch, transitional interfaces must be designed such that mental connections between realities are maintained.

In pursuit of this idea, this chapter explores an *asynchronous hybrid user interface*, where heterogeneous (i.e., non-immersive and immersive) devices are used *exclusively*. While the exclusive use does not immediately offset the individual device-specific disadvantages with the advantages of another interface, it allows users to choose the appropriate device for a given sub-task. To ground this work on a real-world analysis scenario, this chapter explores the use of such asynchronous hybrid user interfaces in the context of analyzing MR studies.

[1] https://support.apple.com/en-us/HT209455, last accessed 2025-05-12.

6.1.1 Use Case

Given its rapid growth, MR environments are seeing an ever-increasing need for user studies and field evaluation, analyzing not only MR-specific metrics such as use of space [22], movement patterns [62, 215], or interaction ergonomics [117], but also more traditional HCI performance metrics such as time and error. For this purpose, researchers gather an abundance of quantitative and qualitative data such as interaction or movement logs (e.g., spatiotemporal data), audio/video recordings, and questionnaire responses in mixed-method evaluation approaches that allow for data triangulation. However, MR tools and techniques are particularly challenging to evaluate given their device heterogeneity, diversity of use, and mobile deployment. Additionally, many metrics are affected by the original environmental context [60, 215] or may be hard to analyze in a 2D context [62, 117]. Thus, Ens et al. [114] recently identified the establishment of a general evaluation framework as one of the current grand challenges in immersive analytics [76]. Several immersive analytics tools have already emerged that aim to streamline this process (e.g., [60, 215, 289]), offering increased immersion and flow [306]. This can be especially beneficial when analyzing spatiotemporal data (e.g., simulations of the actual study scene [60]), analyzing data within their environmental context [60, 215], or viewing 3D visualizations [60, 222, 289] **in situ**. However, "*being 'in the data' at times prevents an outside-in view that may be needed to get an overview of the data*" [60].

In contrast to these novel immersive approaches, non-immersive visual analytics tools are widely used to analyze data through a more traditional **ex-situ** approach – where the analyst is detached from the original study environment – that can outperform immersive counterparts for overview tasks [223]. The familiar input methods in these non-immersive tools allow for precise interaction (e.g., via mouse pointer) and facilitate the creation of specialized analysis workflows (e.g., via computational notebooks [13]). In addition, the 2D environment is suited towards a range of relevant study analysis tasks, such as analyzing video data, transcribing audio, pre-processing data, or exporting results for use within other applications (e.g., for statistical tests).

We argue that a holistic analysis of MR study data can therefore benefit from both immersive and non-immersive visual analytics tools, as they complement each other well [109, 114, 424]: On the one hand, immersive analytics tools excel in an in-situ analysis, allowing researchers to reconstruct the context of the original setting (akin to a crime scene investigation) and facilitating the analysis of inherently 3D data; on the other hand, non-immersive visual analytics tools excel in an ex-situ analysis, providing a holistic overview of the data and inter-compatibility with other tools, and allowing users to define their own specialized analysis pipelines to compare data across multiple participants or conditions. Yet, there is a missing link between both approaches: Researchers have to reconstruct their (mental) workspaces when switching from one tool to the next, thus barring any kind of serendipitous findings that might occur if this transition were seamless.

In addition, there has been little research on which tasks are best suited for immersive or non-immersive settings as well as on how to best transition between these environments [114]. More specifically, which aspects of the analysis process, if any, benefit from immersive representation, and which aspects are instead better served by the use of a non-immersive visual analytics interface? How can we transfer context information and support users when switching from, for example, an immersive VR device to a non-immersive desktop environment and back?

6.1.2 Research Questions

To investigate the aforementioned challenges, this chapter introduces ReLive: a *mixed-immersion* visual analytics framework [339] that combines both immersive and non-immersive views to enable the holistic exploration and malleable analysis of MR user studies (see Figure 6.1). ReLive offers an **in-situ VR view** suited for immersing the user in an interactive spatial recording replicating the original study setting (cf. [245]). Similar to prior work (e.g., [60, 215, 289]), users can walk through the scene, create visualizations based on entities and events within the scene, and view the study data within its original environmental context. Unlike prior work, which uses AR to visualize spatiotemporal data (e.g., [60, 289]), we argue for the use of VR, which does not require access to the original study setting (cf. [60, 289]), can simulate studies across the whole virtuality continuum [202, 276], and allows for higher immersion, which, in turn, may provide deeper insights into the study participant's environmental context. To cover both exploratory and analytical procedures of the analysis process and identify which aspects may be better served by an immersive analytics view or a 2D view, we complement our VR view with an **ex-situ desktop view**. Users can use this desktop view as a visual analysis workbook, taking advantage of a toolkit of visual analytics techniques for summarizing, linking, and exploring details of spatiotemporal, event, and nominal data to make comparisons between study sessions. The desktop view also allows for the playback of audio and video media, and offers a 2D window into the current VR view. Both the desktop and VR view are synchronized in real time, facilitating the switch between the different views – thus representing an *asynchronous* hybrid user interface [187] with *exclusive* temporal usage. Although our focus in this work is solely on a single-user system, such cross-platform environments can also open up the design space for asymmetric collaboration [77, 129, 344].

For ReLive, we were particularly interested in the interplay between in-situ and ex-situ analysis (contributing towards **RO1: Transitioning Between Devices**), and which tasks are best suited for which level of immersion (contributing towards **RO2: Task Allocation**). Although we also investigated the general applicability of this concept, these explorations are beyond the scope of this thesis and are thus omitted from this chapter.

RQ3.1 Interplay Between Ex-Situ and In-Situ Analysis
How do in-situ and ex-situ analysis complement each other?

RQ3.2 Task Allocation
Which analysis tasks benefit from immersive analytics, which tasks are better suited for non-immersive visual analytics?

We evaluated a prototype of ReLive in a two-step evaluation process to identify the benefits and challenges of combining immersive and non-immersive views for the analysis of MR user studies. In the first step, a guided design walkthrough was conducted with the authors of this paper to analytically evaluate and validate the concepts. In the second step, we invited 5 MR experts in an empirical user study to see how they use the interplay between immersive and non-immersive visual analytics for the analysis of study data. To ground this study in realistic data, we first explored relevant metrics for evaluating user studies conducted in MR environments and examined the level of immersion required for meaningful analysis.

6.1.3 Metrics for Evaluating Mixed Reality Studies

When evaluating MR studies, researchers employ a variety of different metrics, the choice of which is heavily influenced by the study type and research objectives [114, 289]. Performance metrics such as task completion time, accuracy, or error rate are often used for MR studies and offer a well-understood point of reference [32]. However, established metrics for 2D interaction may not be easily transferred to these immersive environments: For example, applying ergonomics metrics created for interacting with vertical displays [168] to mid-air interaction may result in misleading results, especially when visualized in situ [117]. In addition, most MR studies "*need to cover many more factors than studies of non-immersive surroundings*" [32], such as place illusion and world awareness [363], environmental constraints [60, 117], or novelty bias [114]. Although an extensive analysis of MR study metrics exceeds the scope of this work, recent works [32, 114, 274] point towards a lack of standardization for MR study metrics.

An interview with domain experts conducted by Nebeling et al. confirms that "*[researchers] mentioned many types of data specific to their projects but relatively few concrete metrics*" [289]. Rather, many MR systems are evaluated qualitatively based on interview data, observations, or bespoke visualizations of available study data, such as movement data [62, 282]. To aid in these observations, some studies employed a passive observation client [179, 394, 415], which can provide more insights into the digital environment than the user's point of view.

Another aspect of evaluating MR studies is capturing and calculating specific metrics: Nebeling et al. [289] automated the calculation of a set of global metrics (e.g., task completion time, distance moved, area coverage of user movement) in a Unity framework; Kloiber et al. [215] integrated a clustering algorithm to automatically detect keyframes in a recording with high spatial activity; and Lilija, Pohl, and Hornbæk [245] enable the user to step through notable changes of a selected object. Prior work also supports creating annotations and tags [60], recording of gestures and voice commands [289], or defining tasks for calculating metrics [289]. Yet, some MR systems may not rely on evaluating user movement, but instead investigate olfactory (e.g., [23]) or taste (e.g., [1, 401]) feedback.

In summary, well-established performance metrics work especially well for simple interactions such as pointing and dragging, but are usually insufficient to completely characterize more complex activities in MR. The possibilities of multi-modal interaction, multi-user scenarios, and multi-device environments therefore demand an analysis environment that not only allows for the calculation of classical measures, but can also offer richer ways to capture, visualize, and analyze these complex activities (e.g., using a 3D environment [117]).

6.1.4 Degree of Immersion for Analysis of Mixed Reality Studies

When analyzing user studies, researchers usually turn towards (non-immersive) desktop analysis tools, offering a wide range of different well-established software suitable for analysis, the choice of which depends on the data set, the goal of the analysis, and the expertise of the user group [13]. Existing approaches offer a fully-featured graphical user interface which may require data preprocessing, or provide a powerful development environment (e.g., R [130] or Python) that allows users to both calculate their metrics and output visualizations. These non-immersive approaches are highly configurable, as simple visualizations may be expressed using a common specification (e.g., Vega-Lite [348]) to replicate the visualization in a wide range of analysis tools, or use powerful toolkits (e.g., D3.js [47]) to generate bespoke visualizations. To give more context to these visualizations, computational notebooks (e.g., [13, 143, 311, 323]) are often used, employing the concept of literate computing [277] to narrate the analysis process by combining explanations, code, and resulting visualizations as a visual analysis workbook.

Specific to the analysis of MR user studies, past efforts have also investigated non-immersive visualizations of spatiotemporal data. For visualizing user movement and orientation, such prior work has augmented top-down views [57, 80, 98, 383, 395] and 3D scene views [58, 98, 316] with different visualizations, including trajectory plots [57, 58, 80, 98, 383], heatmaps [57, 80, 395], or field of view frustums [57, 80, 316]. Often (e.g., [57, 58, 80, 98, 268, 395]), visualizations can be controlled by a timeline in combination with playback controls. These visualizations of movement data are also often complemented by the playback of one or more video recordings (e.g., [57, 58, 98, 268, 395]), synchronizing the visualization of user movement with the actual video recordings. Furthermore, several works [57, 58, 268, 395] visualize calculated or manually annotated events (time points and time periods) as part of a timeline.

In contrast to these (non-immersive) desktop-focused approaches, recent work has also introduced various immersive prototypes and toolkits that facilitate the analysis of spatiotemporal data from AR [60, 289] and VR studies [215, 245]. The capability of these prototypes depends on the individual research focus, which affects the choice of visualization: Here, 3D trajectory plots and 3D trails are often used to visualize the position and speed of objects of tracked devices [62, 245], participants' heads and hands [215, 245], or gaze cues [316]. In addition, Nebeling et al. [289] use 3D point plots to visualize events such as the position and direction of users, tracked objects, and physical markers, while a tablet shows an overview of the events in a 2D visualization. Büschel, Lehmann, and Dachselt [60] also enrich spatial 3D trajectories with additional videos and 2D visualizations (e.g., heatmaps, scatterplots) that can be placed in the AR environment, while Lilija, Pohl, and Hornbæk [245] use these 3D trajectories as an interactive, non-linear time slider.

In summary, both non-immersive and immersive analysis tools represent viable choices for analyzing data from MR user studies: Non-immersive tools offer flexibility and reproducibility (e.g., via computational notebooks) and are well-integrated in a rich ecosystem of established applications, for example allowing users to export their results to their research paper. In contrast, immersive tools can reveal the environmental context [60], increase spatial understanding [222], help in understanding physical measurements [233], and aid in the decision-making [328], but still suffer from novelty factors and discomfort issues [73, 310] such as HWD weight [416], temperature [416], or simulator sickness [209] which can make these immersive approaches unattractive for some users over longer periods of time. We therefore argue that a holistic analysis of MR user studies requires both immersive and non-immersive approaches – allowing users to choose and transition between different levels of immersion based on their current analysis task. Yet, most existing tools only support one or the other; users have to therefore either choose one or spend significant effort in migrating their current analysis workflow to a different reality.

6.2 ReLive

We propose ReLive: a visual analytics framework [339] that combines immersive with non-immersive views to enable the holistic exploration and malleable analysis of MR user studies. Based on trends in immersive and visual analytics study analysis, ReLive provides an immersive VR view for in-situ analysis with a synchronized, non-immersive desktop view for ex-situ analysis (see Figure 6.1): The VR view allows users to relive an interactive spatial recording replicating the original study (cf. [60, 245, 289, 316]), while the complementary desktop view facilitates the malleable analysis of aggregated study data (cf. [13, 45, 311, 323]).

To ground ReLive in authentic data and realistic evaluation scenarios, we refer to five reference studies (RS) that served as testbeds for the conceptual design, prototypical development, and evaluations of ReLive:

RS1 In this study [22], domain experts were tasked with individually exploring a domain-specific data set, configure their visualizations in a 3D space, and present their findings. Here, the use of space or the resulting complexity of visualizations was studied, focusing on quantitative measurements in a within-subjects design.

RS2 In this study [408], two co-located participants were asked to collaboratively position and rotate virtual furniture to pre-defined target positions using handheld AR tablets. This allowed to study the influence of different interaction techniques on task performance (e.g., accuracy or task completion time), focusing on quantitative measurements in a within-subjects design.

RS3 In this study [286], two participants (either co-located or remote) were instructed to remember the position of task objects using different display configurations: The study compared a handheld AR condition (tablets superimposing digital content on their real-world camera) with handheld VR condition (tablets showing a completely digital world), both controllable through egocentric navigation. Both display conditions featured visual landmarks in the form of digital furniture. The display configuration represented the within-subjects factor and the spatial dispersion (e.g., co-located or remote) represented the between-groups factor. The authors studied the influence of the display configuration on measures such as the perceived social presence (i.e., for remote collaboration) and task completion times, while combining other qualitative and quantitative measurements in a mixed design (between-groups and within-subjects factors).

RS4 In this study [179], participants were tasked to individually explore and analyze a visualization of 3D parallel coordinates using the various input modalities provided by the handheld tablet (e.g., touch or tablet orientation) and AR HWD (e.g., head-gaze or voice input) to investigate the applicability of this multimodal interaction approach, focusing on qualitative measurements in a within-subjects design.

RS5 In this study [125], two remote participants were asked to jointly decide on a travel destination and plan a trip. In this realistic negotiation task, the influence of the location and representation of the remote peer allowed to study user experience and subjective perception of presence, focusing on qualitative measurements in a within-subjects design.

The selection of these reference user studies was guided by the aim to cover a wide range of MR user study situations to ensure the general applicability of ReLive. It ranges from single-user studies using an AR/VR HWD (RS1, RS4), over co-located collaborative user studies using handheld AR/VR (RS2, RS3), to remote collaborative user studies using AR HWDs (RS5), each focusing on various types of measurements and applying different study designs.

To facilitate replication in ReLive, we created a data specification that unifies the recorded data from these different studies. Existing data was converted to this specification using bespoke preprocessing pipelines and imported to ReLive. To support and validate our concepts in real world scenarios, we also created a data logging toolkit to easily recreate and analyze future studies in ReLive.

The following sections describe the *data specification*, the *component templates and instances* that allow for a malleable analysis, the *non-immersive desktop view* and the *immersive VR view*, and the *transition between in-situ and ex-situ analysis*. Lastly, we describe our *data logging toolkit* to support ReLive for future user studies.

Throughout this chapter, we will refer to the following example scenario to showcase the possible benefits of using ReLive, which is further illustrated in the figures in this chapter. The scenario is based on findings that were unveiled when analyzing data from RS3 with ReLive during development.

Scenario *The HCI researcher Sarah finished the mixed reality mixed-method experiment RS3 with 32 participants and two conditions. Before investigating the gathered data with statistical tools such as IBM SPSS, she starts exploring the qualitative and quantitative data with ReLive to get a first impression and to see how the display configurations were adopted by the participants. For this, she opens up a new analysis notebook in the desktop view and selects RS3, which then loads all data based on the pre-defined dependent variables on the desktop and on the complementary VR view.*

6.2.1 Data Specification

Our data specification was designed to holistically reflect the data from a user study, allowing ReLive to reconstruct the study as accurately as possible. The data format differentiates between four types of logging messages: *sessions*, *entities*, *events*, and *states* (see Figure 6.2). Additional data, such as audio or video, may be included as attachment, but is stored as an ordinary file to facilitate external access.

Sessions. A session represents a self-contained subset of the study data pertaining to one study session containing multiple entities and events. Depending on the study design, a session can represent a trial, an experimental condition, or an entire study session.

Entities. Each entity describes an actor, input device, or object with a visual appearance, which should be replicated. Entities may change over time (e.g., their

Figure 6.2: Hierarchy and relation of data specification. Each logging type shows a data example from RS2. Relational and technical metadata (e.g., timestamps, foreign keys) were omitted for brevity.

position or rotation, see *states*) and are usually represented by a 3D model of the original object. Entities may also contain media (e.g., screen capture of a tablet) that can be used for a more holistic replication (e.g., displaying a tablet's screen recording directly on its virtual replica).

Events. Events represent ephemeral actions that typically do not have a physical representation (e.g., task start, touch event), but occur at a specific point or interval in time. Similar to entities, events may occur at a specific position within a room (cf. [289]) and may contain media data, such as screenshots. Events may optionally refer to specific entities: For example, a *touch* event may refer to a tablet *entity* to later identify the origin of the event.

States. States capture the changed properties of any entity or event over time, thus recording each individual change.

Scenario *In RS3, each pair of participants is divided into two sessions, one for each condition: handheld AR and handheld VR. In this study, entities are represented by the participant's tablets, the virtual furniture acting as landmarks, and the cubes of the memory task. New events are created at the position of a participant's tablet whenever a cube is unveiled by the participant.*

6.2.2 Component Templates and Instances

To establish a malleable analysis workflow that works in both VR and on a desktop, we adapt the concept of *components* from computational notebooks (cf. [13, 45, 323]). Instead of having to program each component separately, our concept differentiates between *component templates* that are programmable and *component instances* which execute the template's code with data provided by the analyst.

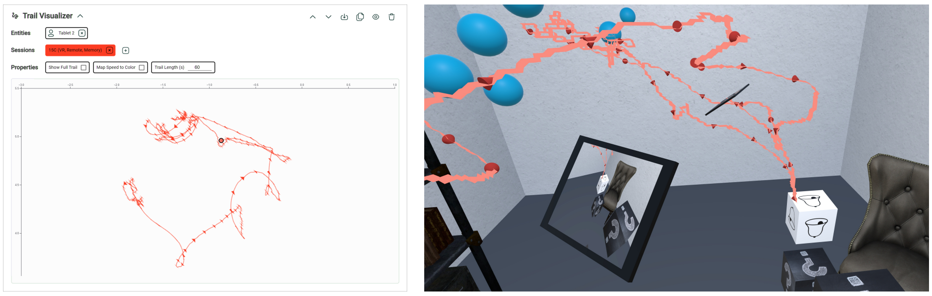


Figure 6.3: Component instance. Used on the desktop (top) and in the 3D study replication (bottom), created from a *trail visualizer* component template. On the desktop (top), the component requires one entity, as well as at least one data set (session) to create a top-down 2D visualization of the tablet's movement. The component exposes three properties to further customize the current visualization. Additional buttons in the top right corner allow for general control (e.g., delete component, export data). In VR (bottom), the same component instance also adds a 3D visualization in the 3D study replication, which can reveal additional information such as user behavior.

ReLive thus combines the benefits of a desktop environment for programming, while still offering a flexible analysis workflow in VR.

Component templates can be programmed by analysts to calculate and visualize metrics, or to add custom behavior to the interactive 3D study replication – thus allowing for the creation of custom analysis tools. By using the data specification, templates are generic and only need to specify how many entities or events are required for successful code execution. In addition, analysts can specify optional parameters to further customize the calculation. Component templates facilitate the distribution of metrics across different analysis workflows, as these templates can be easily packaged (cf. [45]) and shared with others. For example, consider a component template that visualizes the user's movement as a trail (cf. [60, 245], see Figure 6.3): Here, the analysts can specify that the component requires one entity, add code to convert the entity's movement data to a visualization, and provide additional parameters (e.g., length of visible trail in seconds) that can customize the resulting visualization. In future work, we aim to make these templates more easily programmable by the user, for example by offering an editor similar to existing computational notebook, allowing components to output a Vega-Lite specification [348] depending on their specified input parameters.

In contrast, *component instances* (see Figure 6.3) can be created from existing templates. Instances can be customized by the analyst by adding study data or adjusting parameters. Once configured, components automatically execute the code defined in their corresponding template. Component instances either add custom behavior to the 3D study replication, or visualize the calculated metrics. Visualizations can differ between the non-immersive and immersive environment akin to multiple coordinated views, making the best use of each environment.

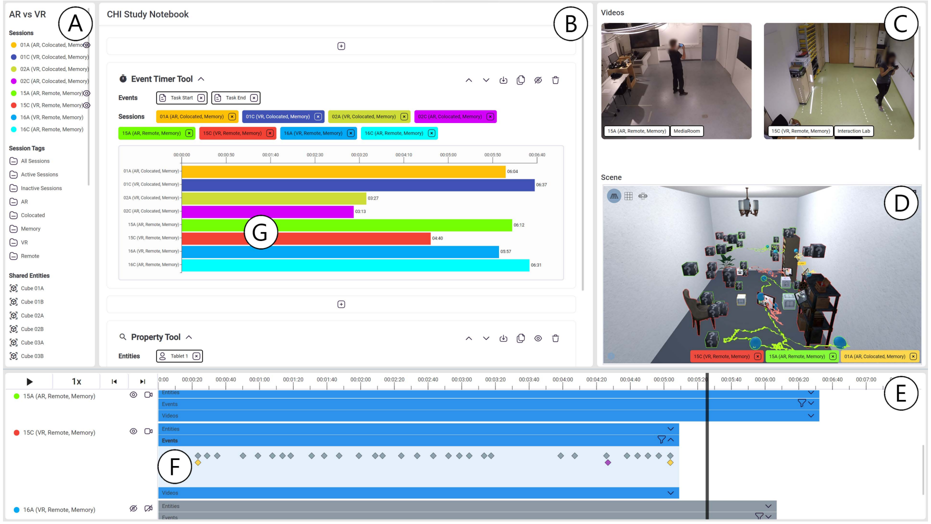


Figure 6.4: ReLive desktop view. This view combines a computational notebook approach with a video editor interface, specializing in an ex-situ analysis of aggregated study data. Access to additional media data such as videos and 3D study replication allows for a holistic insight into the user study.

For example, when instantiating the trail component template, analysts have to specify a entity for this instance and can adjust the defined properties, such as trail length. Once one or more sessions (i.e., data subsets) are added to the instance, the template's code is executed, and a visualization appears in the component (see Figure 6.3 (top)). In addition, a corresponding 3D visualization appears in the 3D study replication if the affected entity is visible (see Figure 6.3 (bottom)).

6.2.3 Non-Immersive Desktop View

The non-immersive desktop interface of ReLive is designed to provide analysts with a holistic ex-situ overview of the study's available data and is suited for an analytical evaluation of aggregated study data using a components-based approach (see Figure 6.4). Our concept takes inspiration from computational notebooks (e.g., [13, 45, 311, 323]), which use an interactive programming environment to display results such as visualizations inline (cf. literate computing [277]). Here, analysts can create component templates and instances, allowing analysts to, for example, narrate their analysis process, calculate and visualize metrics (see Figure 6.4 (B)), or export data from these components for external applications.

This notebook approach is combined with a video playback interface akin to state-of-the-art video editors to better support the complementary in-situ analysis. Analysts can display available 2D media data (see Figure 6.4 (C)), or see into a MR study session via a reconstruction of the study session (see Figure 6.4 (D)). The current playback status is synchronized across all available media data and can be controlled with a timeline.

The desktop interface is divided into five panels to merge the computational notebook approach with a video editor (see Figure 6.4): a *data panel* (A), a *component panel* (B), a *timeline panel* (E), a *video panel* (C), and a *3D scene panel* (D).

Data Panel. The data panel shows an overview of all available sessions, entities, and events within a study (see Figure 6.4 (A)), allowing analysts to assign this data to other panels via drag and drop (e.g., to configure component instances). To facilitate comparisons across study sessions, data with similar attributes (e.g., identical name) is automatically grouped together: *Session tags* act as a smart folder of sessions with similar attributes (e.g., the same experimental condition), allowing analysts to drag and drop multiple sessions at once; *shared entities* and *shared events* group together similar entities or events, respectively, across sessions and thus allow for an easy comparison of data across different sessions.

Component Panel. A component panel allows analysts to visualize study-specific metrics as visualizations via customizable components (see Figure 6.4 (B)). Here, analysts can create new component instances from existing component templates. Created component instances are organized linearly like a computational notebook, allowing analysts to narrate their analysis process. In addition, the visualizations can provide contextual information about the current playback time (e.g., as interactive playhead in a time series line chart, see Figure 6.4 (G)), linking the ex-situ analysis with the available videos and 3D study replication.

Timeline Panel. Similar to a video editor, a timeline panel at the bottom controls the current playback time, speed, and status (see Figure 6.4 (E)). To facilitate the comparison across study sessions, the current playback time is synchronized across all sessions – each session is therefore akin to a single video track in a video editor. In addition, each session timeline can be expanded to reveal an overview over the lifetime of all entities and a timeline visualization of events (see Figure 6.4 (F)), which can be filtered (cf. [289]). Static data (e.g., screenshots contained within events) can be inspected by hovering over the data, while videos can be displayed in the *video panel*.

Video Panel. The video panel provides additional insights into the real-world context of the study (see Figure 6.4 (C)). Here, videos enabled in the timeline panel are displayed as separate videos and synchronized with the playback time displayed in the timeline panel, akin to a video editor.

3D Scene Panel. The 3D scene panel provides additional insights into the virtual context of the study (see Figure 6.4 (D)). Based on the available data, the original study is reconstructed as an interactive 3D scene suited for in-situ analysis (cf. [316]). This 3D scene panel provides controls similar to common 3D editing programs such as free camera movement and isometric (e.g., top-down) camera perspectives, but can also provide additional insights, for example by replaying

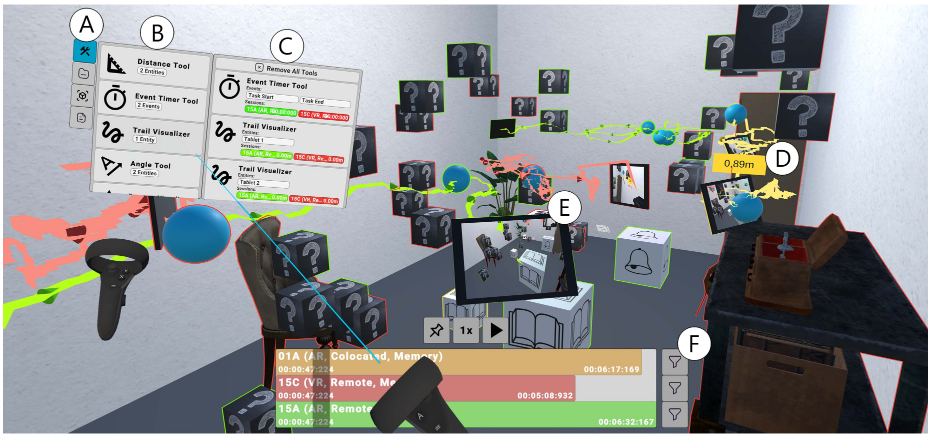


Figure 6.5: **ReLive VR view.** This view immerses analysts in an interactive reconstruction of the original study. A user interface is affixed to the analyst's left controller and a timeline at the bottom allows analysts to control the simulation. In addition, metrics and utilities from component instances are visualized in situ. Once multiple sessions are visible simultaneously, a colored outline helps in differentiating entities and events from different sessions.

the scene from a user's point of view. In addition, analysts can drag and drop data to instantly locate the corresponding 3D object. Lastly, visualizations from corresponding component instances are displayed in the 3D scene window, allowing for a quick glance into the VR view.

Scenario *As Sarah is interested in task completion times, she creates a new component instance of an already available "Event Timer" component template in the component panel. To get an initial overview, Sarah adds the data from all participants to this component instance by dragging and dropping the "All Sessions" tag from the data panel to the newly created component instance. In addition, Sarah specifies a start and end point by dragging in the representative events from one of the sessions in the timeline panel into the component instance. Once the data has been defined, a 2D barchart automatically appears – revealing that one participant pair was considerably slower in their first trial (see Figure 6.4 (G)). To investigate this outlier, Sarah looks at the video data by dragging in the outlier sessions into the the video panel to show the real world video, and then dragging the same session into the 3D scene panel to show a digital reconstruction, controlling their playback with the timeline panel. However, viewing the replay on a desktop makes it difficult to understand how the participants moved through the digital environment, as Sarah has to constantly adjust the camera to understand the objects' spatial relations.*

6.2.4 Immersive Virtual Reality View

In contrast to the desktop view which focuses on ex-situ visual analytics, the VR view focuses on in-situ immersive analytics (see Figure 6.5). The VR view is inspired by prior work (e.g., [60, 245, 289, 316]), which enriches the in-situ analysis with environmental context. To that end, analysts can relive an interactive replication of the original study, which is enabled by our data specification: Session backgrounds (e.g., 3D model of the room) can be added to the data set and displayed in VR; entities are reproduced using, for example, their 3D model and move around the scene based on their captured movement data; and events are visualized in situ as colored spheres (cf. [289]).

For better cohesion with the desktop view, the user interface is structured similarly to its desktop counterpart (cf. [71]), yet also geared towards an in-situ analysis. Analysts can interact by clicking and pointing on user interface elements or objects within the scene, mirroring the features of the desktop where possible. In the future, we aim to provide an ex-situ view in VR similar to the desktop view, allowing for a more detailed control over the analysis. The interface of the VR view provides analysts with access to the *study data*, the created *components*, a *timeline*, and displays available *media* data within the scene.

Study Data. Analysts can reveal additional information about entities and events as a floating tooltip by pointing at a corresponding 3D object within the scene. Analysts can also use a user interface anchored to their left controller to get a basic overview of all entities and events within a scene (see Figure 6.5 (A)). This menu also allows analysts to focus on specific entities (e.g., by hiding all irrelevant entities). Lastly, analysts can browse through all available sessions and load the data into the scene. When multiple sessions are active at the same time, a colored outline is added to each object in the scene, allowing analysts to map entities to their corresponding data set.

Components. To allow for an analysis of the data, analysts can create new component instances through a 2D menu anchored to the left controller (see Figure 6.5 (B)). However, as the input modalities for VR are unsuited for text-based programming, creation of new component templates is restricted to the desktop view. Thus, component instances can be created by selecting a template, then clicking on the corresponding entities or events within the 3D scene.

The user interface shows existing component instances (see Figure 6.5 (C)), which can also be placed anywhere in the 3D space. The resulting visualizations may appear in situ (movement trail, see Figure 6.5 (D)), as a freely placeable visualization, or as a single value depicting the current value (see Figure 6.5 (D)).

Timeline. Similar to the desktop view, a timeline allows analysts to control the playback state, speed, and time (see Figure 6.5 (F)). In addition, analysts can scrub back and forth by using the VR controller's joysticks. The timeline can also provide an overview of all events, which can be filtered based on event properties.

Media. To simulate a device's screen during the actual study, each device shows their current view on the virtual scene on their screen (see Figure 6.5 (E)). We aim to visualize more recorded media data in future work, for example by replaying screen recordings from a tablet on the corresponding replica's screen, or adding another video playback panel to the *components* menu.

Scenario *To better investigate the outlier, Sarah immerses herself in VR by putting on a VR HWD – putting her into the same environment as visible in the desktop's 3D scene panel. For comparison, she loads in another participant pair by opening the study data interface on their left controller and activating another session. Thus, another set of tablets appears, which are now highlighted with a yellow outline to distinguish them from the tablets of the other session (outlined in red and green). To better visualize the participant's movement, Sarah selects a "Trail Visualizer" component instance from a predefined component template and attaches this to each tablet by pointing and clicking on each tablet with her right controller. As a result, each tablet now shows a trail of its movement over the last few seconds (see Figure 6.5). Using the timeline interface, Sarah quickly scrubs through the session, allowing them to relive the study. By observing the tablet's position and screen, Sarah notices that one participant's tablet was occluded by the digital furniture placed in the room, which could be a potential cause for the longer task completion time. She also notices that the participants communicated by holding their tablets into the cubes they wanted to select (see Figure 6.3 (bottom)). For both cases, Sarah selects the "Camera" from the components panel to take a photo, thus saving these incidents within ReLive for further analysis.*

6.2.5 Transitioning Between In-Situ and Ex-Situ Analysis

Both the non-immersive desktop view and the immersive VR view were designed to suit different analysis workflows. Depending on the task and context, analysts may prefer to work in situ (e.g., in the VR view) or ex situ (e.g., on a desktop). ReLive therefore aims to facilitate the transition between in-situ and ex-situ analysis and follows the guidelines proposed by Carvalho, Trevisan, and Raposo [71].

Most importantly, application state (e.g., components, playback time) is synchronized across both views in real time, thus "*[making] users aware of the system state*" [71]: Component instances created in the desktop view are instantly visible in the VR view and vice versa. This is further supported as analysts can explore the 3D study replication on the desktop view (see Figure 6.4 (D)). In addition, the desktop view allows analysts to drag data (e.g., entities, events) from the ex-situ overview to the 3D scene replication, causing the camera to focus on the related object in the 3D scene. Analysts can therefore instantly explore the environmental context from within the ex-situ overview.

In future work, we also want to implement and examine *cross reality linking and brushing*: Here, users could mark outliers in the component instance of the desktop view, which is then instantly highlighted in the VR view and vice versa – akin to linking and brushing [208] *across realities*.

Scenario *While investigating the outlier in VR, Sarah also notices that the "cube selection" events of one pair of participants are clustered within one corner of the room – while the events of other pairs are spread throughout the entire room. To* *quantify this, Sarah attaches a "Property" component instance to the tablets of each session and configures the component instance in the component menu to show the "total distance moved" – thus displaying the current distance above each tablet. By scrubbing to the end of the session, Sarah can now compare their values between the currently loaded sessions, revealing a substantial difference between the different sessions. To better generalize this across all 32 participants, Sarah switches back to the desktop view. Here, Sarah can seamlessly resume her workflow with the previously created "Property" component instance, which shows a line chart of the "distance moved" metric. By dragging and dropping the "All Sessions" tag to this component instance, Sarah can easily generalize her findings for all participants and continue their analysis, for example by correlating the movement with task completion time or exporting her findings and visualizations to her research paper.*

6.2.6 Data Logging Toolkit

To utilize ReLive in a real world scenario, we developed a holistic logging and malleable evaluation environment, supporting researchers throughout all stages of a study (see Figure 6.6). At the center of this logging framework is a data specification that aims to captures the abundance of qualitative and quantitative data collected during a study (e.g., spatiotemporal data, interaction events, video/audio recordings, 3D models), allowing researchers to *relive* the original study as accurately as possible. With the increased availability of 3D scanning hardware (e.g., LIDAR sensors in Apple iPad), scans of the real world environment can be added to the resulting data set [316]. In comparison to prior work that offers similar workflows (e.g., [60, 289]), our concept follows two main design principles:

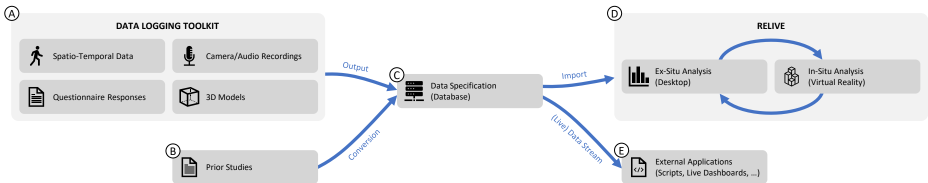


Figure 6.6: Logging and evaluation overview. This environment supports researchers throughout the all stages of a study: (A) An open data logging toolkit can facilitate data capturing, including spatiotemporal data (e.g., movement data from AR HWD), video data (e.g., imported or automatically captured from network cameras), questionnaire responses, and 3D models. (B) Data from prior studies can be converted to the data specification for use within ReLive. (C) Our data specification contains holistic data about a study, and can be easily accessed by external applications. (D) ReLive uses the data specification to enable a holistic study data analysis and reconstruct the original study setting. (E) Data from an ongoing study can be streamed to (and parsed by) external applications, enabling an analysis using bespoke scripts or live dashboards.

(1) Openness: Our concept was designed to run on all platforms (e.g., Unity, web) and integrates well with external applications by relying on established standards (e.g., JSON for interaction data, standardized file formats where possible). Our aim is to encourage open science by making it easy to share and reuse data sets.

(2) Extensibility: To offer support beyond the status quo and accommodate a wide range of MR studies, the toolkit is designed to be easily extensible without requiring changes to the underlying code.

6.3 Evaluation Process

We evaluated ReLive in a two-step evaluation process: First, a guided design walkthrough (Section 6.4) allowed us to analytically investigate ReLive in a formative evaluation; second, an expert user study (Section 6.5) provides deeper insights into the real world applicability of ReLive in an empirical evaluation. The overall goal of this evaluation process was to better understand the interplay between an in-situ and ex-situ analysis as exemplified by the VR and desktop view of ReLive, respectively, and how this combination can support the analysis of MR user study data. In line with the research goals of *task allocation*, *interplay between ex-situ and in-situ analysis*, and *applicability*, we created an *evaluation prototype* that implements the core features of ReLive.

6.3.1 Evaluation Prototype

To address our research questions, we created an evaluation prototype of ReLive (see Figures 6.1–6.5), which was iteratively improved. First, we focused on implementing core concepts for the guided design walkthrough. This first version already supported both a limited desktop view and a limited VR view, which were synchronized: The desktop view offered predefined component templates (see Table 6.1) that produced 2D visualizations, a timeline panel, and a 3D scene view; the VR view also supported the same predefined component templates that produced in-situ visualizations and a timeline. Based on the feedback from the guided design walkthrough, we refined our evaluation prototype to support an authentic evaluation of our reference studies. This study prototype includes all features described in Section 6.2, with the exception of user-programmable component templates, which were intentionally replaced with predefined *tools*. This reduced the complexity of this study and allowed us to focus on uncovering initial challenges relating to *RQ3.1: interplay* and *RQ3.2: task allocation*, especially as potential participants would not be familiar with the used reference studies.

We decided on seven predefined tools listed in Table 6.1, based on the evaluation requirements of our reference studies RS1–RS5. The tools are aimed to address a wide range of analysis tasks and allow for an easily understandable, yet authentic analysis of our five reference studies. Consequently, we intentionally decided against implementing other ideas for very specific components that we

Table 6.1: Overview of predefined tools. We defined seven component templates as predefined tools for our final prototype. Tools were derived based on the evaluation requirements of our reference studies. The chosen tools showcase either visualizations available in both desktop and VR view, or utilities for analyzing the study in VR. The *event timer tool* was not part of the prototype for the guided walkthrough.

Tool	Visualizes	Desktop view	VR view	Use Cases / References
Distance Tool	Distance between two entities	Line chart (distance vs. time)	Virtual measuring tape connecting the entities with label showing the current distance	• Position error during docking task, as in RS2 • Finding F-Formations
Angle Tool	Angle between two entities	Line chart (angle vs. time)	Angle visualization connecting the entities, projected to the floor	• Rotation error during docking task, as in RS2 • Finding F-Formations
Trail Visualizer	Entity movement over time	Top-down view on 2D trail (xz graph)	3D trail visualization in the VR scene	• Analyzing movement behavior / spatial activity, as in RS1, RS4, and RS5 • Detecting tracking issues
Event Timer Tool	Time between two events	Bar chart	Line connecting the events with label showing the time between them	• Task completion time as in RS1, RS2, RS3 • Action duration as in RS4
Property Tool	An entity's properties and derived metrics	Line chart (property value vs. time)	Label with the properties' values hovering over the entities	• Displaying raw data and metrics during analysis, such as distance moved per minute, as in RS1
Frustum Visualizer	An entity's field of view within the 3D scene	Configuration only	Lines visualizing the frustum directly at the entity in the VR scene	• Visualize what participant has seen, as in RS5 • Attention grouping
Camera	Screenshots of the 3D study replication	Gallery of taken screenshots and possibility to take screenshots from scene view	Virtual camera for taking screenshots in the VR scene	• Documenting insights (e.g., outliers)

discussed during our design walkthrough. To align with our research questions *RQ3.1: Interplay* and *RQ3.2: Task Allocation*, each tool either provides a visualization on both the desktop and in VR, or utility within the 3D scene (Frustum Visualizer, Camera).

6.3.2 Technical Implementation

For the ReLive evaluation prototype, we use a client/server structure with two separate clients (immersive VR and non-immersive desktop). Our server uses a multithreaded Node.js v12 runtime and is backed by MongoDB. The server is responsible for serving media files as well as preprocessing (e.g., data compression) and calculating metrics to ensure that the clients remain responsive. For this, we employ different transport protocols, such as HTTP for fetching static data, TCP/WebSockets for real-time data synchronization, and WebRTC for video transmission. For the desktop interface, we utilize web technologies for rapid prototyping (e.g., Angular), while the immersive VR client was written in Unity. The 3D scene in the desktop interface is streamed as a native HTML video from a rendertexture within Unity. The ReLive evaluation prototype, data specification, logging toolkit, and sample data is available as open-source project on GitHub[2].

[2] https://github.com/hcigroupkonstanz/ReLive

6.4 Design Walkthrough

We analytically evaluated the initial prototype of ReLive to validate our concepts, develop ideas for additional features that are essential for the analysis of MR user studies, and verify the general applicability of ReLive. We conducted design walkthroughs [165] using each of our five reference user studies (RS1–RS5) as testbeds. Each session lasted about 1.5 hours and 4–7 authors participated in the design walkthroughs. We prepared both request and evaluation sheets for task-related features for systematic feedback during the sessions, and used screen sharing in a video conferencing tool to share the same point of view, while following all ethical and sanitary guidelines provided by our universities.

Roles. During each session, attendees were given specific roles: A *moderator* who moderated and guided the design walkthrough and kept track of time; a *note-taker* responsible for filling out evaluation sheets for task-related features; a *presenter* who was involved in the reference user study (i.e., one of the authors) and could provide insights into the prior study analysis; and an *analyst* who was not involved in the reference user study and tried to replicate the analysis within the ReLive prototype – mimicking a potential user. Other participants without a dedicated role participated by discussing ideas and noting down possible issues by observing the analyst. We switched roles for each reference user study to, for example, avoid that a person who was involved in the prior study analysis takes on the role of the analyst as this might have influenced the workflow.

Procedure. First, the presenter briefly introduced their reference user study by presenting the study's goal, research questions, and other user study related aspects (e.g., procedure, data gathering methods, apparatus). This introduction was concluded with a list of research objectives and metrics that were investigated during prior analysis. This list was then briefly discussed and extended with additional analysis ideas of all participants. Next, the analyst started the ReLive application and progressed through each of the research objectives. All participants discussed possible solutions to each research objective and evaluated the use of existing features, which were noted down by the notetaker.

Results. We collected a total of 59 requests and 16 evaluation sheets for task-related features from all walkthroughs to prepare the prototype for an expert user evaluation. We first estimated the priority and implementation effort for each feature request and tagged it using the following tags: (1) Web, (2) VR, (3) Bridge, (4) Component, and (5) Visualization. We then ranked each feature request based on its generalizability, combined possible duplicates, and filtered out requests that did not fit within the scope of the following expert user study. Remaining features were further organized based on common topics, using a semantic clustering approach. Lastly, we discussed and sketched out possible ideas, before implementing 18 requests that were in line with our research questions. This extended prototype was then used for our expert user study. The outcome of our walkthrough is merged with the insights from our expert user study and are discussed in Section 6.6.

Aside from revealing opportunities to improve usability, our requests were mainly concerned with *new visualizations and tools for study analysis* and *interaction techniques to facilitate the transition between environments*. In terms of *new visualizations and tools for study analysis*, many ideas were linked to visualizing the available data types: For example, transcribed audio data could be presented as a wordcloud in 2D for overview, while showing the exact position where the words were spoken in 3D. Similarly, a waveform visualization could be combined with a 3D trajectory trail in the 3D scene replication, highlighting areas where participants talked.

Concerning *interaction techniques to facilitate the transition between environments*, we gathered requests to increase the cohesion between the in-situ and ex-situ analysis: For example, users can drag and drop entities directly into the desktop's 3D scene reconstruction to zoom in on the relevant object; show the point of view of the VR user or different entities within the 3D scene (e.g., showing the exact point of view of an AR tablet); or change the position of the VR view directly in the desktop view. To further increase cohesion, ReLive could support linking and brushing on both the desktop and VR (i.e., *across realities*): Here, areas of interest can be marked in an ex-situ 2D visualization, which would automatically highlight relevant areas within the corresponding in-situ visualization. However, the implementation and evaluation of such a *cross reality linking and brushing* exceeds the scope of this work.

6.5 Expert User Study

To empirically evaluate ReLive, we conducted an expert user study, focusing on participants that had prior experience in conducting and analyzing MR studies. Our goal was to evaluate ReLive guided by our two research questions. Our tasks were based on real data from two of our five reference user studies. We collected qualitative and quantitative data to gain insights into participants' workflows.

6.5.1 Participants

We recruited 5 male MR experts between 25 and 38 years ($M = 29.80$, $SD = 5.17$) from different research labs as participants. We intentionally looked for researchers with prior experience in analyzing or conducting MR studies and who were not affiliated with any of the current works of the authors. Thus, we invited researchers from a data analysis, an immersive analytics, and a virtual reality lab from the University of Konstanz. We also invited a colleague from the human-computer interaction lab at the University of Konstanz who was not involved with the design or implementation of ReLive. All participants had degrees in computer science, 4 with master's degrees and 1 with a doctorate. All of them were working in academia (e.g., as lecturers or research assistants) and had conducted MR user studies and analyzed their results before. All participants had normal or corrected to normal eyesight and did not suffer from color blindness; consequently, they had no problems with text sizes and the color-coding used in the different visualizations. We asked them to rank their experience in conducting AR/VR studies ($M = 4.20$, $SD = 0.45$) and analyzing the results ($M = 3.60$, $SD = 0.55$) on a scale from 1 (very inexperienced) to 5 (very experienced). On the same scale, they also ranked their experience with computational notebooks ($M = 2.40$, $SD = 0.55$) and virtual reality applications ($M = 4.60, SD = 0.55$). Four of them already developed a VR application on their own. We also asked participants if they prefer to analyze MR study data on their own (n=3) or together with others (n=2). All participants interacted with ReLive for the first time during the expert user study – this means, none of them were involved in the design nor the implementation.

6.5.2 Apparatus

All studies took place in one of our labs in which we allotted a walkable area of approximately 2×2 m where participants could freely move. At one side of this area, we set up two tables (each 1.40×0.80 m) with the long sides aligned. At one table, the participant was seated. The experimenter sat down at the opposite table. We equipped the participant's table with a 27″ 4k display, a mouse, and a keyboard connected to a desktop PC (simulating a desktop workspace, similar to Figure 6.1) that they used to work with the non-immersive desktop view of ReLive and to fill out questionnaires. Additionally, we provided participants with a tethered VR HWD (Oculus Quest 2) and the accompanying controllers to work with the immersive VR view of ReLive. The connection cables of the HWD were mounted via ceiling trusses to provide participants freedom in their movement while avoiding tripping hazards. On the experimenter's table, we placed two displays (both 27″ 4k), a mouse, and a keyboard connected to the participant's desktop PC. One display was mirrored with the participant's display and the other display showed the VR scene.

6.5.3 Procedure

Participants were welcomed and provided with introductory documents explaining the purpose and procedure of this study. They signed a consent form and filled out a demographic questionnaire. Using a slide show, the experimenter then introduced participants to ReLive and the data set of RS3. After that, participants started with the guided phase (see tasks below) to familiarize themselves with ReLive. Then, participants received a short introduction to the data set of RS4 before starting with the *free phase* (not guided by the experimenter). At the end of the free phase, participants filled out the System Usability Scale [53] and each session was concluded with a semi-structured interview, which included a subjective rating of the interplay of both views. Sessions took approximately 1.5 h and participants received compensation for their time. We followed all ethical and sanitary guidelines provided by the local institution at the time of the study.

6.5.4 Tasks

Participants started the tasks sitting in front of the desktop view. They were allowed to stand up and move through the allotted walkable area at any time. For the guided phase, we carefully ensured that all tasks were balanced between the desktop view and VR view. This means that the guided phase was not favoring one of the views. For the free phase, participants were given the opportunity to use both views as they suited them for their analysis workflow.

Guided phase. Participants were guided by the experimenter to solve the following six tasks using step-by-step instructions. For each task, we summarize in italics which instructions the participants received. The tasks resembled authentic analysis scenarios based on the actual analysis of RS3. Participants started with the desktop view and over the course of these tasks, they were required to switch between the desktop and VR view 5 times and therefore had 3 phases in the desktop view alternating with 3 phases in VR.

1. **Task completion time** (one session). Visualize the task completion time of a given session by measuring the time between the start and end event. *Participants solved this task using the desktop view. They created an event timer tool and added the given session, its start event, and its stop event to compute and visualize the session's task completion time.*

2. **Task completion time** (all sessions of a condition). Visualize the task completion time of all sessions of a given condition. *Participants solved this task using the desktop view. They added the given condition's session tag to the event timer tool they created in the previous task.*

3. **Accuracy** (one session). Measure how accurately the participants of a given session placed a given entity in comparison to a given target position. *Participants started with the desktop view to select the session and find the entities' position within the scene using the 3D scene panel. They were then instructed to switch to the VR view and create a distance tool by visually connecting the two entities. This also involved winding through time and teleporting through the 3D scene.*

4. **Accuracy** (all sessions of a condition). Measure how accurately all participants of a given condition placed a given entity in comparison to a given target position. *Participants switched back to the desktop view for this task. Here, they added the session tag for the given condition to the distance tool they created in the VR view during the previous task.*

5. **Events.** Investigate where the participants of a given session were located in the room when a given event happened. *Participants switched back to the VR view, filter the events, and use the VR controller's joystick to wind forward and backward to investigate the events' position and occurrence.*

6. **Tracking Issues.** Investigate if a given tablet (entity) in a given session had tracking issues by visualizing the speed (unrealistic high speed is an indicator for tracking issues). *Participants switched back to the desktop view and used the property tool to investigate the time series line chart for the speed of the given tablet. In the chart, they identified segments where the tablet moved with an unrealistic high speed. They then switched to the VR view and created a trail visualizer to investigate the direction of the jumps in context.*

Participants needed between 15 min 8 s and 25 min 36 s (M = 20.25 min, SD = 3.53 min) for the guided phase – excluding the times for switching between desktop and VR, as they were not representative due to hygiene requirements. In total, they spent between 6 min 38 s and 13 min 44 s (M = 9.25 min, SD = 2.52 min) in VR. While E1 and E2 always stood up when in VR, E4 remained seated during all tasks. E3 stood only briefly during the last phase in VR. E5 started the first use of VR seated but then decided to stand up for the remaining time in VR. In total, participants were standing between 0 min and 10 min 6 s (M = 6.01 min, SD = 4.16 min).

Free phase. Participants explored and analyzed the data set of RS4 on the their own and were free to use either environment on their own accord. As a starting point, the experimenter suggested the following four analysis goals. However, participants were free to follow their own analysis approaches.

1. Investigate the distance between HoloLens and the Interaction Tablet over time. Did the distance increase while the user was holding the tablet?
2. Which participant moved the most?
3. Were there any tracking issues?
4. When holding the Interaction Tablet vertically, at which height did the participants roughly hold their Interaction Tablet: At eye-level, or at shoulder level?

The free phase had a soft limit of 10 minutes. However, all participants decided to continue their individual analysis at the end of these 10 minutes and spent between 11 min 20 s and 17 min 14 s (M = 13.37 min, SD = 2.19 min) in the free phase – again excluding the times for switching between the views. All participants decided to start their analysis using the desktop view. E3 switched 3 times between both views, while the others switched only once. Participants spent between 2 min and 10 min 21 s (M = 5.70 min, SD = 2.89 min) in VR. Only E3 remained seated during the whole free phase. All other participants stood up for using the VR view. In total, they were standing between 0 min and 10 min 50 s (M = 4.93 min, SD = 3.86 min).

6.5.5 Data Collection and Analysis

We used two ceiling-mounted cameras with opposing views on the scene to capture video data. We recorded the content of the participant's display and HWD. For audio recordings (e.g., interviews), we placed a dedicated microphone in the center of the tables. Further, we transcribed all interviews and analyzed them following an inductive thematic analysis approach [51]: After familiarization with the data, one author identified data extracts that are relevant to our research questions and generated descriptive codes to label them. These descriptive codes were then counterchecked and validated by two other authors. We then thematically clustered the codes to identify potential themes. In the further process, we iteratively revised and refined the themes to ensure that they were in line with the dataset and our research questions.

6.5.6 Findings

Here we present the findings of the expert user study. We use the five themes that resulted from the thematic analysis of the data collected in our semi-structured interviews as structure to report our findings – this means that we present the scores of the subjective rating and system usability scale thematically aligned.

Theme #1: Desktop view for overview, and VR view for reasoning in the environmental context

All experts generally expressed that the desktop view is better suited for getting an overview of the study data compared to the VR view. For example, they said that "*the advantage of the [desktop] view is overview*" – E1 or that "*the 2D charts provide a quick overview*" – E3. In addition, E1 noted that the "*VR view lacks [this] overview over the data*" – E1 and that the "*VR tools are not powerful enough to quickly get the big picture*" – E1. Two experts (E1, E4) emphasized that this enabled them to find points of interest (e.g., outliers) in the data that they then could analyze more in detail using the VR view. Besides, E3 explained that the desktop view is better suited for performing statistical analyses and to "*correlate different factors*" – E3. Also, experts (E1, E3, E4) mentioned that the desktop view was faster to use when configuring components.

Concerning the VR view, experts noted that it allows you to immerse yourself in the data (E3) and to relive the study in its spatial context (E3, E5). Our experts mentioned various situations and tasks for which this can be beneficial. For example, E3 noted that reliving the study can be especially helpful when analyzing the study data weeks after you conducted it: "*[Then the VR view] offers the possibility to dive back into the study. It's like being inside the study again […] and maybe seeing aspects that were no longer present*" – E3. Two experts (E2, E4) additionally mentioned that they found the VR view helpful to get a better overview of the course of the sessions. According to our experts, this spatial-visual impression (E1, E2) helps to assess where entities and events were located in the scene (E2) and to see what the participants of the study have seen (E1, E2). Although E5 stated that he did not use the VR view during the free phase as the desktop view already provided an interactive representation of the VR scene, he described the potential utility of the VR view for tasks that benefit from an increased depth perception. In this context, experts (E1, E2, E4, E5) also positively highlighted the stereoscopic 3D view in VR as an advantage over the scene view in the desktop view. They perceived it beneficial for spatial measurements (E2, E4, E5) and assessing the entities' and events' depth and spatial constellations (E1, E5). "*The [desktop scene view] helps a lot. [...] I can explore it like a game. However, only as 2D representation of a 3D world with which I don't get the spatial impression. I only get that when I put on the headset. For example, when I look at [these objects] in the 2D scene view it is hard to assess their depth. I get that much, much better in VR*" – E1.

With that, the VR view helps to explain the data (E2, E3, E4, E5), for example to find reasons for outliers and also allows for exploring the data to discover points of interest that you would not see in the 2D diagrams presented in the desktop view (E1, E3). Participants' descriptions of their workflows during the free phase reflected these different strengths of the two views: Three experts (E3, E4, E5) stated that they started in the desktop view to get an overview of the data and then – if needed – switched to VR for reasoning (E4, E5) or to explore the data further (E3): "*For me, the [desktop] view is overview, and if I need details, I'll switch to VR, and that's a coherent workflow for me*" – E5. In contrast, E2 stated that he would start in VR to better understand the scene and switch to the desktop afterwards. E1 could imagine to use both workflows and E1, E3, and E5 agreed that the workflow "*depends on the research questions and the data*" – E3 to be analyzed. We also discovered three different patterns of participant behavior when using VR: (1) Remaining seated, (2) standing up, and (3) switching between standing and sitting. Interestingly, participants behavior partly changed between the guided and the free phase, indicating that their behavior not only relates to personal preferences but also to the type of task (cf. [28]).

All in all, experts stated that both views complement each other well (E1, E3), that their combination has no disadvantages (E1, E2, E4), and that the VR view is necessary to solve the tasks (E4).

Theme #2: Interplay between in-situ and ex-situ visual analytics for analysis workflow

A common theme during the evaluation was the interplay between the different views (i.e., desktop for ex-situ, VR for in-situ) and, thus, analysis types (i.e., in situ or ex situ). Three experts (E2, E4, E5) appreciated the instantaneous synchronization between the desktop and VR environment, with one expert noting that: "*If I start in 2D and then see something [...] interesting within the data, I'll jump there, put the headset on and look at it*" – E1. Another expert (E4) considered the synchronization the most important feature for the interplay. Three experts

(E1, E2, E4) noted that, while the interplay was "*not perfect*" – E1, it worked "*flawlessly*" – E1 and there were "*no problems*" – E2, E4, removing the need to configure the VR view (E1). One expert (E1) highlighted the consistency of both views in terms of their icons and interaction with the components.

When transitioning between devices, problems occurred for 2 experts (E3, E4) due to loss of orientation: "*You somehow have a cut, and then you have to reorient yourself*" – E3. Similarly, one expert (E4) also lost touch with the desktop's position in the real world. Here, E3 suggested showing the real desktop device in VR, which could act as an anchor point and could enable the use of the keyboard in VR. In addition, one expert (E4) noted that the transition could be more fluid, while two experts noted that the transition "*worked well*" – E1, E2 and was "*quite fast*" – E2. Here, two experts (E2, E4) mentioned that the 3D scene panel in the desktop view helped them when switching between devices, noting that it was "*really helpful*" – E2. Another expert (E1) appreciated the combination of a "*familiar*" – E1 desktop view with its link to a native 3D application.

The availability of a desktop view in VR for quick access to the ex-situ analysis view was appreciated by four experts (E1, E2, E3, E4): "*I thought the possibility to jump directly to the desktop view was very cool. I can imagine that it helps in ensuring a quick transition*" – E3. However, three experts (E1, E2, E3) requested a deeper integration of this desktop view into the VR environment. Similarly, four experts (E1, E3, E4, E5) missed the availability of their component's 2D charts in VR, for example for finding the exact point in time when an outlier occurred: "*I would've liked to see [the 2D charts] in VR! Because then I get the impression in 2D: okay, here the curve rising, then I can jump to this point and view it in detail*" – E1. During the concluding interview, we also asked experts to rank the interplay between both views on a scale from 1 ("*not very useful*") to 10 ("*very useful*"). Here, experts ranked the prototype with an average score or 8.00 ($SD = 1.73$).

In terms of using both in-situ and ex-situ view for their analysis workflow, opinions were divided: Two experts (E1, E5) expressed a preference for working in the desktop view, especially as the desktop still offers a higher resolution (E1) and as HWDs can be uncomfortable (E1, E5). E5 therefore preferred working solely on the desktop view, treating both views separately and commenting that "*the [desktop view] can do everything, why should I still switch to VR?*" – E5. In contrast, E1 could also imagine staying in VR if a well-integrated desktop view was available. Nevertheless, all experts used both views during the free phase of our expert study and, on average, spent almost half their time in VR. In addition, two experts (E1, E2) were positive about integrating the switch between desktop and VR into their analysis workflow, noting that each view has its distinct advantages and disadvantages: "*You can completely use the [desktop view] [...], and then for parts where there is an advantage for VR, you can use VR, and vice versa*" – E2. E3 highlighted the potential for future hardware, which could offer better ergonomics and enable more fluid transitions.

Theme #3: Modularity to suit individual analysis requirements

Throughout the interview, experts mentioned several features that are required for analyzing their specific study data, "*[because] everyone has different requirements for what is analyzed*" – E1. Here, experts highlighted the variety of different analysis use cases from their own work, such as crime scene investigation (E2, E5), collective behavior (E2), ergonomics (E3), balance issues (E1), or general asynchronous remote tasks (E4). Two experts (E1, E4) therefore emphasized the system's extensibility, with one expert (E1) appreciating that ReLive uses a similar approach to a computational notebook. Experts also mentioned integrating additional visualizations, such as aggregated trails (E5), 3D heatmaps (E1, E5), or eye-tracking data (E4). To support this, experts noted that the data specification should support eyetracking data (E4), questionnaire data (E4), interaction data from non-immersive systems (E5), and audio recordings (E5).

Furthermore, experts suggested different metrics that could provide further insights, such as metrics to calculate a user's balance (E1), speed (E2), room coverage (E2), task completion time (E2), metrics to analyze ergonomics (E3), and the user's height as an indicator if users have bent down (E4, E5). Here, one expert highlighted the need for an export functionality for intercompatibility with other tools: "*If I really want to calculate something with the data, I would want to do that in a different tool*" – E1. Experts also suggested extending the timeline to support aligning different sessions (E2, E3) or dividing sessions (E3) to facilitate comparisons. In addition, experts (E1, E4) suggested different forms of video playback in VR, including 360° videos (E1), displaying screen recordings on their corresponding devices (E1, E4), or adopting different points of view (e.g., from individual participants) (E5). Lastly, four experts (E1, E2, E4, E5) emphasized that importing the data must be as easy as possible, highlighting the need for a data logging toolkit: "*If it's easily applicable, then I would definitely like to use it. I even would have liked to use if for my last study!*" – E2.

Theme #4: Despite the limited extent of the prototype, the concept was appreciated

Generally, all experts (E1–5) were positive about the ReLive framework, noting that ReLive was "*intuitive*" – E1, E2, "*very positive*" – E4, "*very useful*" – E3, "*cool*" – E2, E3, E5, "*easy to use*" – E1, E2, and well-suited for the analysis of lab studies (E3). One expert cited ReLive's adaptiveness as an advantage, "*[enabling] the analysis of all kinds of studies*" – E5. Furthermore, E5 appreciated having the data in one place, while E4 mentioned that the data specification helped to create a mental map of the data. E4 also highlighted the potential for open science several times, mentioning ReLive's potential for whiteboxing analysis, providing data provenance, and helping with replication and reproducibility of research results: "*[...] and I think that is indeed a step into this direction [of replication and, generally, reproducible research]. Because for VR, it's often the case that [...] it's flexible, but not traceable*" – E4.

Since our evaluation prototype did not cover our entire concept, experts also mentioned current limitations. Experts expressed that the system is "*unfinished*" – E4 and that several details should be improved (E2, E3, E5), noting various minor usability issues (E1, E3, E4, E5). Specifically, E3 mentioned that the mapping between sessions and participants was unclear and the terminology should be clarified, while three experts (E1, E3, E5) felt that it was unclear when a session was visible in the 3D scene reconstruction. One expert also noted that pointing at objects to interact with them can be imprecise (E1). Experts also addressed the complexity of ReLive: four experts (E1, E2, E3, E5) mentioned that a preceding tutorial or training is mandatory, especially as the system was "*overwhelming*" – E5 at first. Here, experts provided the "*feature-richness*" – E5 as a reason, and ranked ReLive as an expert system with many different use cases (E2).

The experts' impression of our prototype was also reflected in the scores of the System Usability Scale [53]. With a mean Usability Score of 74.50, the usability of our prototype was rated as good [21]. Additionally, all experts expressed that they would like to use ReLive for the analysis of their next MR study.

Theme #5: Potential for collaborative analysis

A recurring topic across all experts was ReLive's potential for a collaborative analysis, which was unanimously seen as advantage. Experts provided several possible collaboration opportunities, such as looking together at the desktop (as opposed to sharing VR glasses) (E2), symmetric collaboration in VR (E2), or asymmetric collaboration with one user at the desktop and another user in VR (E2, E4, E5). Here, user could take on different roles (E1, E2, E5): for example, the desktop user could act as a director, guiding the VR user to look at the data in context (E5). This asymmetric collaboration could also remove the need to transition between a desktop and VR (E5), allowing users to maximize the advantages of both VR and desktop (E2). Experts also pointed at additional opportunities during collaboration, such as the use of a private space (E3) or handing over control (E1).

6.6 Insights and Implications

This section provides *design insights* (D 3.1–D 3.4) and *research implications* (I 3.1–I 3.4) based on our design of ReLive (Section 6.2), as well as findings from our design walkthrough (Section 6.4) and expert user study (Section 6.5). The structure follows our research questions of *RQ3.1: interplay between in-situ and ex-situ analysis* and *RQ3.2: task allocation*.

6.6.1 Interplay Between In-Situ and Ex-Situ Analysis

Although ReLive's desktop view focused on an ex-situ analysis, adding an in-situ view (i.e., 3D scene panel) proved useful, providing an at-a-glance window into the environmental context. Depending on the task complexity, this can render the full transition to VR obsolete, as the cost of switching displays may outweigh the benefits of full immersion and stereoscopy provided by VR **(D 3.1)**.

While the desktop's 3D scene panel allowed for an instant glimpse into the VR scene, the transition between the desktop and the VR environment can still be disorienting. Here, future research could investigate a more explicit switch, requiring the user to specify a position in the VR scene before allowing the user to transition into VR **(I 3.1)**. In addition, anchors in VR (e.g., displaying a virtual desktop in the actual position of the physical desktop) can provide spatial context to the user, aiding in both transition into VR and back to the desktop (cf. [284]) **(D 3.2)**. Future devices may further help this transition by gradually fading between VR and real world **(I 3.2)**, or allowing the user to work exclusively in VR.

Furthermore, a task queue can be helpful in further reducing the need to transition between devices. Here, a *cross reality linking and brushing* technique could be investigated, that could, for example, allow users to mark areas of interest (e.g., outliers) in the ex-situ view which are then highlighted in the in-situ view **(I 3.3)**. While the ex-situ view in VR should be easily accessible for an at-a-glance overview (cf. [60]), the 3D visualizations can be further improved with an on-demand 2D overview visualization.

In terms of video playback, both ex-situ and in-situ view complement each other: The ex-situ can act as a familiar 2D video player, allowing users to investigate video details (e.g., with a high-resolution desktop display). In contrast, the in-situ view helps in immersing the user in the original study setting, as videos can be played back in its original context (e.g., displaying screen recordings on a replica of the original device). The in-situ replication also enables analysts to view the study from different points of view, for example reliving the study from a participant's perspective. Considering the limited field of view of current AR and VR devices, this emulation can provide additional insights into what was actually in the user's view.

Design Insights – Interplay Between In- and Ex-Situ Analysis

- **D 3.1** Avoid needless transitions by offering basic in-situ view in desktop (and vice versa for VR).
- **D 3.2** Use anchors when transitioning between VR and desktop (cf. [284]).

Research Implications – Interplay Between In- and Ex-Situ Analysis

- **I 3.1** Investigate requiring an explicit switch before allowing transition into VR.
- **I 3.2** Investigate effect of continuous transition between the real world and VR.
- **I 3.3** Investigate benefits and challenges of cross reality linking and brushing.

6.6.2 Task Allocation

All experts agreed with our intended distribution of using a desktop view for an ex-situ overview of the data and using VR for an in-situ analysis of the details in its environmental context, akin to an overview + detail visualization: "*For me, the [desktop] view is overview, and if I need details, I'll switch to VR, and that's a coherent workflow for me*" – E5. The desktop view therefore allows experts to quickly compare data across different study sessions, find outliers, and create an analysis pipeline **(D 3.3)**. In contrast, the in-situ perspective (i.e., both 3D scene reconstruction on the desktop view and the VR environment) was used for sensemaking and reasoning tasks, focusing on *why* something happened – thus providing additional (environmental) context that cannot be found in the overview **(D 3.4)**. This is in line with prior research, which suggests that 2D outperforms immersive environments for overview tasks (cf. [223]). Here, further research could be conducted in directly comparing in-situ or ex-situ tasks in either environment (cf. [394]) **(I 3.4)**. However, the use of VR for an in-situ view on the data (as opposed to using the 3D study replication on the desktop) depends on both the task complexity (cf. [222]) as well as the user's own preference.

Design Insights – Task Allocation

D 3.3 Non-immersive visual analytics desktop view allows for analysis, ex-situ overview, finding outliers.

D 3.4 Immersive analytics view suited in-situ analysis, looking at data in context, reasoning, but also exploration.

Research Implications – Task Allocation

I 3.4 Comparison of tasks in in-situ and ex-situ scenarios (cf. [223]).

6.7 Limitations and Future Work

Our evaluation prototype of RELIVE was tailored towards evaluating the interplay between in-situ and ex-situ analysis of MR user study data. As the analysis of unfamiliar MR user study data with a novel analysis environment can be challenging, we intentionally tried to reduce the complexity of the prototype, the data, and its visualizations. While this lowered the threshold of interacting with RELIVE, it also left several aspects of the overarching concept untouched: Conceptually, the desktop view can be organized linearly like a computational notebook, allowing analysts to, for example, configure or program components that suit their analysis. We did not implement programmable component templates to reduce complexity. Further studies are necessary to study the general feasibility and applicability of such components for a mixed-immersion analysis workflow (cf. [13, 311]). In this regard, prior work by Borowski, Rädle, and Klokmose [45] has already demonstrated the benefits of providing packages (cf. components) for computational notebooks and how they can even support shared activities [46].

Future work could investigate this concept, allowing analysts to access, share, or contribute to a public collection of available metrics, visualizations, or scripts. This would in turn facilitate the idea of open science, allowing for a whiteboxing of the analysis (supporting reproducibility and data provenance) by, for example sharing the analysis notebook along with the available study data, in addition to preregistering.

To further reduce complexity, our evaluation prototype only provides simple, yet easily understandable predefined visualization templates (see Table 6.1). Given that participants were not familiar with the used reference studies, we intended to not overwhelm participants further. This allowed us to study the suitability of in-situ and ex-situ analysis for different analysis tasks. While our insights already allude to the necessity of 2D overview visualizations inside VR, additional studies are necessary to investigate the role of 2D and 3D visualizations in the different views (e.g., desktop, VR). Furthermore, future work could investigate the mapping between 2D visualizations and their 3D equivalents and how established techniques such as linking and brushing could facilitate the interplay between in-situ and ex-situ analysis.

The reduced evaluation prototype and relatively small amount of expert participants also limits the extent of our results and discussion. Due to the predefined tasks and missing comparison to existing tools, our evaluation provides only limited insights into the real world efficacy of ReLive. Thus, we concentrated on investigating the transition between different environments instead of quantitatively comparing if ReLive provides any ecologically valid benefits over existing tools. However, given that participants were experts in their respective fields, they provided valuable insights into initial opportunities and challenges. In addition, although the tasks artificially forced a transition between desktop and VR, they were grounded on our own experience when analyzing MR studies.

Although ReLive is designed to provide access to a wide range of data, gathered using an abundance of available quantitative and qualitative methods, we intentionally limited the scope of available data for the evaluation. This reduced complexity facilitated especially the free phase in the expert user study and reduced the overall duration of the user study. However, we see a great potential for the interplay between in-situ and ex-situ analysis of, for example, qualitative data such as annotating and coding of video material: Traditionally, the coding of video material is a part of the user study analysis workflow done in a desktop environment. The analyst carefully watches recordings of user study sessions to understand participants' activities, annotates remarkable events, and assigns codes to them to quantify the material for further analysis. However, having a pre-defined and static perspective of the camera can be a limitation (due to occlusion or limited resolution). Here, the in-situ analysis approach can complement the ex-situ analysis by investigating the activities from any desired point of view. Future research could investigate to what extent sophisticated video analysis tools (e.g., [216]) in combination with in-situ approaches can support MR user study analyses either *post-hoc* after completing all study sessions or even with the help of additional experimenters *ad hoc* during the runtime of a study session (e.g., adopting the participant's view).

Lastly, we intentionally focused on a single-user scenario to investigate the interplay between in-situ and ex-situ analysis. However, the findings of our expert user study highlighted the potential for ReLive as a collaborative system, supporting different constellations (e.g., collaborating on a desktop, or assuming different roles such as for overview on desktop and detail in VR), with one participant expressing that "*[...] you need another person, and, I think, then it'll be really great*" – E5. Further studies are necessary to investigate the potential benefits and challenges of such a collaborative system for the analysis of MR user studies (cf. heterogeneous remote assistance systems [77, 129]).

6.8 Follow-up Research Trajectories

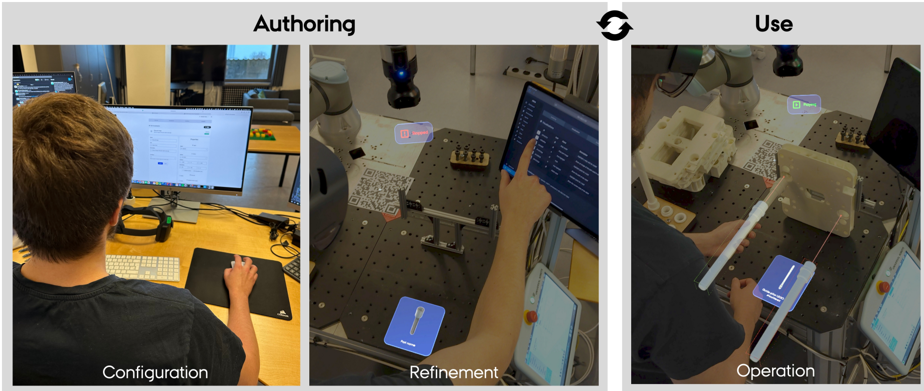
The idea of asynchronous hybrid user interfaces has since been extended and validated across a range of scenarios [3, 197, 255, 387]. One example is our follow-up work on ARTHUR, which applies this approach to the field of human–robot interaction (HRI).

In HRI, recent research has demonstrated the benefits of collaborative robots (cobots) for assisting with assembly tasks [256, 264], as they enable close collaboration between humans and robots without extensive safeguards [79]. However, despite great interest and desire to adopt these new technologies, the manufacturing industry struggles to integrate collaborative robots into their processes for various reasons, such as lacking knowledge and safety concerns. To overcome such limitations and make "*genuine*" human–robot collaboration possible, the operator must be made aware of the ongoing and planned robot procedures and be able to coordinate these. Information presented in AR can be beneficial for conveying the robot's intent [308, 378], visualizing safety information [167, 204], and highlighting the user's tasks and procedures [103, 256].

However, the process of designing, testing, debugging, and refining AR content for HRI is often tedious and challenging: AR setups for HRI are typically authored on a desktop system, requiring accurate simulation of the real-world environment in which the virtual content should be anchored. Even then, there is often a notable discrepancy between the simulation of AR content and its situated visualization upon deployment on the AR HWD. To break out of the sequential design-deploy-refine process that requires time-consuming alternation between a desktop and an HWD, prior work has demonstrated the potential of authoring visualizations directly in AR (i.e., *in-situ*) [258]. Yet, such in-situ authoring in AR can limit the user's effectiveness compared to desktop-based interaction.

To address these challenges, we created ARTHUR, an open-source[3] AR-based **A**utho**r**ing **T**ool for **Hu**man–**R**obot collaboration scenarios that supports the creative workflow through a hybrid user interface (i.e., across a desktop, tablet, and

[3] https://gitlab.au.dk/arthur



Supplemental Video

Figure 6.7: ARTHUR is an authoring tool for augmented reality-supported human–robot collaboration. It supports (1) creating the initial system configuration in the web interface on a PC (left), (2) refining the setup in-situ using a hybrid of web interface (tablet) and AR interface (center), and (3) testing and using the authored system on the AR interface through an AR HWD (right).

HWD, see Figure 6.7). We expand upon prior work by facilitating in-situ authoring not only of visualizations but also of user actions and conditions, thus creating a holistic AR environment for designing human–robot collaboration processes. Besides enabling the general setup and in-situ authoring, our proposed system fluidly facilitates switching to the operation phase, allowing users to instantaneously try out their authored workflows on one or more robots. ARTHUR thus combines not only *exclusive* temporal usage by switching between a desktop (configuration phase) and AR environment (refinement and operation phase), but also utilizes the advantages of a synchronous hybrid user interface during the refinement and operation phase.

Although the findings of the evaluation are once again not part of this thesis, we demonstrated the potential of ARTHUR by replicating a variety of scenarios from prior work and evaluated the usefulness of the supported features and the suitability of our hybrid user interface approach in a usage evaluation with experts. Follow-up research by Lunding, Feuchtner, and Grønbæk [257] further validated ARTHUR in an in-the-wild study, confirming its feasibility and the effectiveness of such an asynchronous hybrid user interface.

Together, ReLive, ARTHUR, and subsequent efforts by other researchers illustrate the broad applicability of integrating existing workspaces and practices with novel immersive MR hardware. While asynchronous use may not take full advantage of the benefits of a hybrid user interfaces (as discussed in Section 3.5.6) and such *exclusive* usage might be better described as *cross reality*, bridging the gap between environments still allows users to benefit from the strengths of each – ultimately enabling more *fluid* workflows.

However, the physical transition between devices remains challenging. Users still experience a physical (and thus cognitive) break when switching from a desktop setup to a VR environment. Future HWDs may be able to replace traditional desktops as they become more comfortable. Yet, the principles of connecting 2D and immersive environments established in this chapter remain relevant and transferable to such scenarios [387], regardless of their actual output modality. In terms of input, however, switching between dedicated input devices, such as mouse and VR controllers, continues to be a source of friction that has not been addressed thus far.

6.9 Conclusion

The "*ReLive*" exemplar showcases the use of an asynchronous hybrid user interface with *exclusive* device usage for analyzing MR user studies. The immersive VR environment allows users to relive an interactive recording of a replica of the original study, providing the possibility for in-situ analysis of the data. In contrast, the non-immersive desktop view facilitates the analysis of aggregated study data and provides a holistic overview of the available study data. ReLive also supports the transition between the VR and desktop environment, for example by synchronizing both environments in real-time and offering a glimpse of the VR environment on the desktop and vice versa – thus representing an asynchronous hybrid user interface.

By investigating the distinct roles of devices for different tasks, ReLive contributes to the research objectives **RO1: Transitioning Between Devices** and **RO2: Task Allocation**. However, achieving a fully fluid interaction remains challenging due to the inherent break when switching between devices. The following chapter explores a potential solution to further bridge this gap, tightly integrating these two complementary environments.

Chapter 7

Exemplar: SpatialMouse

Chapter Content

This chapter is based on the publications:

Sebastian Hubenschmid, Johannes Zagermann, Robin Erb, Tiare Feuchtner, Jens Grubert, Markus Tatzgern, Dieter Schmalstieg, and Harald Reiterer. "SpatialMouse: A Hybrid Pointing Device for Seamless Interaction Across 2D and 3D Spaces." In: *Proceedings of the 2025 31st ACM Symposium on Virtual Reality Software and Technology*. New York, NY, USA: ACM, 2025, pp. 1–13. ISBN: 979-8-4007-2118-2. DOI: 10.1145/3756884.3766047

Sebastian Hubenschmid, Johannes Zagermann, Niklas Elmqvist, Tiare Feuchtner, Jens Grubert, Markus Tatzgern, Dieter Schmalstieg, and Harald Reiterer. "Revisiting Hybrid Input Devices for Immersive Analytics." In: *2025 IEEE Conference on Human Factors in Immersive Analytics (HFIA)*. IEEE, 2025. DOI: 10.1109/HFIA68651.2025.00011

Figure 7.1: The *SpatialMouse* combines a desktop mouse with a virtual reality controller, enabling seamless interaction across 2D and 3D information spaces in immersive mixed reality environments.

7.1 Chapter Context

By combining the distinct input and output modalities of 2D and 3D devices, hybrid user interfaces leverage the strengths of both paradigms for their respective environments. This works especially well when used *synchronously* (i.e., *parallel* or *serial* temporal usage, as discussed in Section 3.5.6 and exemplified in Chapters 4 and 5), as potential downsides of one device are immediately compensated by the other device – forming a symbiosis of interfaces [424].

Although ReLive (see Chapter 6) demonstrates the value of linking two exclusive environments (i.e., desktop and VR environment) in an *asynchronous* hybrid user interface, it also highlights the significant gap between these two environments compared to synchronous solutions: Despite combining the benefits of a desktop with VR environments, users still need to physically switch between a desktop and VR environment, regardless how interconnected the software is. This practical inconvenience poses a significant disruption to the user's flow – and thus hinders fluid interaction.

Current HWD can overcome this gap in terms of output, as they can now sufficiently simulate physical 2D screens [155, 202, 310]. Yet, the affordances of physical input devices are more difficult to simulate as their capabilities and constraints vary widely. As a result, users must still either intentionally switch between input devices for different information spaces – incurring significant transitioning costs, even with hand tracking [40, 87] – or work with limited input devices, for example, using a desktop mouse for 3D input [16, 86].

This situation is unfortunate, since mouse input still offers high accuracy and familiarity for 2D surfaces despite lacking the capabilities for full spatial interaction. In contrast, VR controllers excel at 3D manipulation in immersive environments, but face the challenges of spatial interaction: Users may face hand jitter [36] or a Heisenberg effect of spatial interaction [48, 411]. Free-hand interaction using hand tracking [40, 87] or touchpads [359] could facilitate device switching. However, physical input devices offer significant advantages in terms of task completion time and accuracy [69, 156, 260] in addition to superior ergonomics and user preferences [260]. Hybrid user interfaces overcome this dichotomy by combining multiple devices through software – why not address this challenge by combining a mouse and VR controller into one *hybrid* input device?

This chapter presents one potential solution by introducing the *SpatialMouse* (see Figure 7.1) – an input device for the "*era of spatial computing*". The *SpatialMouse* enables seamless interaction between 2D surfaces and 3D spaces, thus representing a hybrid pointing device suitable for both indirect 2D pointing as well as for 3D input with six degrees of freedom (6DoFs). Although not a hybrid user interface per se, the *SpatialMouse* has the potential to support fluid interaction within asynchronous hybrid user interfaces by combining the capabilities of a desktop mouse with a VR controller as one unified device: When used on a table, the *SpatialMouse* acts as a conventional 2D mouse suitable for on-surface operation; when the device is lifted, it automatically switches to a spatial interaction mode, acting as a VR controller with 6DoFs suitable for 3D spaces.

In addition to eliminating a switch of input devices, the *SpatialMouse* enables several novel hybrid interaction techniques. We demonstrate the capabilities the *SpatialMouse* using two approaches: First, we implemented an initial physical prototype, which was evaluated in a user study with 12 participants. Second, we discuss a set of novel interaction techniques for 2D surfaces and 3D spaces enabled by this hybrid input approach.

7.1.1 Research Questions

Our overarching goal in developing the *SpatialMouse* was to eliminate the need to switch between input devices. While the novel interaction techniques enabled by this hybrid input device contribute to **RO3: Interaction Techniques**, the *SpatialMouse* directly addresses **RO1: Transitioning Between Devices**. Towards this goal, the *SpatialMouse* must offer performance comparable to established, dedicated input devices. Accordingly, our user study focuses on the following three research questions to investigate performance metrics:

RQ4.1 Performance
How does the *SpatialMouse* affect *performance* when switching between 2D surfaces and 3D spaces?

RQ4.2 Task Load
Does the *SpatialMouse* impact the user's perceived *task load* in tasks that require transitioning between 2D and 3D environments?

RQ4.3 Usability and User Experience
In what way is *usability and user experience* influenced by the *SpatialMouse*?

To inform the design requirements of such a hybrid input device, we first examined three key areas: First, we explored specific *transitional workspaces* to identify relevant task demands; second, we analyzed existing *dedicated input devices* to guide the needs of the hybrid device; and third, we investigated *shared input paradigms across 2D and 3D spaces* to inform the seamless transitions.

7.1.2 Transitional Workspaces

Prior work in the area of cross reality finds that "*most cross reality workflows involve short, temporary movements between VR and desktop*" [406] and that users "*were more willing to transit between [desktop] and VR when the transition cost was lower*" [387]. Several systems have therefore explored the tight integration between 2D monitors and MR environments, such as analyzing data from motion capture [60, 186] and volumetric medical scans [43, 367] or 3D modeling [333, 334]. All these areas represent potential use cases for the *SpatialMouse*.

Although prior work often considers the combination of physical monitors with immersive environments, a contemporary HWD can simply show "*virtual monitors [that] can be used now for real-world work*" [310], thereby enabling purely digital workspaces. Such 2D surfaces can offer distinct advantages beyond the restrictions of physical monitors [152, 310], such as arbitrary sizes or placement within the user's environment. For example, ZIGEN [213] demonstrates how holistic window management for VR environments could integrate transitions between 2D windows and 3D models. Zwin[1] demonstrates how a holistic operating system for VR could involve both 2D windows and 3D "*windows*", thus requiring not only precise 2D interaction but also 3D manipulations.

7.1.3 Dedicated Input Devices

While there is an abundance of devices and modalities for 2D input (e.g., touch, stylus, gaze), this chapter focuses on the computer mouse, with its long history in HCI[2] as the most established input device for desktop environments. Although alternatives are commonplace for certain devices (e.g., touchpads for laptops, direct touch for smartphones), prior research has shown that they can be significantly slower [359] and less accurate [341] than mouse input.

The consumer market is saturated with custom mice, addressing not only minutiae such as weight and latency [397, 417] but also offering diverse form factors with different trade-offs: While horizontal mice offer more performance, vertical mice promote ergonomic wrist posture [2, 251, 295, 319, 373]. Subsequent research therefore proposes a slanted mouse that optimizes ergonomics without compromising performance [251]. Since the *SpatialMouse* should be

[1] www.zwin.dev, last accessed 2025-05-12.

[2] Please refer to Bill Buxton's unpublished manuscript "*Human Input to Computer Systems: Theories, Techniques and Technology*" (www.billbuxton.com/inputManuscript.html) and Horst Oberquelle's "*Computer Museum*" (www.fundus.uni-hamburg.de/en/search/expert?collection=computer&q=mouse), both last accessed 2025-05-12.

easily grippable when acting as a VR controller, we base our initial prototype design on such slanted mice, which are akin to off-the-shelf VR controllers.

Similarly to the mouse for 2D input, physical 6DoF controllers have established themselves as the leading device for input in MR environments [177], despite the wide variety of available 3D input devices [160, 351]. Although free-hand input (i.e., hand tracking, gaze [313]) is becoming more prevalent, studies show that physical controllers can outperform free-hand input [260], in part due to their tactile feedback [36] and limitations of hand-tracking hardware [69, 156].

Given the advantages of dedicated input devices, prior research has investigated explicitly switching between devices. All studies conclude that such a switch incurs a significant overhead [40, 87, 315]: For example, a recent study by Cools et al. [87] shows that switching between free-hand interaction and mouse input incurs significant penalties for tasks that require constant switching between 2D and 3D spaces. We assume that this overhead is even more pronounced for physical controllers, especially since users may not want to pick up another device after working with mouse and keyboard [400]. As a result, prior work proposes a "*universal mouse-pointer to handle both 3D and 2D interaction [to reduce] functional discontinuity in the transition between [3D manipulation] and WIMP*" [71].

7.1.4 Shared Input Paradigms Across 2D and 3D Spaces

Adapting traditional 2D input devices for interaction within immersive 3D spaces is an ongoing research topic [272]. Several recent studies demonstrate the usefulness of traditional input devices for text entry in MR [153, 154, 270]. Prior research has also explored translating mouse input into 3D space [10, 86, 290, 430] or using touch interaction techniques to extend to 3D space [30, 85, 95]. These approaches work especially well with spatial AR [351], which projects the mouse cursor on physical surfaces rather than allowing it to hover in free space [211, 267, 337]. However, given the 2DoF of mouse input, such approaches fall short of 6DoF devices for manipulation tasks [230, 428]. It has been argued that the "*lengthy and dominant reign of the 2D mouse is at an end when it comes to 3D systems*" [203].

The opposite may be true for input devices with 6DoF: They show great potential for 3D manipulation tasks [428], but perform notably worse than dedicated 2D input devices (mouse, touch) to interact with 2D information spaces [54, 87, 194]. This can be partially alleviated by providing users with a desk to stabilize their arms [78, 100]. Note that this approach implicitly provides the user with the option to switch to a dedicated mouse.

Research has therefore explored other input devices suited to both 2D and 3D spaces. Pen input devices, such as the Eye of Ra [43, 44] or SpatialTouch [429], can offer a natural mapping for direct 2D and 3D input. While some researchers suggest that pen input can be a good choice for certain scenarios (e.g., medical use cases [265, 324]), others show that pen input performs worse than 6DoF controllers in task performance and user experience for general 3D sketching and docking tasks [275]. Alternatively, hybrid user interfaces can be beneficial for interacting with 2D and 3D environments, but their ergonomics and efficiency are limited due to the physical constraints of the 2D devices (e.g., a touch screen).

Thus, we still see the mouse and traditional 6DoF controllers as the optimal choice in their respective environments. To bridge the gap between these two devices, previous research has adapted mice to support more than 2DoF: Examples include the roller mouse [393], Rockin'Mouse [16], VideoMouse [170], two-ball mouse [263], GlobeMouse [133], or the commercially-available SpaceMouse[3]. However, these examples employ isometric approaches (i.e., with resistance), which are slower for interaction in 3D space than isotonic inputs (i.e., without resistance) [404]. In contrast, "*flying mice*" [427] such as the bat [403] or the Cubic Mouse [134] support isotonic input with 6DoF but seem more comparable to contemporary MR controllers, as they lack support for mouse operation. The devices closest to our *SpatialMouse* are Logitech's "*MX Air*"[4], Sodial's "*Air Mouse*"[2], and Simgraphics' "*Flying Mouse*"[2], which supports both mouse operation and isotonic input with 6DoF. To the best of our knowledge, there are no studies that evaluate the efficacy or ergonomics of this device compared to a combination of mouse and MR controller or investigate potential interaction techniques offered by such a device. Our *SpatialMouse* further differs by its design for spatial interaction in MR environments, taking the form of an MR controller and slanted mouse for improved ergonomics. In contrast, devices such as the "*Flying Mouse*" retain the shape of a horizontal mouse, which could restrict its suitability as VR controller.

7.2 Use Cases

Our work is motivated by two main scenarios from prior work that span 2D surfaces and 3D spaces, requiring users to switch between these environments. Apart from our presented following use cases, we envision the *SpatialMouse* as a viable input alternative for various scenarios that span across 2D surfaces and 3D spaces.

7.2.1 Immersive Analytics

This use case draws inspiration from cross reality workspaces such as HybridAxes [356] and ReLive [186] (see Chapter 6). Given the distinct strengths and limitations of 2D and 3D visualizations [50, 360, 379], we argue using both can enable more effective and flexible workflows: 2D visualizations are well-established and effective for providing an overview of dense information, which requires precise interaction. In contrast, 3D visualizations are well-suited for inherently spatial data (e.g., volumetric structures, motion trajectories) or providing environmental context (e.g., situated analytics [112]). In addition, a complementary 3D view to a 2D view can offer higher precision in spatial tasks [388].

[3] https://3dconnexion.com/spacemouse/, last accessed 2025-05-12.
[4] https://ifdesign.com/winner-ranking/project/mx-air/36413, last accessed 2025-05-12

We envision data analysts fluidly [111] transitioning between 2D surfaces and 3D spaces throughout their workflow. For example, to evaluate a mixed reality user study, an analyst might start with a 2D overview of aggregated data to identify patterns and outliers [186]. In this phase, mouse input is ideal due to its precision and low ergonomic demand, enabling efficient selection and filtering even in dense datasets. To explore spatial data such as motion patterns, the analyst can drag relevant 2D visualizations into the surrounding 3D space [327, 356], where they automatically transform into interactive 3D representations [234, 244, 355]. These visualizations can then be spatially arranged or manipulated using a VR controller, allowing the analyst to examine data from different perspectives or select specific regions within the 3D visualization. By seamlessly switching between 2D and 3D views, the analyst can triangulate insights across dimensions by leveraging established techniques such as linking and brushing [208].

7.2.2 3D Modeling

3D modeling is central in domains such as computer-aided design and visual effects. This use case is primarily motivated by cross reality workspaces such as Myr [34] and ZIGEN [213], which tightly integrate desktops with immersive VR environments. Spatial interaction using 6DoF controllers supports 3D modeling through one-to-one mapping for 3D manipulation that mimics real-world object handling and natural navigation [224]. However, research has also highlighted key limitations of spatial input for design tasks. Given the ergonomics of extended mid-air interaction [168, 307], designers tend to produce simpler 3D models to reduce strain [278]. In addition, fine motor control can be challenging without arm support [78, 100], and biomechanical constraints may cause translations to be accompanied by unintentional rotations [169]. Research thus proposes "*to split intrinsically 3D operations from operations that are ideally suited to 2D*" [100], for example by implementing constraints in software [126]. Unsurprisingly, 2D mouse input can provide higher precision for fine-grained adjustments in 3D modeling [31]. As a result, prior work has explored using mobile devices to provide more precision and control in 3D modeling applications [279, 333].

We envision 3D designers switching between 2D and 3D input as their task demands. A typical workflow might begin on a 2D desktop, where a designer uses familiar mouse-based input to navigate files and open a project. From there, relevant models can be dragged from the 2D surface into the surrounding 3D space, where they are transformed into fully interactable objects [213]. Designers can thus perform spatial operations such as extrusion and sculpting, or isolate parts of the 3D model for detailed refinement. However, for fine-grained 3D manipulations (e.g., during translations), designers may prefer the precision and stability afforded by 2D mouse input [31], taking advantage of its physical constraints. In addition, designers may need to constantly switch between editing 2D texture maps and adjusting their placement on the 3D model [224], highlighting the need for a seamless transition between mouse operation and 6DoF controllers.

7.3 SpatialMouse

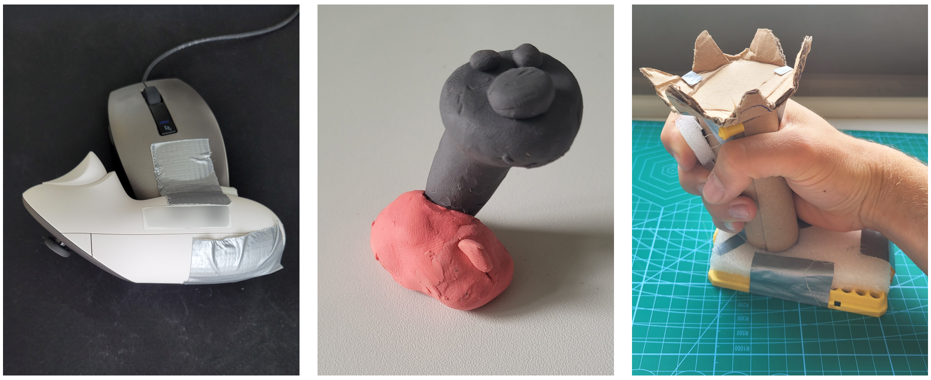
With the increasing relevance and viability of transitional workspaces and asynchronous hybrid user interfaces, we see the need for an appropriate input device that supports users in the seamless transition between 2D surfaces and 3D spaces. Our *SpatialMouse* represents one possible solution (see Figure 7.2), which allows users to seamlessly switch between indirect 2D pointing input for mouse operations and 3D input with 6DoFs for spatial interactions.

7.3.1 Interaction Concept

As its name implies, the *SpatialMouse* differentiates between two distinct modes: *mouse mode* and *spatial mode*. The user maintains the same hand posture across both modes, eliminating the need for grip adjustments when switching.

Mouse Mode. The *mouse mode* is automatically activated when the *SpatialMouse* is placed on a surface, allowing the user to operate the *SpatialMouse* like a slanted mouse. Users can indirectly control a cursor on a 2D surface (e.g., desktop monitor) by moving the device like a mouse. In *mouse mode*, input to the cursor is relative and may require clutching: The clutch can be activated by lifting the *SpatialMouse* slightly and moving it back to a neutral position.

Figure 7.2: The *SpatialMouse* offers controls similar to a mouse and VR controller. The controls are mapped depending on the current mode: In *mouse mode*, the buttons are mapped to the left and right click, while the joystick acts as a mouse wheel replacement. In *spatial mode*, the buttons are mapped to the primary trigger and secondary button, while the joystick allows for 2D input. A spatial tracker at the top ensures consistent tracking in 3D space, while a baseplate on the bottom provides a wrist rest in *mouse mode*.



(a) Low-Fidelity Mockup (b) Clay Modeling (c) Cardboard Prototype

Figure 7.3: We followed an iterative design process for the *SpatialMouse* prototype: (a) starting with a low-fidelity mockup by attaching a VR controller to a mouse to explore ergonomics and input needs, (b) using clay modeling to test form factors for both mouse and spatial usage, and (c) creating cardboard prototypes to refine button placement, hand positioning, and overall size.

Left and right clicks can be triggered by pressing the upper and lower buttons with index and middle fingers, respectively. A joystick mounted near the user's thumb allows for scrolling input, similar to a scroll wheel: moving the joystick up and down moves the content up or down at a speed relative to the joystick's tilt angle, providing precise and adjustable scrolling control. Similarly, moving the joystick back and forth allows for horizontal scrolling, while pressing the joystick simulates clicking on a scroll wheel.

Our initial prototype does not include any additional buttons as often found in many commercial mice (e.g., back, forward). This omission is due to space constraints inside our physical prototype and could easily be addressed by industrial design methods.

Spatial Mode. The *SpatialMouse* automatically activates the *spatial mode* when the device is picked up. In *spatial mode*, the device provides 6DoF for translating and rotating virtual objects in 3D space. This can be used for direct or indirect 3D manipulations (e.g., via raycasting). Given the lack of standardized control schemes for VR controllers (e.g., touchpads or joysticks, button placement, amount and type of buttons), we designed the *SpatialMouse* to support a basic set of inputs, with a primary and secondary button, as well as a joystick. While the button mapping is application-dependent, the upper button acts as the primary input, similar to the trigger of a contemporary VR controller such as HTC Vive or Meta Quest 3 controllers. The lower button acts as a secondary button similar to the grip button on Meta Quest 3 controllers or clicking on the touchpad of an HTC Vive controller. In our prototype, the primary button is mapped to grabbing an object, while the secondary button allows object deletion. Similarly, the joystick can be mapped to various 2D inputs for fine-grained control in mid-air.

7.3.2 Implementation

In the following, we describe the *form factor*, *hardware components*, and *software* of our *SpatialMouse*.

Form Factor. We followed an iterative design approach for the form factor of our *SpatialMouse* (see Figure 7.3). Starting with a VR controller attached to a mouse to gain an initial understanding of general ergonomics and input placement, we built several clay models and cardboard prototypes to refine the shape of the device before arriving at our current design. Given that the *SpatialMouse* is partially intended to be used as a VR controller, we found the form factor of horizontal mice to be insufficient, especially when used for mid-air interaction. Instead, we adopted a design similar to slanted mice and VR controllers such as the HTC Vive controllers, thus promoting ergonomic wrist posture in *mouse mode* [2, 251, 295, 319, 373] and ensuring a secure and comfortable grip in *spatial mode*.

Our current prototype has a total height of 17.5 cm and weighs 282 g with an HTC Vive tracker mounted on top. The case was 3D printed in individual parts, then manually assembled using bolts and glue to allow easy access to its internal components (see Figure 7.4). A baseplate at the bottom allows users to rest their hands when operating in *mouse mode*. A standardized screw on the top supports the mounting of different tracking technologies.

Hardware Components. We use off-the-shelf commodity components with an Arduino Nano ESP32 for input processing and Bluetooth LE 5, ensuring our work can be easily replicated and improved upon. We opted for a PMW 3389 optical sensor (as opposed to disassembling a commercial mouse), as it allows for the

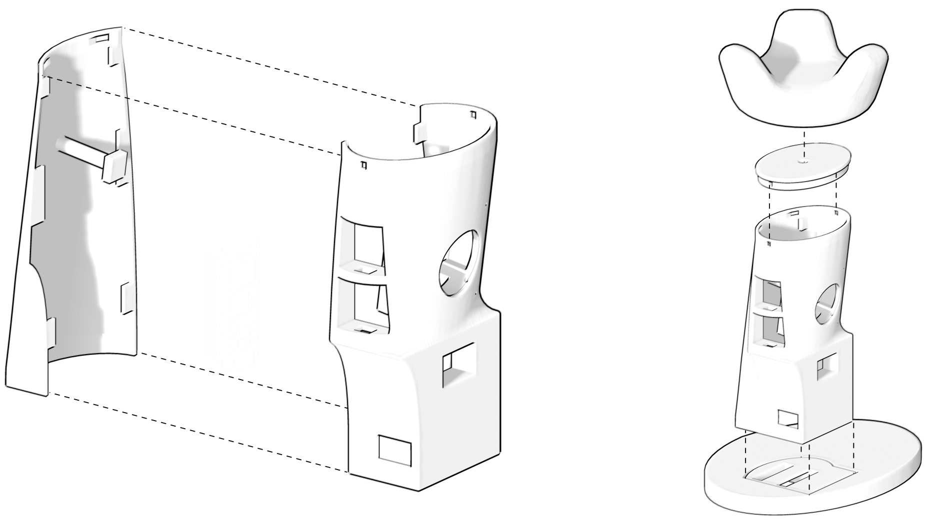


Figure 7.4: The *SpatialMouse* uses a modular 3D-printed case, allowing for easy assembly and access to internal sensors and batteries. A standardized UNC 1/4″-20 screw at the top allows for secure mounting of different tracking technologies.

integration of custom firmware, which is necessary for automatically switching between modes. The mouse sensor, buttons, joystick, and on/off switch are directly wired to the Arduino module. To ensure optimal spatial tracking, we mount a separate tracker on the top of the *SpatialMouse*, which is otherwise not connected to the internal electronics.

Software. The *SpatialMouse* uses custom firmware for the Arduino to process and send sensor input to the desktop as a "*mouse*" human-interface-device. Most notably, we use surface quality readings from the optical sensor to automatically switch between modes: If the surface quality is below a predefined threshold (i.e., mouse is lifted from surface), the internal logic switches to *spatial mode*. This allows us to keep the *SpatialMouse* self-contained and send different button events based on the mode over its external interface.

7.4 User Study

We conducted a user study to evaluate the feasibility of the *SpatialMouse* and to understand how such a hybrid pointing device can facilitate the transition between 2D surfaces and 3D spaces (**RO1: Transitioning Between Devices**). We compared the performance, perceived task load, usability, and user experience of our *SpatialMouse* with a combination of dedicated input devices (mouse and VR controller). We decided on a within-subjects study design to allow for a direct comparison with dedicated input devices as a state-of-the-art baseline.

7.4.1 Conditions

Our study involves two conditions: The **baseline** condition includes dedicated input devices (i.e., a mouse and a VR controller) for 2D surfaces and 3D spaces, respectively – involving switches between the distinct modalities. In contrast, the ***SpatialMouse*** condition does not involve switching between distinct modalities, but its usage leads to *mouse mode* or *spatial mode*. To account for learning and carryover effects, we counterbalanced the order of conditions and tasks (i.e., half of the participants started with the baseline condition, while the other half started with the *SpatialMouse*).

7.4.2 Task

We designed an abstract task to model workflows that involve frequent transitions between 2D and 3D information spaces (see Section 7.2 and Figure 7.5). Participants interact with a simulated *desktop* (i.e., a 2D WIMP interface, cf. [87, 315]) and its surrounding 3D VR environment. We chose VR instead of an AR environment to avoid issues with screen legibility and disparities when transitioning cubes from 2D to 3D, thereby following research on VR workspaces [38, 202, 296]. The task combines drag-and-drop, docking, and construction operations: In each step, a colored icon is dragged from the 2D desktop and dropped outside in

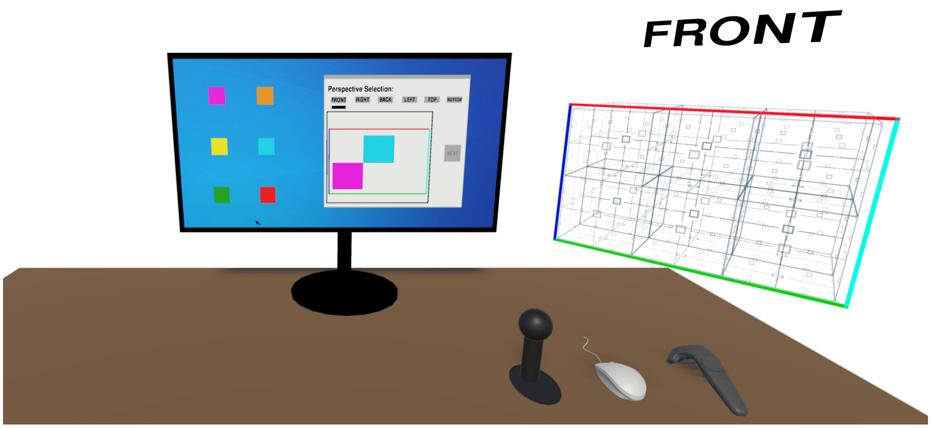


Figure 7.5: In this study task, users had to switch between a simulated desktop (left) and its surrounding 3D environment (right). 3D models of the *SpatialMouse*, mouse, and VR controller were shown in VR, depending on the condition.

3D space to instantiate a 3D cube, which is then placed in a 3D grid. We used a step-by-step workflow that transfers one item at a time from 2D to 3D, enforcing sequential interaction and ensuring frequent environment transitions. We use the terms *mouse* and *spatial* to refer to the *SpatialMouse*'s input modes, corresponding to traditional mouse input and VR controllers, respectively.

The desktop displays a pool of icons and an interactive plan with six 2D perspectives (top, bottom, front, back, left, right) of the 3D grid. Participants switch between views to determine the correct position of each cube. A color-coded border around the 3D grid aligns with the active perspective to support spatial mapping. Each cube must be placed within a threshold of 1 mm (position) and 5° (rotation). Incorrect cubes can be deleted via the VR controller using the secondary button.

To enforce device transitions, desktop interaction is restricted to the mouse (or *mouse mode* of the *SpatialMouse*), while 3D manipulation uses the VR controller (*spatial mode*). Although the *SpatialMouse* supports seamless interaction, we deliberately split the task to match the baseline's device switching. Once all cubes are correctly placed, the grid is highlighted in green, and participants proceed by clicking the "*next*" button on the desktop. During training trials, visual feedback is provided on individual placements. We designed two task sets with each six different yet comparable tasks (i.e., variations of item positions and combinations). We balanced the order of task sets by alternating the starting task set depending on the individual condition.

7.4.3 Measurements

We group our measurements based on our research questions of *performance*, *task load*, and *usability and user experience*. We evaluate *performance* based on

(a) Mouse (b) VR Controller (c) SpatialMouse

Figure 7.6: Devices used during the study[5]. The mouse was fitted with a custom 3D-printed attachment for Optitrack targets.

task completion time, accuracy of the 3D objects placed, and usage of input modality. Task completion time was measured in milliseconds from the start of the task to the successful completion. Accuracy was calculated as the Euclidean and angular distance between the final and target positions and rotations of each cube. We also analyzed the usage of input modes by measuring the time spent in mouse, switching, or VR modes, based on device movement. For the baseline, switches were counted when one device stopped and the other began moving. The *SpatialMouse* required no switching; instead, input events (e.g., button clicks) distinguished clutching from *spatial mode*. We assess *task load* using the raw NASA TLX [164] and *usability* with the SUS questionnaire [53] at the end of each condition. We asked participants about their *user experience* in a final semi-structured interview.

7.4.4 Apparatus

For all conditions, participants wore a Valve Index HWD (1440 × 1600 px per eye, 120 Hz refresh rate, 108° field of view) connected to a desktop computer (Intel i7 7700K, Nvidia GTX 1080 Ti, 32 GB RAM). Four Valve Base Stations in each corner of the room ensured consistent tracking. For the *baseline* condition, participants switched between a Microsoft Surface Precision Mouse and an HTC Vive Pro controller. Participants were asked to hold a separate, powered-off VR controller in their left hand to ensure they switched between devices. The VR controller has a similar shape to the *SpatialMouse* and does not contain a wrist strap. The mouse was connected via Bluetooth and fitted with targets for Optitrack cameras (see Figure 7.6). We manually aligned the mouse model once to match its position in the VR scene. In contrast, the *SpatialMouse* condition used the prototype as described in Section 7.3. The *SpatialMouse* was tracked with an HTC

[5]VR controller image taken from vive.com/us/accessory/controller/, last accessed 2025-05-12.

Vive Tracker, ensuring the same tracking quality as the VR controller. The weight of the *SpatialMouse* was 282 g, the VR controller was 205 g, and the weight of the mouse, including a tracking attachment, was 148 g.

The project was implemented in Unity 2022.3 and is available as an open-source project. Within the VR scene, the desktop monitor simulates a common 16:9 display (~90×48 cm). Each 2D icon on the desktop is 5×5 cm in size and expands to a 30×30 cm cube in the 3D environment, resulting in a size of 90×60×60 cm for the 3D grid.

7.4.5 Participants

We recruited 12 participants from the local university (1 diverse, 2 female, 9 male) between the age of 21–29 ($M = 24.17$, $SD = 2.37$). All participants were undergraduate students from different fields (e.g., computer science, biology). All participants indicated to use their right hand for mouse operation. On a Likert scale from 1 (inexperienced) to 5 (experienced), they rated their mouse proficiency as mostly experienced ($M = 4.42$, $SD = .67$). Prior VR experience was mixed ($M = 3.33$, $SD = 1.5$), with three participants having no prior VR experience. We recruited participants using flyers advertising the *SpatialMouse*.

7.4.6 Procedure

Participants signed a consent form, completed a demographic questionnaire, and received an introduction to the study, task, and input methods. They were instructed to complete the task as quickly and accurately as possible. After putting on the VR HWD, they completed two training tasks (with two cubes, then five cubes) for each condition to become familiar with the system and resolve potential issues. This was followed by six main tasks (five cubes each; see Section 7.4.2). After each condition, participants filled out the raw NASA TLX [164] and SUS [53] questionnaires. Each session ended with a semi-structured interview covering the *SpatialMouse*, missing features, and potential use cases. Sessions took 60–70 minutes, and participants were compensated. The study followed university guidelines for ethics and safety.

7.4.7 Results

Our results are summarized in Figure 7.7. We analyzed the data with a non-parametric approach using a pairwise Wilcoxon test, where appropriate, as a Shapiro-Wilk test indicated that our data did not follow normal distribution. We therefore indicate the medians (*Mdn*) and standard deviations (*SD*) where applicable and assume $\alpha = .05$ for statistical significance. Due to a technical issue, we imputed the final baseline task of Participant 1 with the median of their previous baseline tasks [145]. We transcribed audio recordings from the concluding interview using OpenAI Whisper and verified used quotes manually.

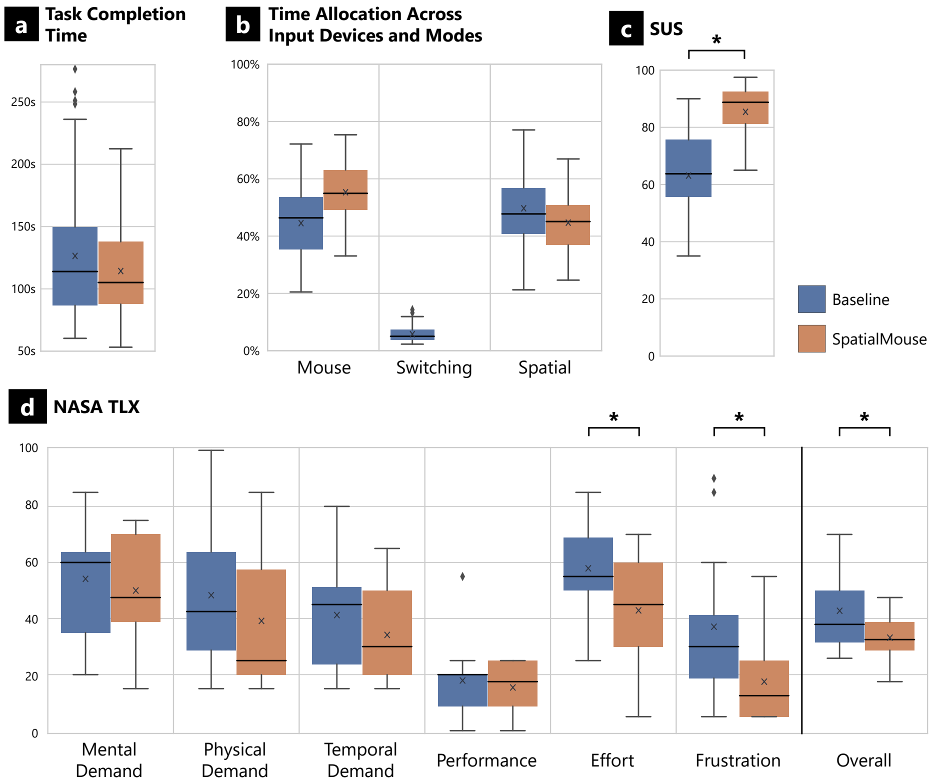


Figure 7.7: Summary of results. * marks significant differences, lines indicate medians, and crosses represent averages.

Two authors iteratively clustered participants' statements thematically using an affinity diagramming approach.

Performance. Task completion time shows no significant differences ($z = -1.448$, $p = .149$) between the baseline ($Mdn = 113.71\,s$, $SD = 51.24\,s$) and *SpatialMouse* ($Mdn = 105.11\,s, SD = 39.34\,s$) condition. There were also no significant differences in position accuracy ($z = -0.722$, $p = .47$) between baseline ($Mdn = 9\,mm$, $SD = 6\,mm$) and *SpatialMouse* ($Mdn = 9\,mm$, $SD = 5\,mm$). Similarly, we found no significant differences in rotational accuracy ($z = 1.124$, $p = .261$) between baseline ($Mdn = 2.91°$, $SD = 1.9°$) and *SpatialMouse* ($Mdn = 2.9°$, $SD = 2.03°$). In terms of time spent in each mode, participants spent between 2.29–14.31% of each task switching between devices ($Mdn = 5.01\%, SD = 2.68\%$) in the baseline condition. Participants spent less time using the mouse during baseline ($Mdn = 46.367\%$, $SD = 11.99\%$) than in its equivalent *mouse mode* during the *SpatialMouse* condition ($Mdn = 54.92\%$, $SD = 9.66\%$). In contrast, participants spent more time using the VR controller during baseline ($Mdn = 47.74\%$, $SD = 12.05\%$) than in *spatial mode* during *SpatialMouse* ($Mdn = 45.08\%$, $SD = 9.66\%$). Since there is no switch in the *SpatialMouse* condition, we refrain from reporting on potentially misleading statistical analyses.

Task Load. The overall score reveals significant differences ($z = -2.118$, $p = .038$) between baseline ($Mdn = 37.92$, $SD = 15.95$) and *SpatialMouse* ($Mdn = 32.5$, $SD = 9.69$). Effort is significantly higher ($z = -2.314$, $p = .022$) in baseline ($Mdn = 55$, $SD = 18.64$) than in *SpatialMouse* ($Mdn = 45$, $SD = 19.82$). Frustration was also significantly higher ($z = -2.141$, $p = .036$) in baseline ($Mdn = 30$, $SD = 27.51$) than in *SpatialMouse* ($Mdn = 12.5$, $SD = 15.3$). We found no other statistically significant differences for the subscales mental demand, physical demand, temporal demand, and performance.

Usability and User Experience. Analysis of the SUS questionnaire shows statistically significant higher scores ($z = -2.746$, $p = .007$) for the *SpatialMouse* ($Mdn = 88.75$, $SD = 9.88$) compared to the baseline condition ($Mdn = 63.75$, $SD = 16.96$). Overall, participants described the *SpatialMouse* as "*intuitive*" ($n = 7/12$) and "*comfortable*" ($n = 6/12$). However, participants noted that the baseplate interfered with wrist movement in *spatial mode* ($n = 6/12$), that the mouse sensor ($n = 5/12$) and buttons ($n = 4/12$) felt worse than those of a commercial mouse, and that the *SpatialMouse* felt heavy ($n = 4/12$). Participants suggested making the *SpatialMouse* more ergonomic by shaping the device to fit one's hand better ($n = 2/12$) and using a more slanted ($n = 1/12$) or shape-changing ($n = 2/12$) design.

All participants appreciated that the *SpatialMouse* reduced the effort of switching, describing the switch as "*instinctual*" ($n = 1/12$): "*The inhibition threshold was lower to switch to VR*" – P11. In this context, participants positively highlighted that they no longer have to grab different devices ($n = 10/12$), allowing them to focus more on their task without interruptions ($n = 5/12$). They were thus more willing to verify the position of the cube on the 2D desktop and click through different perspectives: "*You're more likely to look [at the desktop] again than with the [baseline condition] because you can check it so easily*" – P08. As a result, one participant perceived our intentional study restriction of having to pick up each cube again after dragging it out of the desktop as a "*bug*", with four other participants expressing a similar impulse: "*intuitively [...] I lifted the [SpatialMouse] while dragging [the cube] out of the monitor*" – P03.

7.4.8 Discussion

We organize our findings based on our research questions.

RQ4.1: Performance. We designed our task to focus on frequent transitions between 2D and 3D interactions. Given that the *SpatialMouse* does not require a device switch, we expected significantly improved task completion times. Yet, despite saving about 5% time spent switching between devices, the *SpatialMouse* did not significantly outperform the baseline. Our results show that users spent proportionally more time in the *SpatialMouse*'s *mouse mode* than using the mouse, indicating that the *SpatialMouse* performed worse than a dedicated mouse for 2D input. This performance gap may stem from using lower-quality sensors or the prototype's form factor, which could be refined in future iterations. In contrast, spatial interactions were comparable in accuracy to those performed with a dedicated VR controller, suggesting that the *SpatialMouse* is well-suited for 3D input but still requires optimization for 2D tasks.

RQ4.2: Task Load. Our findings show consistently higher scores for the *SpatialMouse* in the NASA TLX, with significant improvements in effort, frustration, and overall score. While we observed no significant improvements in mental effort, we anticipate that more complex tasks, such as immersive analytics, will benefit substantially more from eliminating device switching. However, the NASA TLX questionnaire captures only a retrospective and subjective assessment of task load. To better understand users' mental demands during interaction, we recommend that future work incorporate real-time measures, such as built-in eye tracking available in off-the-shelf HWDs [220]. Participants reported feeling more focused on the task rather than being distracted by switching input devices. These findings suggest that the *SpatialMouse* effectively reduces cognitive and physical friction, thereby promoting flow.

RQ4.3: Usability and User Experience. The *SpatialMouse* significantly improves usability and user experience in the context of our abstract task. Given the frustration of switching between devices [40, 87, 315], we anticipate similar benefits in real-world scenarios (e.g., immersive analytics). However, leveraging this potential depends on ensuring long-term comfort and usability for mouse operation and spatial usage. Participants suggested improvements to the prototype's ergonomics, which we see as directions for future refinement. In addition, such seamless transitions must be equally supported by the interaction design and accompanying interaction techniques: without this alignment, users may experience friction or frustration, as was occasionally observed under the artificial constraints of this study setup.

7.5 Insights and Implications

Based on our prototype and study, we extract design insights (D 4.1–D 4.3) for further developments of the *SpatialMouse*, as well as research implications (I 4.1–I 4.3) for topics needing further investigation.

7.5.1 Physical Prototype

Our initial prototype was built with commodity components to explore the feasibility of our concepts and allow others to replicate and improve upon the device. We therefore see several opportunities for improvements.

Input Controls. Building the prototype with commodity components limited the overall performance of the device in terms of sensor accuracy, weight, and button tactility. While this performance gap can be partially addressed through industrial molding and proprietary sensors, such hybrid devices may also face inherent trade-offs: For example, mouse operation relies on discrete actuators, whereas VR controllers often employ analog triggers. Similarly, our design emulated the mouse wheel with a joystick, but many users may be reluctant to give up the familiar feel of traditional scrolling. While compromises are possible, specialized devices may still offer the best experience within their respective domains.

Form Factor. The *SpatialMouse* adopts a slanted form factor to provide a good grip that facilitates the transition between *mouse mode* and *spatial mode*. Despite the ergonomic benefits of a slanted design [251], some participants felt unfamiliar with this design, which may have impacted their performance. Likewise, we added a baseplate to support the hand in *mouse mode*, but this limited wrist articulation during *spatial mode*. Although we found the horizontal mouse shape unsuitable for spatial interaction, future iterations could balance the ergonomics of a horizontal mouse with the secure grip required for *spatial mode* **(D 4.1)**. Note that the optimal form factor is likely dependent on individual factors such as hand size: "*For me, I have a very large hand [...] and usually mice kind of don't really fit that well. But [the SpatialMouse] fits really nicely*" – P04.

Grip Change. Prior research [43, 44] and commercial products (e.g., Logitech MX Air[4]) require users to change their grip when transitioning between 2D and 3D operation. While this may offer ergonomic benefits, our findings suggest that such grip changes introduce friction and thus hinder a seamless transition between interaction modes: "*I actually found the design easy, because you could just put it down in a nice position without having to turn your hand in any strange way*" – P08. By removing the need for users to consciously switch modes – whether through grip adjustments, buttons, or other mechanisms – the *SpatialMouse* enables a seamless transition **(D 4.2)**. Future research should further explore the role of grip changes and weigh their ergonomic benefits against their potential to disrupt fluid interaction **(I 4.1)**.

7.5.2 Seamless Transition

Despite shortcomings in the initial physical prototype, enabling a seamless transition between 2D surfaces and 3D spaces significantly improved user experience and reduced perceived task load. We discuss how our concept improves upon *switching between devices* and is therefore *promoting fluidity*.

Switching Between Devices. Prior work has primarily investigated switches between 2D surfaces and 3D spaces through combinations of mouse and free-hand interaction [40, 87, 315]. Consistent with this work, we found that switching between a mouse and a VR controller introduces significant overhead and frustration: "*The most critical [difference] is the fact that I constantly have to let go and pick up another object [during baseline]. And that's really annoying. Because it basically immediately stops me doing the task and makes me focus on picking up the other device*" – P04.

Prior research concludes that "*hand gestures are the prevailing input modality for transitions with 3D objects*" [327]. We argue that this prevalence is in part due to the lack of a suitable unified input device. Our *SpatialMouse* represents a possible device to challenge this status quo: By eliminating a switch between devices entirely, the *SpatialMouse* addresses this limitation and provides seamless interaction between 2D surfaces and 3D spaces. However, further comparative analyses are necessary to unveil potential trade-offs between the *SpatialMouse* and established methods, such as switching between mouse and free-hand input **(I 4.2)**.

Promoting Fluidity. By removing the need for a physical switch between devices, the *SpatialMouse* significantly reduces the perceived task load and improves usability. Unsurprisingly, not having to manage device transitions consciously enabled participants to maintain greater focus: "*You don't interrupt your workflow by reaching over. You can concentrate more on the task at hand. I noticed, for example, that I could remember more things during the second trial [with the SpatialMouse]*" – P02. Overall, our findings indicate that the SpatialMouse aligns closely with the principles of fluid interaction [111], which emphasizes promoting flow, enabling direct manipulation, and minimizing gulfs of actions **(D 4.3)**.

Notably, by reducing friction – both mental and physical – the *SpatialMouse* encouraged participants to shift between interaction spaces naturally. For example, one participant attempted to interact with the 2D surface using a VR controller but experienced no such impulses with the *SpatialMouse* due to its effortless switching. In addition, several participants expressed frustration over intentional study restrictions that prevented fluid cross-dimensional interactions (e.g., cross-dimension drag and drop). Although these constraints were necessary to avoid confounding factors, participants' feedback highlights the potential of fluid interaction as supported by the *SpatialMouse*.

7.5.3 Application Scenarios

With the increasing viability of MR HWDs, we envision future workspaces replacing physical desktop monitors with virtual surfaces [152, 310]. For example, game development workflows could benefit from fluid transitions between precise 2D operations (e.g., accurately editing objects in isometric views) and spatial manipulations (e.g., positioning objects within a virtual scene). Such scenarios could reveal additional opportunities and challenges to guide the design of the *SpatialMouse* **(I 4.3)**.

The *SpatialMouse* can also be employed in general MR use cases that do not rely on frequent transitions. For example, MR environments enable the placement of monitors beyond physical restrictions [394]. Yet, mouse usage in such environments can be challenging, as users may need to switch between distant 2D surfaces or navigate large spaces. Here, our hybrid interaction techniques (see Section 7.6) could be beneficial and enable the integration of other modalities, such as gaze, to traverse and interact within such spaces.

Design Recommendations

- **D 4.1** Optimize form factor depending on primary usage space.
- **D 4.2** Minimize user's burden of transition, for example by avoiding grip changes.
- **D 4.3** Support fluid interaction by promoting a seamless switch between 2D surfaces and 3D spaces.

Research Implications

- **I 4.1** Investigate trade-offs on grip change on ergonomics and workload when switching between modes.
- **I 4.2** Compare *SpatialMouse* against prevailing input methods for transitions (e.g., combinations of mouse and free-hand interaction).
- **I 4.3** Study usage of *SpatialMouse* in holistic usage scenarios.

7.6 Hybrid Interaction Techniques

Based on the insights from our user study and our experience in developing the *SpatialMouse*, we describe four interaction techniques that use the available capabilities of transitional workspaces (see Figure 7.8). As our work extends well-established interaction techniques such as clutching, they need careful evaluation of their efficacy in real-world applications, which we consider outside the scope of this work. Additionally, we consider our described interaction techniques an initial foundation for further exploration and refinement of interaction techniques enabled by the *SpatialMouse*.

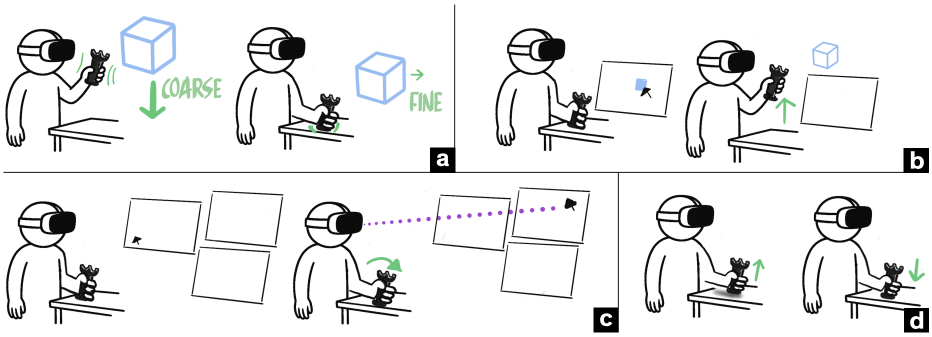


Figure 7.8: The *SpatialMouse* opens up opportunities for novel interaction techniques, such as (a) *hybrid input*, (b) *cross-dimension drag and drop*, (c) *gaze-based clutching*, or (d) *stamping*.

Hybrid Input. Prior work shows that 6DoF devices outperform mouse input for basic manipulation tasks in 3D [428]. Still, mouse input can be more accurate for placing objects in 3D [31], especially as our biomechanical constraints may cause translations to be accompanied by unintentional rotation during spatial operations [169]. Using the *SpatialMouse*, we can fluidly switch between fine-grained mouse interaction or efficient 6DoF input, depending on the required level of precision. For example, 3D manipulations during 3D modeling may be achieved by first coarsely positioning an object using the 6DoF in *spatial mode*, followed by fine-grained adjustments in *mouse mode* (see Figure 7.8 a).

This approach also extends to interactions on large or spatially distributed 2D surfaces. For example, moving an object between 2D surfaces may require multiple clutching operations to reach the target. In such a scenario, the user may briefly engage the *spatial mode* to position an item coarsely via mid-air pointing, then return to *mouse mode* for fine adjustment.

Cross-Dimension Drag and Drop. Immersive workspaces require frequent transitions between 2D and 3D, such as dragging an item from a 2D surface to its surrounding 3D environment [86, 327, 407]. For example, this study task required users to drag and drop a cube from a 2D surface into 3D using first the *mouse mode*, then picking up the cube again in 3D using *spatial mode* for positioning. While we deliberately split this into two distinct actions for comparability, the *SpatialMouse* enables a seamless transition between 2D and 3D spaces, and vice versa: Users can start dragging in *mouse mode*, then lift the device to enter *spatial mode* while holding the button to perform a *cross-dimension drag-and-drop* operation (see Figure 7.8 b). Users can also grab an object in *spatial mode*, then put the *SpatialMouse* down to enter *mouse mode*, thereby transforming the 3D object back to its 2D representation on a 2D surface. Unlike our study task, this procedure no longer requires items to be dragged outside of the bounds of the 2D surface first; instead, the *SpatialMouse* can be lifted as soon as the 2D item is dragged, regardless of its position to transform it to 3D and vice versa, thereby avoiding potential errors due to proximity with other 2D surfaces.

Gaze-Based Clutching. Because the *SpatialMouse* provides relative input during *mouse mode*, users must perform clutching operations by slightly lifting the device, resetting it to a neutral position, and then putting it down again. This works for conventional mice, as the sensor does not report any movement while in the air. Unlike conventional mice, however, the *SpatialMouse* is also tracked in space, allowing the detection and usage of clutching as an implicit form of input. With eye-tracking prevalent in new HWD models, the *SpatialMouse* could therefore combine this with clutching to position the cursor: Once the clutching ends or *mouse mode* is initiated (i.e., the *SpatialMouse* is put down), the cursor jumps to the user's current gaze location (cf. [227]). While this can help to position the cursor on a single 2D surface, we also see the potential in switching between multiple 2D surfaces (e.g., multiple visualization dashboards [394]), which may be arbitrarily arranged in the MR environment (see Figure 7.8 c).

Stamping. Similarly to clutching, the process of lifting the *SpatialMouse* in *mouse mode* can serve as an explicit input. Unlike clutching, this technique refers to a rapid vertical motion, akin to *stamping* a document: lifting the device briefly and returning it to the table without horizontal movement (see Figure 7.8 d). This can be mapped to deliberate commands such as "*undo*", providing a quick yet intentional interaction method.

Due to the higher physical effort involved compared to conventional inputs such as button clicks, *stamping* may be particularly suited for deliberate actions where accidental input should be avoided. However, implementation of this technique requires careful considerations of detection thresholds (e.g., time and movement) to avoid ambiguities with clutching detection. In addition, lifting or putting down the *SpatialMouse* may cause movement of the 2D cursor (cf. Heisenberg effect of spatial interaction), which could limit its practical application.

7.7 Limitations and Future Work

As our work represents an initial exploration of the *SpatialMouse* concept, it has several limitations that provide opportunities for future research.

While our user study demonstrated that our hybrid pointing device outperforms the combination of each device (i.e., mouse and VR controller) in terms of task load and user experience, we did not investigate how the *SpatialMouse* compares to each device individually. As we cannot claim that our design matches the decades of industry research into mouse ergonomics, the current prototype compromises on mouse operation (e.g., ergonomics) and spatial usage (e.g., weight, maneuverability). However, future studies could employ standardized tasks (e.g., Fitts's law evaluations for 2D [127, 262] and 3D [6, 24, 385]) to provide precise benchmarks and identify concrete design improvements for both mouse and spatial usage. Potential directions include shape-changing tangibles [426] that dynamically adapt their form to the current interaction mode or modular attachments. In addition, given the potential ergonomic challenges associated with extended spatial interactions [168, 307], future research should assess the long-term comfort and usability of the *SpatialMouse* in real-world scenarios.

Another limitation is that this study only involved the usage of the primary and secondary buttons, leaving other controls, such as the joystick, unused. Since the specific controls of commercial VR controllers vary between vendors, it may be necessary to tailor the type and layout of these controls based on the expected use case scenarios. In addition, our prototype lacks haptic feedback (i.e., vibrations), which is common in VR controllers. Integrating vibration motors to provide haptic feedback increases the weight of the *SpatialMouse*, but could prove useful for both *spatial mode* and *mouse mode* [303].

Lastly, our user study intentionally used an abstract task that emphasized switching between 2D surfaces and 3D spaces. While this allowed controlled comparisons, it may not fully reflect holistic, real-world usage contexts. However, ecologically valid usage also includes many confounding factors, such as pointer speed and hand sizes, which provide options for in-depth investigations in future work. In this context, our investigation only involved right-hand interactions. Similarly, we did not consider how off-hand usage factors into device transitions: For example, mouse usage on a desktop is often combined with a keyboard for text input; likewise, manipulations in 3D spaces may use bi-manual interaction using two controllers.

7.8 Chapter Conclusion

The "*SpatialMouse*" exemplar demonstrates a hybrid pointing device that combines the capabilities of a desktop mouse with a virtual reality controller, allowing for a seamless transition between 2D surfaces and 3D spaces in asynchronous hybrid user interfaces. The *SpatialMouse* is thus suitable for both indirect 2D pointing of a desktop mouse with 3D input with six degrees of freedom, allowing users to fluidly switch between both by simply lifting or putting down the device. By removing both the physical and cognitive barriers associated with device switching, the *SpatialMouse* directly addresses **RO1: Transitioning Between Devices**. In addition, removing the barriers for switching between indirect 2D pointing and direct 3D manipulations opens a range of novel interaction techniques, contributing to **RO3: Interaction Techniques**.

While this exemplar deviates from the others by focusing on a hybrid *input device* rather than a hybrid *user interface*, it highlights important directions for future research. First, STREAM (see Chapter 4) demonstrates that hybrid user interfaces can also treat 2D interaction components as spatial inputs, combining touch-based 2D interaction and 3D manipulation. Here, the insights gained from the *SpatialMouse* can inform the broader design of hybrid user interfaces. Second, as we approach Sutherland's ultimate display [377], the need to combine output devices may vanish. Yet even with the ultimate display, both 2D surfaces and 3D spaces will continue to play essential roles, requiring appropriate input devices, such as the *SpatialMouse*.

Part III

SYNTHESIS

Chapter 8

Reflections

Chapter Content

Parts of this chapter are based on the following publications:

Sebastian Hubenschmid*, Marc Satkowski*, Johannes Zagermann*, Julián Méndez*, Niklas Elmqvist, Steven Feiner, Tiare Feuchtner, Jens Emil Sloth Grønbæk, Benjamin Lee, Dieter Schmalstieg, Raimund Dachselt, and Harald Reiterer. "Hybrid User Interfaces: Past, Present, and Future of Complementary Cross-Device Interaction in Mixed Reality." Submitted to: *IEEE Transactions on Visualization and Computer Graphics*. 2025

Sebastian Hubenschmid*, Daniel Immanuel Fink*, Johannes Zagermann, Jonathan Wieland, Harald Reiterer, and Tiare Feuchtner. "Colibri: A Toolkit for Rapid Prototyping of Networking Across Realities." In: *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Sydney, Australia: IEEE, 2023, pp. 9–13. ISBN: 9798350328912. DOI: 10.1109/ISMAR-Adjunct60411.2023.00010

* Contributed equally

8.1 Chapter Context

The exemplars presented throughout Part II reveal design insights and research implications that are rooted in their specific context. This chapter aims to synthesize those context-dependent findings into generalized insights that can inform future work across the wider field of hybrid user interfaces. Towards this goal, this chapter will first revisit the overarching research objectives addressed by each exemplar, complementing the findings with insights from the literature survey (see Part I). Then, it extracts common design insights and research implications from the presented exemplars. Since hybrid user interfaces often focus on specific technological combinations, this chapter also explores several considerations regarding their implementation.

8.2 Research Objectives

The work in this thesis was motivated by the following overarching research goal:

> **Research Goal**
> How can hybrid user interfaces support fluid interaction in immersive analytics workflows?

This goal was further split into three research objectives to address the core challenges and opportunities of hybrid user interfaces. As each exemplar contributed towards these objectives within its specific context, this section reflects on the research objectives from a more generalized perspective, synthesizing findings across all case studies and incorporating results from our literature survey (see Chapter 3). Given the vast design space, note that these results are largely limited to the design dimensions covered by the previously presented exemplars (see Figure 8.1).

8.2.1 RO1: Transitioning Between Devices

> How can we design transitions between tightly coupled devices to support fluid interaction in hybrid user interfaces?

Hybrid user interfaces draw their strength from the combination of complementary interaction components, leveraging the distinct benefits of heterogeneous devices. While this can have many benefits, it also implies a transition between different devices. To support fluid interaction, we thus need to consider the overall *spatial layout*, manage *visual attention*, and offer techniques for *seamlessly transitioning* between devices.

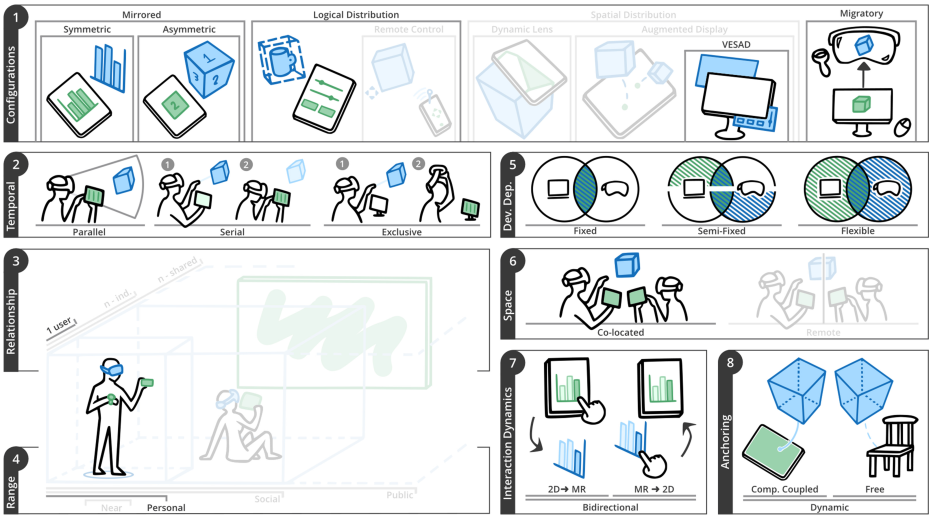


Figure 8.1: Overview of hybrid user interface characteristics explored throughout all exemplars. Uncovered characteristics are faded out.

Spatial layout. In multi-display environments, the spatial layout and relationships between displays significantly influence visual attention and system usability [325]. While traditional setups are bound by physical constraints, hybrid user interfaces offer greater flexibility (e.g., VESAD configurations), which also introduce new challenges. For example, prior work has shown that large distances between a smartphone and its corresponding AR content can introduce mental overhead [19]. In other cases, such as situated analytics, spatial placement may be dictated by proximity to physical referents to preserve contextual relevance. Thus, the design of hybrid user interfaces must carefully consider the appropriate spatial layout between 2D interaction component and MR environments.

An important factor that could inform the required spatial layout – and thus hybrid user interface configuration – is the *degree of complementarity* between the 2D interaction component and the MR content. Prior work [19] as well as the exemplars suggest that there may be a correlation between the spatial distance of content and the cognitive effort required for switching. Therefore, the switch must offer sufficient value to justify this cost.

In cases of high complementarity, greater spatial separation may be acceptable: The spatial separation in the "*STREAM*" and "*ReLive*" exemplar is high, but exhibits high complementarity by fully utilizing the benefits of their respective environment (e.g., precise interaction on a 2D interaction component, spatial capabilities of MR environment). Conversely, when the complementarity is low, the content should likely remain in close proximity to minimize cognitive disruption. For example, the "*ARound the Smartphone*" exemplar demonstrated the switching costs for small virtual display extensions can outweigh their benefits, even in a VESAD configuration.

Visual attention. Introducing multiple visual output components divides the user's attention among each component, increasing mental load [326]. Prior work on multi-display environments proposes several metrics to potentially reduce strain [325, 326] which are equally applicable to hybrid user interfaces. In addition to these metrics, current hardware limitations have to be considered for hybrid user interfaces, as several works [106, 150, 292, 407] opt for VST AR HWDs to avoid differences in focal planes, which can further increase mental load [106].

In this context, the "*STREAM*" exemplar demonstrates several promising strategies to reduce the amount of visual attention shifts. First, it leverages the user's proprioception by introducing an *eyes-free interaction* concept that maps different corners of the tablet to specific actions. This allows users to interact with MR content via the tablet without needing to look at it. Second, for actions that do require visual attention on the tablet (e.g., touch-based input for filtering data), "*STREAM*" ensures that users are already looking at the tablet, thereby minimizing unnecessary transitions. Thus, the design of hybrid user interfaces must carefully consider the user's visual attention during their workflow.

Seamless transition. Since switching between devices is an inherent aspect of hybrid user interfaces, minimizing this disruption is an essential part of achieving fluid interaction. While the design of such transitions depends on the configuration and context of each hybrid user interface, the previous exemplars provide potential directions. "*STREAM*" demonstrates a seamless transition through the *tablet lens* interaction, which visually merges the tablet and MR scatterplot, allowing users to fluidly transition between 2D and 3D environments. In contrast, "*ReLive*" facilitates the transition by offering a preview of the MR scene on the desktop.

However, the disruption caused by physically switching between devices (e.g., with *exclusive* temporal usage as in "*ReLive*") remains a significant barrier. Here, the "*SpatialMouse*" represents a promising step towards achieving fluid interaction in *migratory* configurations: By enabling a seamless transition between 2D and 3D interaction within a single device, it unlocks novel interaction techniques, supports flow, and is thus conducive to fluid interaction.

8.2.2 RO2: Task Allocation

> How can an immersive analytics workflow be effectively distributed across heterogeneous devices in a hybrid user interface?

Prior research on hybrid user interfaces often suggests allocating system control to the 2D component and leveraging the MR component for spatial interaction tasks [43]. While this is true in general, hybrid user interfaces must account for more nuanced trade-offs when distributing tasks across devices. Designers need to consider the *cost of switching*, leverage *complementary views*, account for *hardware capabilities*, facilitate *transferring tasks and content between interaction components*, and support the *perceptual synchronization* of content.

Cost of switching. The cost of switching between devices can influence how tasks and capabilities should be allocated. In *exclusive* temporal usage such as "*ReLive*" or "*ARTHUR*", high switching costs may necessitate that both components offer similar capabilities to minimize disruptive transitions. In contrast, when switching is seamless (e.g., in *parallel* or *serial* temporal usage such as "*STREAM*" or "*ARound the Smartphone*"), interaction design can intentionally guide users toward using the most appropriate device for a given task, thus leveraging the strengths of each without incurring excessive cognitive load.

Complementary views. Effective task allocation also depends on the degree of complementarity between components. The idea of complementary views is well-known in visualization research and is equally applicable to allocating tasks in hybrid user interfaces. For example, in immersive analytics, the 2D interaction component can serve as the workspace for quantitative analysis or scripting, while the MR interaction component provides spatial understanding. Likewise, in situated analytics, analysis benefits from the spatial placement and context of data within the user's environment, while the 2D interaction component can offer familiar interaction with mobile devices.

Hardware capabilities. Since hybrid user interfaces are a technology-driven concept, the specific capabilities and limitations of hardware are critical in guiding task allocation. For example, while stereoscopic views can enhance depth perception, they may also introduce perspective distortion. In such cases, monoscopic views offer simplified perception and interaction, as demonstrated by "*STREAM*".

Current trends in commercial MR hardware suggest two divergent trajectories: Optical see-through AR glasses, such as the XReal Air[1] are lightweight and well-suited for mobile use. Yet, they remain limited in terms of visual fidelity, brightness, and virtual field of view. In contrast, video see-through headsets, such as the Apple Vision Pro[2] offer a richer immersive experience and greater display fidelity, but this comes at the expense of portability and prolonged comfort.

Given these diverging capabilities, it can be beneficial to designate a *primary device* within the hybrid user interface. For instance, "*ARound the Smartphone*" and "*Push2AR*" anticipate the limitations of current mobile AR glasses and thus treat the smartphone as the primary device, with mobile AR glasses acting as a supplementary display space. In contrast, "*ReLive*" uses a workstation-based scenario with more capable HWDs and thus treats both the desktop and immersive components as equal.

Transferring tasks and content between interaction components. When distributing tasks and content across devices in a hybrid user interface, it is essential to consider how they can be transferred across devices. While the distribution is often static (e.g., determined by the designer of the system), several systems have demonstrated the utility of transferring content from a 2D interaction com-

[1] https://www.xreal.com/air2, last accessed 2025-05-12.
[2] apple.com/newsroom/2023/06/introducing-apple-vision-pro/, last accessed 2025-05-12.

ponent to the MR environment [4, 355, 400, 413]. Here, our "*Push2AR*" concept demonstrates the use of visual anchors to link MR content back to its original smartphone position.

In the context of immersive analytics, recent research investigated visualization transformations between 2D and 3D [234, 236]. Several transition metaphors have been explored in prior work (e.g., "*pulling*" sticky notes from a smartphone [413]), but further investigations are necessary to standardize and evaluate these techniques [355], especially as the boundaries between devices continue to blur.

Perceptual synchronization. Another challenge is found within the mapping and synchronization of content and interaction across components. While prior research has shown that an asymmetry of interaction (e.g., decoupling spatial interaction [357]) or information (e.g., showing simplified 2D views of 3D content [172, 191]) can be beneficial, this may come at an increased cost of user perception: For example, how does a 2D interaction affect its 3D equivalent, and how can we communicate this asymmetry?

Aside from these conceptual challenges in establishing a consistent mental model, we also have to consider technological challenges. In this context, prior work found that cross-device interaction techniques can be highly sensitive to network latency [248]. This sensitivity may be even more pronounced in hybrid user interfaces, especially if they appear as one conceptual device (e.g., the virtual screen in a *VESAD* configuration may lag behind the real screen). For example, in the "*ARound the Smartphone*" exemplar, network latency was minimized to ensure that the virtual extension remained synchronized with the smartphone display, resulting in a virtually imperceptible delay between devices – thus contributing to the illusion of a virtual screen extension. Still, more research is necessary to investigate how inevitable technological factors (e.g., latency) can confound findings.

8.2.3 RO3: Interaction Techniques

> What novel interaction techniques emerge from the use of complementary devices in hybrid user interfaces?

One of the main advantages of hybrid user interfaces is their ability to leverage established interaction paradigms from heterogeneous devices (e.g., touch gestures). As the exemplars in this thesis demonstrate, hybrid user interfaces can also serve as a foundation for entirely new interaction techniques. These techniques typically emerge by *eliminating the barriers* between 2D and 3D interaction spaces, leveraging *eyes-free interaction*, or *unchaining device capabilities*.

Eliminating Barriers. The "*STREAM*" and "*SpatialMouse*" exemplars illustrate how novel interaction techniques emerge when transitions between 2D and 3D interactions are seamless. By embedding 2D input into a 3D context (as in "*STREAM*" and the “*SpatialMouse*”) or employing 3D input within a 2D workspace (as in the “*SpatialMouse*”), hybrid user interfaces open new possibilities for supporting fluid interaction. However, the investigated interaction techniques merely scratch the surface of what is possible. For example, “*STREAM*” could be extended to support seamless transitions between touch input and mid-air gestures (e.g., seamlessly moving from a touch gesture to a mid-air gesture or vice versa), potentially unlocking new interaction techniques for fluid interaction.

Eyes-free interaction. One of the challenges in hybrid user interfaces is managing divided visual attention between multiple output components. Often, users can only pay attention to one device at a time, leaving the other device unused. The “*STREAM*” exemplar addressed this by introducing an eyes-free interaction technique, leveraging the user’s proprioception and tablet form factor. By mapping different actions to each corner of the tablet, users can benefit from the touch interaction (i.e., haptic feedback) without shifting their visual attention, thus preserving flow and reducing the cognitive load of switching between devices.

Unchaining device capabilities. The past decade has seen a surge in both device variety (e.g., smartwatches, smartrings) and device capabilities (e.g., inside-out spatial tracking), which can also be seen in the increase in interaction component combinations. Yet, 2D interaction components such as smartphones are restricted by their form factor, favoring device ergonomics and portability but limiting their potential output capabilities. By combining them with MR interaction components (e.g., AR HWDs), we can “*unchain*” their capabilities, enabling entirely new interaction possibilities (e.g., offloading menu items as demonstrated in “*Push2AR*” or using their spatial awareness as input as in “*STREAM*”), thereby inching closer to a form of “*universal interaction*” [218] and ultimately, Weiser’s vision of ubiquitous computing [405]. While this can further expand the design space, we also recognize the need to establish guidelines to better understand the trade-offs and potential of each device (cf. [398]).

8.3 Design Insights

Reflecting on the design and evaluation of the exemplars presented in Part II, this section describes generalized design insights applicable to the broader field of hybrid user interfaces.

8.3.1 Complementarity First

When creating a hybrid user interface, the underlying motivation should address specific limitations in the interaction or output capabilities. The value of hybrid user interfaces comes from their ability to combine heterogeneous devices in a way that offsets each device’s individual weaknesses, drawing upon their individual strengths. However, if the devices are not already part of the user’s existing device ecology, a dedicated solution (e.g., bespoke tangibles) may be more appropriate.

The general principle of *complementary interfaces* highlights that devices in a multi-device ecology should not only simply coexist, but meaningfully complement each other. This is reflected throughout all exemplars: Each device or environment was assigned a complementary role within the user's ecology, fulfilling a role that was not suitably covered before. Thus, the goal of a hybrid user interface should not be in a random assortment of technologies that communicate with each other (cf. cross-device interaction), but in a principled integration of heterogeneous devices to form a unified system that is greater than the sum of its parts.

8.3.2 Seamless Transitions

A recurring theme across all exemplars was the seamless transitions between devices by using a conscientious interaction design that guides the user's visual attention, minimizes disruptions, and thus encourages flow. The qualitative feedback across the studies reveals some initial insights into the importance of minimizing friction during transitions. This aligns with recent research advocating for animated and smooth transitions between 2D and 3D views [244, 355]. In particular, the "*SpatialMouse*" shows that removing barriers in transitions can lead to novel interaction techniques and support the fluid interaction in the user's workflow.

However, achieving this fluidity requires not only appropriate interaction design, but also technical considerations. For example, an *exclusive* temporal usage due to device limitations (e.g., as demonstrated by "*ReLive*") prevents a truly seamless transition due to hardware limitations (i.e., putting HWD on). In addition, technical factors such as latency and synchronization must be addressed. Here, prior work has shown that latency in cross-device scenarios can quickly degrade user experience, which is likely equally applicable to hybrid user interfaces.

8.3.3 Linking Content

Given that hybrid user interfaces act as a singular application, they can be seen as different lenses on a shared interactive space (e.g., multiple coordinated views). It is therefore essential to establish clear links between content distributed across different interaction components, allowing users to maintain context when switching between potentially different 2D and 3D representations, which is common in immersive analytics. For example, the "*STREAM*" exemplar implemented this by using colors on the tablet interface that match the selected AR scatterplot, reinforcing their connection through a shared color.

8.3.4 Ever-Shifting Window of Opportunity

With their technological focus, the design of a hybrid user interface is dependent on an evolving technological landscape with an ever-shifting window of opportunity. As such, their effectiveness is determined by the affordances of available hardware: Earlier exemplars, such as "*STREAM*", emphasized interaction with 2D interaction components due to the technical limitations of mid-air interaction at that time.

Yet, as MR hardware continues to mature, the inclusion of more capable spatial inputs becomes viable and provides new opportunities.

However, this means that the design of a hybrid user interface must remain adaptable and anticipate upcoming hardware capabilities. For example, foldable smartphones could replace the need for virtual screen extensions, offering a large screen space without the overhead associated with switching between a virtual and physical display. In contrast, the growing prevalence of smartwatches might shift the utility of VESAD configurations towards combinations of lightweight wearables and AR glasses.

8.4 Research Implications

Drawing from the findings of the presented exemplars and the insights gained from our literature analysis, this section outlines several broader research implications for the design and evaluation of hybrid user interfaces.

8.4.1 (When) Are Hybrid User Interfaces Beneficial?

The "*STREAM*" and "*ReLive*" exemplars highlight the potential of hybrid user interfaces in immersive analytics scenarios. For their evaluation, we intentionally chose qualitative assessment methods to explore their usability and practical applicability. Still, a notable limitation is the absence of direct comparisons against established baselines.

Evaluating hybrid user interfaces against state-of-the-art alternatives presents significant challenges, as real-world scenarios often involve a myriad of confounding variables that make it difficult to isolate the specific benefits of hybrid user interfaces. Decomposing these systems into their individual components (e.g., such as comparing input modalities in isolation) oversimplifies the evaluation and ignores the symbiosis of interfaces that emerge from the combination of heterogeneous devices. As a result, it remains difficult to determine when and where hybrid user interfaces offer a quantitatively measureable benefit.

8.4.2 Exploring Holistic Real Life Applicability

The versatility of hybrid user interfaces allows for their application across a wide range of domains. While this thesis focuses specifically on immersive analytics, the increase in artifacts and evaluations within the survey corpus indicates that the nascent space of hybrid user interfaces is slowly maturing. Research prototypes and their interaction techniques have so far been investigated in isolation, yet "*it is clear that [these] facets should not be discussed in isolation; instead, they are highly interconnected and affect each other*" [226]. Given the wide applicability of existing and future systems (e.g., robotics [254], medical domain [226, 331, 366], explainable artificial intelligence [273]), we need to consider their role within the holistic context of their work environments and unveil their unique challenges and opportunities. Furthermore, consolidating common interaction techniques into a library of design patterns can help researchers and practitioners alike in the design and implementation of holistic hybrid user interfaces.

8.4.3 Collaboration

The exemplars in this thesis as well as the majority of the papers in the survey corpus focus on *single-user* systems. Yet, some papers indicated a collaborative setting, either with a shared interaction component (e.g., a shared output) or using individual interaction components.

Regarding co-located collaboration, the choice whether there is a shared interaction component or individual interaction components can directly influence the type of collaboration: For example, Butscher et al. [66] used a large interactive tabletop display as a shared input device, combined with an AR HWD for each user. Limited by the touch input of the shared device, only one user is able to manipulate the content, thus enforcing a closely coupled collaboration [157, 384]. In contrast, providing users with individual interaction components enables loosely coupled, parallel activities [354]. Similarly, asymmetric device setups might benefit from individual device affordances [131] but also lead to an asymmetry of roles [421].

Lastly, no paper in our corpus features remote collaboration. In such cases, techniques known from MR remote collaboration could be applied: Virtual avatars can represent the remote person to create awareness [148, 149] while methods to align workspaces can allow for deictic referencing [125]. The use of diverse interaction components might require user representations per interaction component, and highly dynamic environments might benefit from transitional user representations [421].

8.4.4 Evaluation, Assessments, and Toolkits

Brudy et al. [56] described that research on cross-device interaction can be considered as a "*constructive problem*" [299]. This is reflected in our corpus: The majority of papers focus on creating new artifacts. These papers "*re-envision and push the boundaries of interaction possibilities*" [56] – a common theme for research in HCI that is often "*much better at proposing new technologies than at validating them*" [175], especially when novelty is expected as a key contribution[3]. This further fragments research on hybrid user interfaces, as artifact and empirical contributions tend to drift apart [56]. While some research addressed the call of Brudy et al. [56] for a frame of reference to compare cross-device interaction techniques [422], this is still missing for research on hybrid user interfaces. One possibility to create such frames of reference and further systematically study (and later compare) research within this space is to consider *experiments as design-informing activities* [298]. Here, different design alternatives (e.g., meaningful combinations of input and output components) can serve as independent variables

[3] sigbed.org/2022/08/22/the-toxic-culture-of-rejection/, last accessed 2025-05-12.

for different use cases. This allows us to study their effect on general (e.g., time and error) and use case-specific dependent variables, such as utilization of devices [353] – focusing on the effect, influence, and utility of each combination, beyond a technical perspective [56].

Hybrid user interfaces differ from single-device user interfaces by involving multiple interaction components, making assessments more complex. For example, cognitive workload (a typical metric in HCI user studies [220]) can be measured post-hoc using subjective questionnaires like NASA TLX [163, 164], but real-time, objective methods such as eye tracking are preferable [220]. As hybrid user interfaces often include an HWD, continuous assessment of eye movements via built-in sensors is achievable. However, switching between interaction components complicates data collection, as some components might not support eye tracking or may have varying data quality, requiring data fusion or repeated calibration. This issue becomes more complex with multiple users frequently switching between interaction components. Although subjective questionnaires simplify user studies, they reduce data accuracy, as participants might forget specific details. Objective, real-time assessments are critical for understanding hybrid user interfaces, but careful study designs are necessary to manage complexity in conducting and analyzing them.

To address this complexity, research has suggested a variety of toolkits that can support this process for MR user studies [60, 186, 289, 332]. For example, the "*ReLive*" exemplar describes one solution to facilitate such data logging for MR user studies. Yet, these toolkits also hold the potential to support the design, conduct, and analysis procedures of user studies involving hybrid user interfaces.

8.5 Implementation Considerations

As hybrid user interfaces are a technology-driven realization of complementary interfaces, their design and development demand careful attention in various implementation aspects. This section outlines key technical considerations and challenges for creating hybrid user interfaces.

8.5.1 Authoring Mixed Reality Hybrid User Interfaces

Creating hybrid user interfaces can be challenging: On the one hand, designers require evidence-based guidelines that focus on the integration of multiple devices [94], such as determining the optimal ergonomic content distribution [117, 425]. On the other hand, developers need to deal with implementations on different platforms. Although web technologies (e.g., WebXR) can standardize development, their support on commercial HWDs is uncertain [63]. In addition, toolkits (e.g., for integration of multiple devices [142], interaction component configuration [343], and synchronization [184]) can help focus on the implementation at hand, rather than worrying about implementation details.

8.5.2 Visualization Synchronization

Another challenge in applying hybrid user interfaces to immersive analytics is in the creation and synchronization of visualizations. While the creation of effective visualizations can already be challenging in itself, hybrid user interfaces potentially double the required effort: visualizations must often be implemented separately for both 2D and 3D environments, which may differ significantly in representation and interaction capabilities.

Here, a common grammar (cf. VEGA [348, 349]) can be useful, allowing content to *responsively* adapt to each device automatically. For example, the open-source[4] RagRug toolkit [128] introduced a workflow for responsive, reactive visualizations and has already been used to replicate the "*STREAM*" exemplar. Still, more research is necessary to build a toolkit that facilitates the authoring of such hybrid visualizations and their transitions (cf. [234, 236]).

8.5.3 Networking

Reliable and low-latency networking is a critical requirement for hybrid user interfaces, as they need to interconnect multiple devices in real time to enable a seamless transition. But despite the proliferation of toolkits in different areas such as visualization [88, 128, 335, 361] or logging [186, 289], networking has been mostly neglected and delegated to commercial solutions. In contrast to commercial applications, research prototypes have distinct requirements regarding the empirical reproducibility, data availability, latency, and privacy – ruling out externally hosted server software while simplifying development by neglecting edge cases (e.g., anti-cheat precautions) or artificial restrictions (e.g., reducing update rate to save on bandwidth).

To address the distinct needs of hybrid user interfaces, we created Colibri (**co**mmunication **libr**ary), an open source[5] networking toolkit for data exchange, model synchronization, and voice transmission. Colibri is a Unity- and web-focused networking toolkit for researchers to quickly prototype hybrid user interfaces, prioritizing low latency to suit the specific needs of hybrid user interfaces. Colibri was iteratively developed and refined over the course of all exemplars presented in this thesis and thus represents the technical foundation of all presented exemplars.

8.5.4 Future Challenges

The development of hybrid user interfaces spans many more areas that require further support. Since sophisticated toolkits can drastically reduce the technological barriers, the following section describes three areas that are not yet adequately supported by existing toolkits.

Alignment of Coordinate Systems. A common task when combining multiple interaction components is the alignment of coordinate systems. Nowadays, many

[4] https://github.com/philfleck/ragrug, last accessed 2025-05-12.
[5] https://github.com/hcigroupkonstanz/Colibri, last accessed 2025-05-12.

devices have powerful integrated tracking capabilities, making external tracking solutions (e.g., Optitrack) superfluous. However, this also results in each device establishing their own coordinate system, which then needs to be aligned with all others. These solutions are usually specific to the hardware configurations of each project and involve a combination of image markers, external tracking systems, or specific workarounds (e.g., starting the application in the same room position). A unified solution that automatically provides coordinate system alignment could greatly reduce the obstacles of combining multiple devices.

Asymmetric Device Capabilities. Hybrid user interfaces span across multiple heterogeneous devices, such as desktop computers, handheld devices (e.g., smartphones or tablets for AR), or head-worn MR devices. Each device in this ecology has different capabilities or limitations and, at times, a distinct role to fulfill. Here, automatically detecting, communicating, and assigning roles based on capabilities could be beneficial. For example, in asymmetric collaborative environments for immersive analytics, a desktop device can be useful for displaying visualizations in 2D, while other MR devices could be better suited for 3D visualizations. Here, we could specify data transformations (e.g., [234, 236]) to automatically convert the data, based on the specific capabilities or roles of the display device.

Merge Policies. In many collaborative scenarios, multiple collaborators may manipulate an object simultaneously, such as two users moving or rotating an object at the same time (see Figure 8.2). In such cases, a merge policy (e.g., averaging or summing up inputs) is necessary to properly incorporate the manipulations of both users (cf. [408]). However, this could be further complicated in hybrid user interfaces, as the object representations between the 2D and 3D environment may not be the same and thus require further transformations and merging strategies.

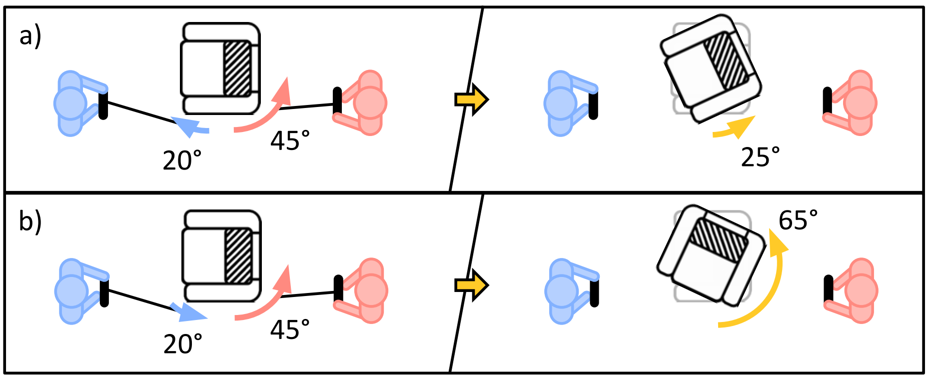


Figure 8.2: A composition merge policy allows two users to simultaneously manipulate an object, for example by summing up users' rotation inputs [408].

8.6 Chapter Conclusion

This chapter provides broader reflections on the context-specific exemplars introduced in Part II. First, it revisits the overarching research objectives and discusses them within a wider context beyond the scope of individual prototypes, integrating the findings from the literature survey presented in Part I. Then, by synthesizing the findings from the exemplars with the theoretical foundations and literature survey, the chapter derives general design principles, outlines research implications, and presents implementation opportunities. Together, these insights aim to inform the future design, development, and evaluation of hybrid user interfaces.

Chapter 9

Conclusion

Chapter Content

Parts of this chapter are based on the following publications:

Sebastian Hubenschmid*, Marc Satkowski*, Johannes Zagermann*, Julián Méndez*, Niklas Elmqvist, Steven Feiner, Tiare Feuchtner, Jens Emil Sloth Grønbæk, Benjamin Lee, Dieter Schmalstieg, Raimund Dachselt, and Harald Reiterer. "Hybrid User Interfaces: Past, Present, and Future of Complementary Cross-Device Interaction in Mixed Reality." Submitted to: *IEEE Transactions on Visualization and Computer Graphics*. 2025

* Contributed equally

9.1 Summary

Mixed reality hardware is gradually pervading everyday life, bringing us closer to Weiser's vision of ubiquitous computing [405] and Sutherland's concept of the ultimate display [377]. Although recent advances have improved visual fidelity and comfort of MR HWDs, spatial interaction remains an inherently limiting factor. Yet, spatial interaction is used as a form of "*universal interaction*" despite failing to accommodate the wide range of user needs in immersive environments. One proposition, which is the focus of this thesis, is to complement MR HWDs with conventional devices, such as tablets and desktop computers, and thus create a hybrid user interface.

Hybrid user interfaces are particularly well-suited for immersive analytics, where the benefits of immersive 3D visualization must be balanced with the precision and familiarity of 2D interaction. In this context, the concept of fluid interaction, which emphasizes user flow, direct manipulation, and minimizing gulfs of action, provides a suitable lens for evaluating hybrid user interfaces. Thus, this thesis is driven by the overarching research goal of: *How can hybrid user interfaces support fluid interaction in immersive analytics workflows?* To address this question, the thesis investigates three core research objectives:

- **RO1:** Designing seamless **transitions between devices**.
- **RO2:** Investigating meaningful **task allocations**.
- **RO3:** Exploring novel **interaction techniques**.

The thesis begins by establishing foundational knowledge in **Part I: Theory**.

Chapter 2 positions hybrid user interfaces within its related fields, thereby introducing the foundation required for a deeper exploration of hybrid user interfaces. This chapter first outlines the key concepts for multi-device interaction of *distributed user interfaces* and *cross-device interaction*, which encompass both homogeneous and heterogeneous device setups. The chapter then discusses *cross reality environments* and highlights the relevance of transitional interfaces. While these areas describe *how* devices and realities can be combined, it is unclear *why* such cross-device and cross reality approaches should be used. Therefore, this chapter introduces *complementary interfaces* as a conceptual lens to describe meaningful combinations. The chapter concludes by framing immersive analytics and fluid interaction as core use cases.

Next, **Chapter 3** explores the concept of hybrid user interfaces in depth. We first examine the history of hybrid user interfaces, based on which we establish a general definition and present general characteristics. We then conduct an extensive literature survey, selecting a corpus of 83 papers that match our criteria of a hybrid user interface. Based on this, we establish a taxonomy with eight dimensions to describe existing hybrid user interfaces and inform future systems. This taxonomy is used as an overarching framework for contextualizing the exemplars in Part II, providing a common foundation through which the exemplars can be connected.

Building on the theoretical foundation, **Part II: Exemplars** presents four implemented hybrid user interfaces, each addressing different aspects of the research objectives through empirical evaluation.

Chapter 4 introduces "*STREAM*", a hybrid user interface combining spatially-aware tablets with AR HWDs to enables fluid interaction with 3D parallel coordinates visualizations. This exemplar focuses on multimodal interaction by leveraging available input methods to interact with the visualization's components, such as the tablet's display for direct touch input, spatial awareness of the tablet for spatial input, and the HWD for head-gaze, voice input, and proxemics. STREAM demonstrates a novel eyes-free interaction technique that allows users to perform actions using their proprioception, which minimizes visual attention switches between devices. STREAM bridges the 2D and 3D representation of the visualization and contributes to all three research objectives (*RO1*, *RO2*, *RO3*).

In **Chapter 5**, the "*ARound the Smartphone*" exemplar explores the extension of a smartphone's screen with a virtual display through an AR HWD. This work focuses specifically on device transitions *(RO1)* by evaluating how varying sizes of virtual screen extensions impact spatial memory, workload, and user experience.

Chapter 6 presents "*ReLive*", an asynchronous hybrid user interface that combines a non-immersive desktop system with an immersive VR environment to analyze data from MR studies. This system investigates task allocation *(RO2)* by distributing immersive, in-situ exploration to the VR environment and aggregated, ex-situ analysis to the desktop. ReLive supports transitions between environments *(RO1)* by synchronizing both environments in real-time and offering a preview of the VR environment on the desktop and vice versa.

In **Chapter 7**, we present the "*SpatialMouse*", a hybrid pointing device that combines a desktop mouse with a VR controller. The *SpatialMouse* thus supports both 2D indirect input and 6DoF spatial input, allowing users to switch by simply lifting or resting the device. While not a hybrid user interface *per se*, the *SpatialMouse* serves as an enabling technology device transitions *(RO1)* in hybrid user interfaces, eliminating physical and cognitive friction associated with switching input devices. We outline a set of novel interaction techniques *(RO3)* enabled by this hybrid input device. Overall, our work addresses interaction barriers between 2D and 3D spaces, contributing a promising design concept toward achieving fluid transitions in mixed reality environments.

While each exemplar explores a specific use case in depth, **Part III: Synthesis** steps back to generalize these findings for the broader design space.

Chapter 8 synthesizes the insights from all four exemplars and incorporates the findings from the literature survey. This chapter first revisits each research objective and provides a general discussion of transitioning between devices *(RO1)*, task allocation *(RO2)*, and potential interaction techniques *(RO3)*. It then distills broader design principles, research implications, and practical implementation considerations, based on findings from the exemplars and drawing from the theoretical foundations discussed in Part I.

Finally, **Chapter 9** (i.e., this chapter) concludes the thesis by reflecting on the role of fluid interaction in hybrid user interfaces, questioning whether hybrid interfaces are a compromise or an opportunity, and contemplating on the future of hybrid user interfaces.

9.2 Overview of Contributions

This thesis follows the contribution types proposed by Wobbrock and Kientz [410], differentiating between *survey contributions*, *theoretical contributions*, *artifact contributions*, and *empirical research contributions*.

9.2.1 Theoretical Contribution: Complementary Interfaces

Our definition of *complementary interfaces* (Chapter 2) represents a *theoretical contribution*, derived from reflecting on our own prior work in cross-device interaction and cross reality environments. This concept underpins the work of this thesis by providing a conceptual lens through which to frame meaningful multi-device ecologies. While hybrid user interfaces are a possible technical realization of this concept and the primary focus of this thesis, the scope of complementary interfaces extends far beyond hybrid user interfaces. It offers a generalizable framework for understanding how different devices and environments can be meaningfully combined, and its theoretical nature allows the concept to extend beyond the potential lifetime of hybrid user interfaces (see Section 9.5).

9.2.2 Survey Contribution: Contextualizing Hybrid User Interfaces

Although the term hybrid user interfaces was coined in 1991 [122], the surrounding technological landscape has evolved considerably since then, while the definition itself has remained largely stagnant. Our *survey contribution* (Chapter 3) thus revisits hybrid user interfaces within the context of the current research landscape. We conducted a literature review of 579 records, selecting 83 papers that align with the characteristics of hybrid user interfaces. This survey highlights how hybrid user interfaces differ from adjacent research areas, identifies core characteristics, and introduces a taxonomy to establish a shared understanding of their key dimensions, design potential, and common challenges.

9.2.3 Artifact and Empirical Research Contributions: Exemplars

To explore hybrid user interfaces in depth, this thesis presents four hybrid user interface exemplars. Each exemplar (Chapters 4 to 7) provides a unique perspective on different facets of hybrid user interfaces. Each exemplar consists of two contribution types:

First, we provide an *artifact contribution* by describing the design and implementation of different hybrid user interfaces (or hybrid input devices). Each artifact explores different design ideas, hybrid user interface configurations, temporal usages, device dependencies, or interaction dynamics, thereby contributing towards the overarching research goal. All artifacts are available as open-source projects[1], allowing for replication and extension of our work. In addition, "*STREAM*" and the "*SpatialMouse*" contribute novel interaction techniques for hybrid user interfaces and hybrid input devices, respectively. Furthermore, "*ReLive*" contributes a technical data logging framework used in subsequent prototypes (i.e., "*ARound the Smartphone*", "*SpatialMouse*").

Second, each artifact is evaluated through an empirical evaluation, resulting in *empirical research contributions*. We employed different evaluation approaches depending on the context:

- "*STREAM*" was evaluated in an exploratory user study with 8 participants to examine how users interact with an immersive analytics visualization using a hybrid user interface. The study highlights the value of familiar touch-based interactions in 3D contexts and indicates the overall potential for fluid interaction in hybrid user interfaces.
- "*ARound the Smartphone*" employed a controlled laboratory experiment with 24 participants in a within-subject design to compare a baseline condition (smartphone with no virtual screen extension) with three commonly found display sizes (virtual screen extension of tablet size, desktop monitor size, and television size). The findings contribute to a better understanding of visual attention switches in hybrid user interfaces, providing foundational insights for designing a fluid transition between devices.
- "*ReLive*" used a two-step process. First, a guided design walkthrough allowed us to analytically investigate ReLive in a formative evaluation, contributing to the overall design of ReLive. Second, a user study with 5 experts provides deeper insights in an empirical evaluation. Our findings support the value of allocating in-situ and ex-situ tasks to different devices in a hybrid user interface, enabling a fluid workflow.
- We evaluated the "*SpatialMouse*" in a controlled laboratory experiment with 12 participants in a within-subject design to compare the *SpatialMouse* against a baseline of a mouse and VR controller. Our results indicate that our initial prototype significantly improves perceived workload and usability in tasks that require frequent switching between 2D and 3D input. The results emphasize that seamless transitions between devices are critical for achieving fluid interaction in hybrid user interfaces.

The results of each empirical evaluation were distilled into design insights and research implications, informing the overarching research goal. These insights are further generalized in Chapter 8 and inform the concluding discussions and future directions outlined in Chapter 9.

[1] See https://github.com/hcigroupkonstanz/ for an overview, last accessed 2025-05-12.

9.3 Fluid Interaction and Hybrid User Interfaces

The overarching goal of this thesis is to explore how hybrid user interfaces can enable fluid interaction within immersive analytics workflows. Although fluid interaction is an "*elusive and intangible concept*" [111] and therefore difficult to evaluate directly, the presented hybrid user interface exemplars reveal several design characteristics that align with the principles of fluid interaction [111].

- **Promoting flow.** Immersive analytics can improve flow through increased immersion. Yet, their sole reliance on spatial interaction can fall short for tasks that require, for example, precise selections or efficient text input. Hybrid user interfaces *promote flow* by allowing users to shift between immersive and conventional environments, using the device that best suits the user's current task. For example, spatial capabilities support in-situ exploration for environmental context, while 2D interaction components enable system control and analysis. This interplay between 2D and 3D "*[ensures] that interaction never 'ends'*" [111] or slows down due to inappropriate input and output modalities, thereby keeping users engaged.
- **Supporting direct manipulation.** While direct manipulation is central to fluid interaction and immersive analytics, its utility can be limited by spatial interaction in terms of ergonomics and accuracy. Hybrid user interfaces overcome this by *supporting direct manipulation* through appropriate modalities (e.g., combining touch for 2D tasks with mid-air gestures for 3D manipulation). This gives users the freedom to employ different kinds of direct manipulation depending on their specific needs.
- **Minimizing the gulfs of action.** Hybrid user interfaces can minimize the *gulf of execution* by offering familiar and appropriate input methods. However, hybrid user interfaces can also increase the *gulf of evaluation*: Feedback must not only be immediately clear on the current device, but also synchronized across all connected devices. Ensuring that changes are reflected instantly on all interaction components is thus critical.

While hybrid user interfaces can be helpful for enabling fluid interaction within immersive analytics, the principles of fluid interaction are equally essential for the design of hybrid user interfaces. Relevant guidelines for fluid interaction [111] can thus be reinterpreted through the lens of hybrid user interfaces:

- **Minimize indirection in the interface(s).** As hybrid user interfaces often distribute content across multiple devices, it is essential to clearly indicate the relationships between content across devices. This can be achieved through explicit visual links and real-time synchronization to reduce cognitive load when switching attention between devices.
- **Integrate user interface components in the visual representation.** Given the necessity of 2D interface components, these user interface components

should be easily available across the hybrid user interface. Depending on the gulf between devices (e.g., *parallel* or *exclusive* temporal usage), a hybrid user interface must either support this through seamless transitions (for *parallel* use) or ensure redundant access to features across devices (for *exclusive* use).

- **Reinforce a clear conceptual model.** Hybrid user interfaces often combine immersive environments with conventional 2D devices. This presents an opportunity to build on users' existing mental models of 2D devices to avoid legacy bias and leverage reality-based idioms [65]. For example, spatial input lends itself naturally for 3D manipulations, while tasks such as precise selections or text input can be reserved for 2D interaction components. A clear, consistent division of tasks and roles between devices (e.g., in-situ and ex-situ analysis, 3D manipulations and 2D system control) helps reinforce user understanding.
- **Avoid explicit mode (or device) change.** Switching between modes – or in this case, devices – introduces friction and disrupts user flow. Hybrid user interfaces should support seamless, voluntary transitions, where interaction continues naturally regardless of which device is in use. Hybrid user interfaces should thus avoid enforcing explicit device transitions and instead leverage novel interaction techniques, preview mechanisms, and eyes-free interaction.

Overall, fluid interaction for immersive analytics and hybrid user interfaces are closely intertwined: Fluid interaction emerged from overcoming limitations of early touch hardware through clever combinations of technologies. Likewise, hybrid user interfaces were created to overcome the limitations of early MR HWDs through the complementary use of device technologies. These similarities suggest that one cannot be fully realized without the other for the case of immersive analytics.

9.4 Hybrid User Interfaces – A Compromise?

Hybrid user interfaces are an intentional combination of off-the-shelf devices that would otherwise be used on their own. This can be seen as one of their core strengths: Rather than requiring bespoke hardware tailored to a narrow range of tasks (e.g., tangibles), hybrid user interfaces leverage everyday devices, with each device being capable of solving the task independently – even if, at times, poorly.

However, this can also be seen as a compromise: Since off-the-shelf devices are designed for general-purpose use, they inevitably involve trade-offs and thus may not perfectly align with the user's requirements. For example, the screen of a smartphone may remain mostly unused when combined with an MR HWD, as users focus most of their visual attention on the immersive content and eyes-free interaction is employed. Would users not be better served by specialized technologies instead of a patchwork of incoherent device technologies?

Indeed, specialized devices can offer better performance for their intended use case. However, this comes at the cost of their versatility and are thus less likely to be used outside their designed environment. Looking at the "*SpatialMouse*", for example, specialized inputs (e.g., mouse and VR controller) are optimized for their respective environment and likely outperform the *SpatialMouse* in terms of ergonomics and individual performance, but fall flat when used outside of their specific context (e.g., using a VR controller for precise 2D input). In contrast, hybrid user interfaces take advantage of concepts and devices that are already within the user's ecology, offering flexibility across a wide range of tasks. Even outside of the context of hybrid user interfaces, these devices remain valuable, as smartphones and MR HWDs, for example, can still be used on their own for their originally intended use case.

Rather than seeing hybrid user interfaces as a compromise, I instead argue that they represent a pragmatic design solution that leverages the heterogeneity of currently available device technologies. They allow designers to utilize the complementary strengths of existing device ecologies and create a *symbiosis of interfaces* that is greater than the sum of its parts. As long as hardware constraints persist, hybrid user interfaces will remain a practical solution to overcome these technological limitations.

9.5 The Future of Hybrid User Interfaces

At the beginning of this thesis, hybrid user interfaces were defined as "*an area of cross-device computing that leverages distinct benefits of heterogeneous interaction components*". As discussed in Chapter 3, this definition was intentionally broad, informed by prior usage of the term "*hybrid user interface*". Over the course of this thesis, the prior research landscape of hybrid user interfaces was surveyed (Part I) and a set of current exemplars was explored (Part II). Based on this, we can now discuss the meaning of the term and explore the future of hybrid user interfaces.

What is a hybrid user interface? The term "*hybrid*" in hybrid user interface introduces ambiguity, as it could describe any combination of multiple interfaces and concepts. This ambiguity and divergent facilitation of the term is further demonstrated by contrasting the original definition of hybrid user interfaces – referring to "*heterogeneous display and interaction*" [122] or "*different interface*" [119] technologies – with its most prevalent use today that was also used throughout this thesis (i.e., combining 2D and MR components). This led to the decision to not sharply define hybrid user interfaces because of the (1) historical developments and usages of the term, its (2) relations and overlaps to adjacent research areas, (3) the abundance of possibilities for combinations as shown in our taxonomy, and (4) potential edge cases. However, we opted to modernize Feiner and Shamash's definition of hybrid user interfaces [122] while specifically focusing on the initial device combinations of 2D and MR interaction components.

Overall, the term hybrid user interface is more akin to a **fuzzy concept**: "*a collection of objects [that do] not have sharp, clear-cut boundaries*" [29]. Thereby, hybrid user interfaces can be characterized with attributes as described in Chapter 3, defining "*the membership [as] a matter of degree*" [29] compared to an arbitrary fixed inclusion or exclusion based on currently available device technologies. In the end, actual manifestations may vary – depending on an ever-shifting window of opportunities – where hybrid user interface boundaries remain adaptable to emerging interaction paradigms, technological affordances, and contextual demands. However, if the term is adaptable in such a way, we must ask:

Will there be hybrid user interfaces in the future? On the bright side, hybrid user interfaces combine complementary device technologies, taking advantage of the strengths of each technology, creating user interfaces where "*the real power [...] comes not from any one of these devices; it emerges from the interaction of all of them*" [405]. This combination can create flexible and modular systems in which the user can freely decide when to employ which type of feature.

However, from a rather critical perspective, hybrid user interfaces could be considered as a compensation mechanism for the deficiencies of individual interaction components, such as the limited precision of gesture recognition in AR environments. By including additional technologies to address the shortcomings of any one input method, these systems can become too complex to design, evaluate, and eventually use. With advances in technology, the need for hybrid user interfaces composed of multiple interaction components may diminish, as an ultimate interface might not require the combination of multiple interaction components to compensate for individual weaknesses. Instead, at some point in the future, there will be self-contained "*devices*" that streamline interaction into a single cohesive system. This leads to the question:

Is there an expiration date for hybrid user interfaces? The reality is more nuanced than a straight yes or no. Looking back at over 30 years of research on hybrid user interfaces and related research streams has shown a maturing of concepts and usage of technology. While some of the pioneering works are often technology-driven, proving the feasibility of hardware-heavy combinations, newer works are more mindful of the usage of technology. This can be attributed to advances in technology (e.g., advanced built-in tracking capabilities) – minimizing the need for additional technical augmentation of hybrid user interface environments. However, our literature survey (see Chapter 3) has shown an increase in various evaluation strategies for hybrid user interfaces over time. With this, the focus shifted from purely technical work towards studying the individual benefits of various design alternatives (e.g., different variants of hybrid user interfaces) to emphasize the value and the utility of individual components and the user experience of hybrid user interfaces. Similarly, current off-the-shelf devices (e.g., Apple Vision Pro[2]) and applications (e.g., Immersed[3]) already allow to use hybrid user interfaces: Enabling a larger display space by *migrating* the visual output from a 2D display to AR. With the maturation of concepts, an increase in empirical contributions, the advent of consumer-oriented hybrid user interfaces, and the development of other areas such as artificial intelligence or brain-computer interfaces, we can now take a step back and look at the bigger picture:

What does the future of hybrid user interfaces look like? With our attributes, we identified a classification of past and current hybrid user interfaces. The interplay of interaction components is fundamental for hybrid user interfaces: A typical hybrid user interface consists of multiple interaction components in heterogeneous roles, complementing each other, and forming a single application. Past and present hybrid user interfaces often consider "*interaction component*" as a synonym for "*device*" – creating a cross-device interaction. However, by reconsidering what constitutes an interaction component, we can draw inspiration from concepts such as *complementary interfaces* [424] and related ideas [109]: These concepts are device-agnostic, opening up the research space for combinations of interaction components beyond devices and avoiding being trapped in technological restrictions. Here, new meaningful combinations of interaction components that include different input and output modalities, combinations of implicit and explicit interaction, and various interaction techniques provide an avenue of research opportunities.

9.6 Closing Remarks

It seems inevitable that MR environments will continue to gain relevance. Yet, when we look at the evolution of similar technologies, such as smartphones, it also becomes clear that this technology is still in its infancy, as a true integration into our everyday lives and technological maturity can take decades. Given the current pace of hardware iterations, it remains uncertain when we will approach the "*ultimate display*" [377] – if we can reach it at all.

While technological evolutions remain uncertain, the field of human–computer interaction often looks past such hardware restrictions of today to envision the user interface of the future. This design is often driven by futuristic visions popularized in science fiction [212]. 20 years ago, such visions consisted of floating displays with gestural interactions in mid-air. In contrast, contemporary visions instead depict a rich ecology of futuristic devices (see Figure 9.1): In "*The Expanse*"[4], for example, designers envision the seamless integration of futuristic smartphones and desktop computers within immersive AR environments, leveraging devices as tools that *complement* each other, rather than being superseded by a singular

[2] https://www.apple.com/apple-vision-pro/, last accessed on 2025-05-12
[3] https://immersed.com, last accessed on 2025-05-12
[4] https://www.hudsandguis.com/home/2021/theexpanse, last accessed 2025-05-12.

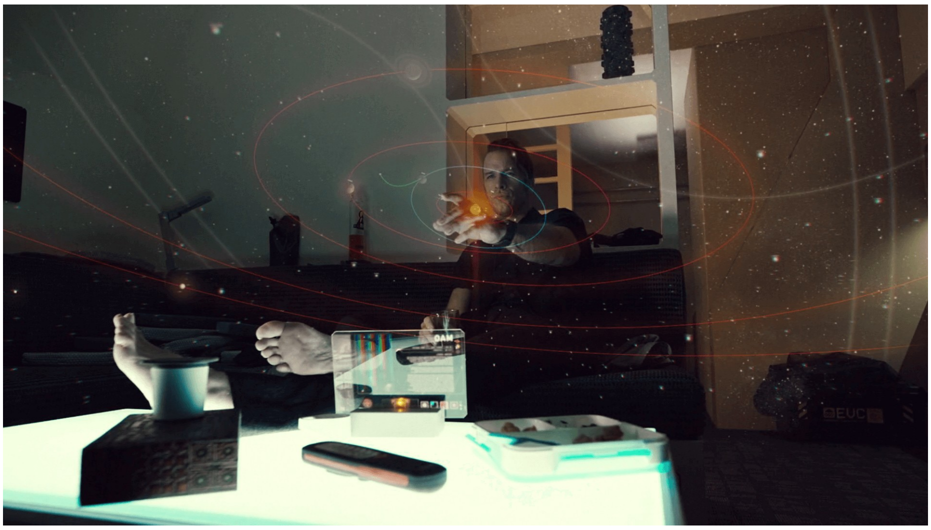


Figure 9.1: User interface design in "*The Expanse*"[4] depicts futuristic cross-device interaction with heterogeneous devices, such as transparent tablets and immersive AR environments.

ultimate display – a vision that shares many similarities with the hybrid user interfaces of today.

Ultimately, whether hybrid user interfaces will persist in the long term is difficult to predict and will depend on how MR technology evolves in the *unforeseeable* future. With the hardware of the *foreseeable* future, however, hybrid user interfaces will remain a pragmatic design strategy within the constraints of today's hardware.

Bibliography

[1] Florian "Floyd" Mueller, Tim Dwyer, Sarah Goodwin, Kim Marriott, Jialin Deng, Han D. Phan, Jionghao Lin, Kun-Ting Chen, Yan Wang, and Rohit Ashok Khot. "Data as Delight: Eating Data." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '21. New York, NY, USA: ACM, 2021, pp. 1–14. ISBN: 978-1-4503-8096-6. DOI: 10.1145/3411764.3445218 (cit. on p. 115).

[2] M. Agarabi, P. Bonato, and C.J. De Luca. "A sEMG-based Method for Assessing the Design of Computer Mice." In: *The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Vol. 3. IEMBS-04. Piscataway, NJ, USA: IEEE, 2004, pp. 2450–2453. ISBN: 978-0-7803-8439-2. DOI: 10.1109/IEMBS.2004.1403708 (cit. on pp. 149, 155).

[3] Michael Aichem, Karsten Klein, Tobias Czauderna, Dimitar Garkov, Jinxin Zhao, Jian Li, and Falk Schreiber. "Towards a hybrid user interface for the visual exploration of large biomolecular networks using virtual reality." In: *Journal of Integrative Bioinformatics* 19.4 (2022). ISSN: 1613-4516. DOI: 10.1515/jib-2022-0034 (cit. on pp. 42, 142).

[4] David Aigner, Nanjia Wang, David Kielmayer, Jürgen Steiner, Julian Hochpöchler, Christoph Heinzl, Daniel Roth, Frank Maurer, and Christoph Anthes. "Cardiac Visualisation Along the RV-Continuum - A High-Fidelity Pilot Study." In: *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2023, pp. 675–680. DOI: 10.1109/ismar-adjunct60411.2023.00145 (cit. on pp. 42, 48, 49, 177).

[5] Gary L. Allen, ed. *Human Spatial Memory*. 1st ed. Psychology Press, 2004. ISBN: 9781135635138. DOI: 10.4324/9781410609984 (cit. on p. 86).

[6] Mohammadreza Amini, Wolfgang Stuerzlinger, Robert J Teather, and Anil Ufuk Batmaz. "A Systematic Review of Fitts' Law in 3D Extended Reality." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '25. New York, NY, USA: ACM, 2025, pp. 1–25. ISBN: 979-8-4007-1394-1. DOI: 10.1145/3706598.3713623 (cit. on p. 167).

[7] Christoph Anthes, Mark Billinghurst, Uwe Grünefeld, Hans-Christian Jetter, Hai-Ning Liang, Frank Maurer, David Aigner, Craig Anslow, Guillaume Bataille, Abraham Campbell, Judith Friedl-Knirsch, Alexander Gall, Renan Guarese, **Sebastian Hubenschmid**, Yue Li, Fabian Pointecker, Andreas Riegler, Daniel Roth, Rishi Vanukuru, Nanjia Wang, Lingyun Yu, Johannes Zagermann, and Daniel Zielasko. "Grand Challenges in Cross Reality." Submitted to: *IEEE TVCG*. 2025 (cit. on pp. x, 19–21).

[8] Rahul Arora, Rubaiat Habib Kazi, Tovi Grossman, George Fitzmaurice, and Karan Singh. "SymbiosisSketch: Combining 2D & 3D Sketching for Designing Detailed 3D Objects in Situ." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '18. New York, NY, USA: ACM, 2018, pp. 1–15. isbn: 978-1-4503-5620-6. doi: 10.1145/3173574.3173759 (cit. on pp. 40, 43, 49).

[9] Jonas Auda, Uwe Gruenefeld, Sarah Faltaous, Sven Mayer, and Stefan Schneegass. "A Scoping Survey on Cross-reality Systems." In: *ACM Computing Surveys* 56.4 (2023), pp. 1–38. issn: 1557-7341. doi: 10.1145/3616536 (cit. on pp. 20, 35).

[10] Hossein Azari, Irene Cheng, and Anup Basu. "Stereo 3D Mouse Cursor: A Method for Interaction with 3D Objects in a Stereoscopic Virtual 3D Space." In: *International Journal of Digital Multimedia Broadcasting* 2010 (2010), pp. 1–11. issn: 1687-7586. doi: 10.1155/2010/419493 (cit. on p. 150).

[11] Teo Babic, Florian Perteneder, Harald Reiterer, and Michael Haller. "Simo: Interactions with Distant Displays by Smartphones with Simultaneous Face and World Tracking." In: *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems*. CHI '20. ACM, 2020, pp. 110–121. doi: 10.1145/3334480.3382962 (cit. on p. 4).

[12] Sriram Karthik Badam, Fereshteh Amini, Niklas Elmqvist, and Pourang Irani. "Supporting Visual Exploration for Multiple Users in Large Display Environments." In: *Proceedings of the IEEE Conference on Visual Analytics Science & Technology*. IEEE, 2016, pp. 1–10. doi: 10.1109/VAST.2016.7883506 (cit. on pp. 67, 72, 77).

[13] Sriram Karthik Badam, Andreas Mathisen, Roman Radle, Clemens N. Klokmose, and Niklas Elmqvist. "Vistrates: A Component Model for Ubiquitous Analytics." In: *IEEE Transactions on Visualization and Computer Graphics* 25.1 (2019), pp. 586–596. issn: 2160-9306. doi: 10.1109/TVCG.2018.2865144 (cit. on pp. 112, 115, 116, 119, 121, 140).

[14] Huidong Bai, Lei Gao, and Mark Billinghurst. "6DoF Input for Hololens Using Vive Controller." In: *SIGGRAPH Asia 2017 Mobile Graphics & Interactive Applications on - SA '17*. SA '17. New York, NY, USA: ACM, 2017, pp. 1–1. isbn: 978-1-4503-5410-3. doi: 10.1145/3132787.3132814 (cit. on p. 68).

[15] Caroline Baillard, Matthieu Fradet, Vincent Alleaume, Pierrick Jouet, and Anthony Laurent. "Multi-Device Mixed Reality TV: A Collaborative Experience with Joint Use of a Tablet and a Headset." In: *Proceedings of the IEEE Symposium on Virtual Reality Software & Technology*. VRST '17. New York, NY, USA: ACM, 2017, pp. 1–2. isbn: 978-1-4503-5548-3. doi: 10.1145/3139131.3141196 (cit. on p. 49).

[16] Ravin Balakrishnan, Thomas Baudel, Gordon Kurtenbach, and George Fitzmaurice. "The Rockin'Mouse: Integral 3D Manipulation on a Plane." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI97. New York, NY, USA: ACM, 1997, pp. 311–318. isbn: 978-0-89791-802-2. doi: 10.1145/258549.258778 (cit. on pp. 147, 151).

[17] Till Ballendat, Nicolai Marquardt, and Saul Greenberg. "Proxemic Interaction: Designing for a Proximity and Orientation-Aware Environment." In: *Proceedings of the ACM Conference on Interactive Tabletops and Surfaces*. ITS'10. ACM, 2010, pp. 121–130. doi: 10.1145/1936652.1936676 (cit. on p. 67).

[18] Daniel Bambušek, Zdeněk Materna, Michal Kapinus, Vítězslav Beran, and Pavel Smrž. "How Do I Get There? Overcoming Reachability Limitations of Constrained Industrial Environments in Augmented Reality Applications." In: *Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces*. Piscataway, NJ, USA: IEEE, 2023, pp. 115–122. isbn: 979-8-3503-4815-6. doi: 10.1109/VR55154.2023.00027 (cit. on p. 37).

[19] Sunyoung Bang, Hyunjin Lee, Seo Young Oh, and Woontack Woo. "AReading with Smartphones: Understanding the Trade-offs between Enhanced Legibility and Display Switching Costs in Hybrid AR Interfaces." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '25. New York, NY, USA: ACM, 2025, pp. 1–20. isbn: 979-8-4007-1394-1. doi: 10.1145/3706598.3713879 (cit. on pp. 59, 107, 174).

[20] Sunyoung Bang and Woontack Woo. "Enhancing the Reading Experience on AR HMDs by Using Smartphones as Assistive Displays." In: *Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces*. Piscataway, NJ, USA: IEEE, 2023, pp. 378–386. isbn: 9798350348156. doi: 10.1109/VR55154.2023.00053 (cit. on p. 49).

[21] Aaron Bangor, Philip Kortum, and James Miller. "Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale." In: *J. Usability Studies* 4.3 (2009), pp. 114–123. issn: 1931-3357 (cit. on p. 138).

[22] Andrea Batch, Andrew Cunningham, Maxime Cordeil, Niklas Elmqvist, Tim Dwyer, Bruce H. Thomas, and Kim Marriott. "There Is No Spoon: Evaluating Performance, Space Use, and Presence with Expert Domain Users in Immersive Analytics." In: *IEEE Transactions on Visualization and Computer Graphics* 26.1 (2020), pp. 536–546. issn: 1077-2626, 1941-0506, 2160-9306. doi: 10.1109/TVCG.2019.2934803 (cit. on pp. 24, 112, 117).

[23] Andrea Batch, Biswaksen Patnaik, Moses Akazue, and Niklas Elmqvist. "Scents and Sensibility: Evaluating Information Olfactation." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '20. New York, NY, USA: ACM, 2020, pp. 1–14. isbn: 978-1-4503-6708-0. doi: 10.1145/3313831.3376733 (cit. on p. 115).

[24] Anil Ufuk Batmaz and Wolfgang Stuerzlinger. "Effective Throughput Analysis of Different Task Execution Strategies for Mid-Air Fitts' Tasks in Virtual Reality." In: *IEEE Transactions on Visualization and Computer Graphics* 28.11 (2022), pp. 3939–3947. issn: 2160-9306. doi: 10.1109/TVCG.2022.3203105 (cit. on p. 167).

[25] Jens Bauer and Achim Ebert. "Virtual Buttons for Eyes-Free Interaction: A Study." In: *Human-Computer Interaction – INTERACT 2015*. Ed. by Julio Abascal, Simone Barbosa, Mirko Fetter, Tom Gross, Philippe Palanque, and Marco Winckler. Vol. 9296. New York, NY, USA: Springer Publishing, 2015, pp. 47–54. isbn: 9783319227016. doi: 10.1007/978-3-319-22701-6_4 (cit. on p. 61).

[26] Jens Bauer, Achim Ebert, Oliver Kreylos, and Bernd Hamann. "Marking Menus for Eyes-Free Interaction Using Smart Phones and Tablets." In: *Availability, Reliability, and Security in Information Systems and HCI*. Ed. by David Hutchison, Takeo Kanade, Josef Kittler, Jon M. Kleinberg, Friedemann Mattern, John C. Mitchell, Moni Naor, Oscar Nierstrasz, C. Pandu Rangan, Bernhard Steffen, Madhu Sudan, Demetri Terzopoulos, Doug Tygar, Moshe Y. Vardi, Gerhard Weikum, Alfredo Cuzzocrea,

Christian Kittl, Dimitris E. Simos, Edgar Weippl, and Lida Xu. Vol. 8127. New York, NY, USA: Springer Publishing, 2013, pp. 481–494. isbn: 9783642405112. doi: 10.1007/978-3-642-40511-2_35 (cit. on p. 76).

[27] Michel Beaudouin-Lafon. "Instrumental Interaction: An Interaction Model for Designing Post-WIMP User Interfaces." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '00. New York, NY, USA: ACM, 2000, pp. 446–453. isbn: 978-1-58113-216-8. doi: 10.1145/332040.332473 (cit. on pp. 71, 75).

[28] Martin Bellgardt, Sebastian Pick, Daniel Zielasko, Tom Vierjahn, Benjamin Weyers, and Torsten W. Kuhlen. "Utilizing Immersive Virtual Reality in Everydaywork." In: *2017 IEEE 3rd Workshop on Everyday Virtual Reality (WEVR)*. IEEE, 2017, pp. 1–4. doi: 10.1109/WEVR.2017.7957708 (cit. on p. 135).

[29] Radim Belohlavek and George J. Klir, eds. *Concepts and Fuzzy Logic*. The MIT Press, 2011. isbn: 9780262298575. doi: 10.7551/mitpress/8842.001.0001 (cit. on p. 196).

[30] Hrvoje Benko and Steven Feiner. "Balloon Selection: A Multi-Finger Technique for Accurate Low-Fatigue 3D Selection." In: *Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces*. Los Alamitos, CA, USA: IEEE Computer Society, 2007, pp. 79–86. isbn: 1-4244-0907-1. doi: 10.1109/3DUI.2007.340778 (cit. on p. 150).

[31] François Bérard, Jessica Ip, Mitchel Benovoy, Dalia El-Shimy, Jeffrey R. Blum, and Jeremy R. Cooperstock. "Did "Minority Report" Get It Wrong? Superiority of the Mouse over 3D Input Devices in a 3D Placement Task." In: *Human-Computer Interaction – INTERACT 2009*. Ed. by Tom Gross, Jan Gulliksen, Paula Kotzé, Lars Oestreicher, Philippe Palanque, Raquel Oliveira Prates, and Marco Winckler. Vol. 5727. New York, NY, USA: Springer Publishing, 2009, pp. 400–414. isbn: 9783642036583. doi: 10.1007/978-3-642-03658-3_45 (cit. on pp. 3, 152, 166).

[32] Joanna Bergström, Tor-Salve Dalsgaard, Jason Alexander, and Kasper Hornbæk. "How to Evaluate Object Selection and Manipulation in VR? Guidelines from 20 Years of Studies." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '21. New York, NY, USA: ACM, 2021, pp. 1–20. isbn: 978-1-4503-8096-6. doi: 10.1145/3411764.3445193 (cit. on p. 114).

[33] Joanna Bergstrom-Lehtovirta, Sebastian Boring, and Kasper Hornbæk. "Placing and Recalling Virtual Items on the Skin." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '17. New York, NY, USA: ACM, 2017, pp. 1497–1507. isbn: 978-1-4503-4655-9. doi: 10.1145/3025453.3026030 (cit. on p. 86).

[34] Christopher Berns, Grace Chin, Joel Savitz, Jason Kiesling, and Fred Martin. "MYR: A Web-Based Platform for Teaching Coding Using VR." In: *Proceedings of the ACM Technical Symposium on Computer Science Education*. New York, NY, USA: ACM, 2019, pp. 77–83. isbn: 978-1-4503-5890-3. doi: 10.1145/3287324.3287482 (cit. on pp. 41, 43, 49, 152).

[35] Lonni Besançon, Mickael Sereno, Lingyun Yu, Mehdi Ammi, and Tobias Isenberg. "Hybrid Touch/Tangible Spatial 3D Data Selection." In: *Computer Graphics Forum* 38.3 (2019), pp. 553–567. issn: 1467-8659. doi: 10.1111/cgf.13710 (cit. on pp. 60, 65, 76).

[36] Lonni Besançon, Anders Ynnerman, Daniel F. Keefe, Lingyun Yu, and Tobias Isenberg. “The State of the Art of Spatial Interfaces for 3D Visualization.” In: *Computer Graphics Forum* 40.1 (2021), pp. 293–326. issn: 0167-7055, 1467-8659. doi: 10.1111/cgf.14189 (cit. on pp. 3, 31, 147, 150).

[37] Narjess Bessghaier and Makram Souii. “Towards Usability Evaluation of Hybrid Mobile User Interfaces.” In: *2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA)*. Piscataway, NJ, USA: IEEE, 2017, pp. 895–900. isbn: 978-1-5386-3581-0. doi: 10.1109/AICCSA.2017.10 (cit. on p. 33).

[38] Verena Biener, Daniel Schneider, Travis Gesslein, Alexander Otte, Bastian Kuth, Per Ola Kristensson, Eyal Ofek, Michel Pahud, and Jens Grubert. “Breaking the Screen: Interaction Across Touchscreen Boundaries in Virtual Reality for Mobile Knowledge Workers.” In: *IEEE Transactions on Visualization and Computer Graphics* 26.12 (2020), pp. 3490–3502. issn: 1077-2626, 1941-0506, 2160-9306. doi: 10.1109/TVCG.2020.3023567 (cit. on p. 156).

[39] Mark Billinghurst, Hirokazu Kato, and Ivan Poupyrev. “The MagicBook - moving seamlessly between reality and virtuality.” In: *IEEE Computer Graphics and Applications* 21.3 (2001), pp. 6–8. issn: 1558-1756. doi: 10.1109/38.920621 (cit. on pp. 20, 21).

[40] Per Bjerre, Allan Christensen, Simon André Pedersen, Andreas Køllund Pedersen, and Wolfgang Stuerzlinger. “Transition Times for Manipulation Tasks in Hybrid Interfaces.” In: *Proceedings of the ACM Symposium on Spatial User Interfaces*. SUI ’15. New York, NY, USA: ACM, 2015, pp. 137–137. isbn: 978-1-4503-3703-8. doi: 10.1145/2788940.2794358 (cit. on pp. 147, 150, 162, 164).

[41] Tanja Blascheck, Lonni Besancon, Anastasia Bezerianos, Bongshin Lee, and Petra Isenberg. “Glanceable Visualization: Studies of Data Comparison Performance on Smartwatches.” In: *IEEE Transactions on Visualization and Computer Graphics* 25.1 (2019), pp. 630–640. issn: 2160-9306. doi: 10.1109/TVCG.2018.2865142 (cit. on p. 47).

[42] Sebastian Boring, Marko Jurmu, and Andreas Butz. “Scroll, Tilt or Move It: Using Mobile Phones to Continuously Control Pointers on Large Public Displays.” In: *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7*. OZCHI ’09. New York, NY, USA: ACM, 2009, pp. 161–168. isbn: 978-1-60558-854-4. doi: 10.1145/1738826.1738853 (cit. on pp. 3, 4).

[43] A. Bornik, R. Beichel, E. Kruijff, B. Reitinger, and D. Schmalstieg. “A Hybrid User Interface for Manipulation of Volumetric Medical Data.” In: *Proceedings of the IEEE Symposium on 3D User Interfaces*. Piscataway, NJ, USA: IEEE, 2006, pp. 29–36. doi: 10.1109/vr.2006.8 (cit. on pp. 21, 32–34, 42, 48, 49, 149, 150, 163, 175).

[44] Alexander Bornik, Reinhard Beichel, and Dieter Schmalstieg. “Interactive editing of segmented volumetric datasets in a hybrid 2D/3D virtual environment.” In: *Proceedings of the IEEE Symposium on Virtual Reality Software & Technology*. VRST06. New York, NY, USA: ACM, 2006, pp. 197–206. isbn: 1-59593-321-2. doi: 10.1145/1180495.1180536 (cit. on pp. 44, 45, 150, 163).

[45] Marcel Borowski, Roman Rädle, and Clemens N. Klokmose. “Codestrate Packages: An Alternative to ”One-Size-Fits-All” Software.” In: *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems*. CHI ’18. New York, NY, USA: ACM, 2018, pp. 1–6. ISBN: 978-1-4503-5621-3. DOI: 10.1145/3170427.3188563 (cit. on pp. 116, 119–121, 140).

[46] Marcel Borowski, Johannes Zagermann, Clemens N. Klokmose, Harald Reiterer, and Roman Rädle. “Exploring the Benefits and Barriers of Using Computational Notebooks for Collaborative Programming Assignments.” In: *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*. SIGCSE ’20. New York, NY, USA: ACM, 2020, pp. 468–474. ISBN: 978-1-4503-6793-6. DOI: 10.1145/3328778.3366887 (cit. on p. 140).

[47] M. Bostock, V. Ogievetsky, and J. Heer. “D^3 Data-Driven Documents.” In: *IEEE Transactions on Visualization and Computer Graphics* 17.12 (2011), pp. 2301–2309. ISSN: 1077-2626. DOI: 10.1109/TVCG.2011.185 (cit. on p. 115).

[48] D. A. Bowman, C. A. Wingrave, J. M. Campbell, V. Q. Ly, and C. J. Rhoton. “Novel Uses of Pinch Gloves™ for Virtual Environment Interaction Techniques.” In: *Virtual Reality* 6.3 (2002), pp. 122–129. ISSN: 1359-4338, 1434-9957. DOI: 10.1007/s100550200013 (cit. on pp. 3, 147).

[49] Eugenie Brasier, Emmanuel Pietriga, and Caroline Appert. “AR-enhanced Widgets for Smartphone-centric Interaction.” In: *Proceedings of the ACM Conference on Mobile Human-Computer Interaction*. MobileHCI ’21. New York, NY, USA: ACM, 2021, 32:1–32:12. ISBN: 978-1-4503-8328-8. DOI: 10.1145/3447526.3472019 (cit. on pp. 41, 49).

[50] Richard Brath. “3D InfoVis Is Here to Stay: Deal with It.” In: *2014 IEEE VIS International Workshop on 3DVis (3DVis)*. Piscataway, NJ, USA: IEEE, 2014, pp. 25–31. ISBN: 978-1-4799-6826-8. DOI: 10.1109/3DVis.2014.7160096 (cit. on pp. 7, 24, 151).

[51] Virginia Braun and Victoria Clarke. “Using Thematic Analysis in Psychology.” In: *Qualitative Research in Psychology* 3.2 (2006), pp. 77–101. ISSN: 1478-0887, 1478-0895. DOI: 10.1191/1478088706qp063oa (cit. on p. 134).

[52] Stephen Brewster, Joanna Lumsden, Marek Bell, Malcolm Hall, and Stuart Tasker. “Multimodal ’eyes-Free’ Interaction Techniques for Wearable Devices.” In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI ’03. New York, NY, USA: ACM, 2003, pp. 473–480. ISBN: 978-1-58113-630-2. DOI: 10.1145/642611.642694 (cit. on p. 61).

[53] John Brooke. “SUS: A ’Quick and Dirty’ Usability Scale.” In: *Usability Evaluation In Industry*. Ed. by Patrick W. Jordan, B. Thomas, Ian Lyall McClelland, and Bernard Weerdmeester. 0th ed. CRC Press, 1996, pp. 207–212. ISBN: 978-0-429-15701-1. DOI: 10.1201/9781498710411-35 (cit. on pp. 131, 138, 158, 159).

[54] Michelle A. Brown, Wolfgang Stuerzlinger, and E. J. Mendonça Filho. “The Performance of Un-Instrumented In-Air Pointing.” In: *Graphics Interface 2014*. Ed. by Paul G. Kry and Andrea Bunt. 1st ed. A K Peters/CRC Press, 2014, pp. 59–66. ISBN: 9781003059325. DOI: 10.1201/9781003059325-8 (cit. on p. 150).

[55] Gerd Bruder, Frank Steinicke, and Wolfgang Sturzlinger. “To Touch or Not to Touch?: Comparing 2D Touch and 3D Mid-Air Interaction on Stereoscopic Tabletop Surfaces.” In: *Proceedings of the ACM Symposium on Spatial User Interfaces*. SUI ’13. New York, NY, USA: ACM, 2013, pp. 9–16. isbn: 978-1-4503-2141-9. doi: 10.1145/2491367.2491369 (cit. on p. 3).

[56] Frederik Brudy, Christian Holz, Roman Rädle, Chi-Jui Wu, Steven Houben, Clemens Nylandsted Klokmose, and Nicolai Marquardt. “Cross-Device Taxonomy: Survey, Opportunities and Challenges of Interactions Spanning Across Multiple Devices.” In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI ’19. New York, NY, USA: ACM, 2019, 562:1–562:28. isbn: 978-1-4503-5970-2. doi: 10.1145/3290605.3300792 (cit. on pp. 4, 8, 17, 18, 20, 31, 33, 37, 39, 42–44, 47, 51, 181, 182).

[57] Frederik Brudy, Suppachai Suwanwatcharachat, Wenyu Zhang, Steven Houben, and Nicolai Marquardt. “EagleView: A Video Analysis Tool for Visualising and Querying Spatial Interactions of People and Devices.” In: *Proceedings of the ACM Conference on Interactive Surfaces and Spaces*. ISS ’18. New York, NY, USA: ACM, 2018, pp. 61–72. isbn: 978-1-4503-5694-7. doi: 10.1145/3279778.3279795 (cit. on p. 115).

[58] Bernhard Brüning, Christian Schnier, Karola Pitsch, and Sven Wachsmuth. “Integrating PAMOCAT in the Research Cycle: Linking Motion Capturing and Conversation Analysis.” In: *Proceedings of the 14th ACM International Conference on Multimodal Interaction*. ICMI ’12. New York, NY, USA: ACM, 2012, pp. 201–208. isbn: 978-1-4503-1467-1. doi: 10.1145/2388676.2388716 (cit. on p. 115).

[59] Petra Büning. “„Quellentaucher“ – Neue Wege in Der Informationsrecherche.” In: *Bibliotheksdienst* 49.6 (2015), pp. 620–628. issn: 0006-1972. doi: 10.1515/bd-2015-0071 (cit. on p. 26).

[60] Wolfgang Büschel, Anke Lehmann, and Raimund Dachselt. “MIRIA: A Mixed Reality Toolkit for the In-Situ Visualization and Analysis of Spatio-Temporal Interaction Data.” In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI ’21. New York, NY, USA: ACM, 2021, 470:1–470:15. isbn: 978-1-4503-8096-6. doi: 10.1145/3411764.3445651 (cit. on pp. 112–116, 120, 124, 126, 139, 149, 182).

[61] Wolfgang Büschel, Annett Mitschick, Thomas Meyer, and Raimund Dachselt. “Investigating Smartphone-based Pan and Zoom in 3D Data Spaces in Augmented Reality.” In: *Proceedings of the ACM Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI ’19. New York, NY, USA: ACM, 2019, pp. 1–13. isbn: 978-1-4503-6825-4. doi: 10.1145/3338286.3340113 (cit. on pp. 71, 76, 78).

[62] Wolfgang Büschel, Patrick Reipschläger, Ricardo Langner, and Raimund Dachselt. “Investigating the Use of Spatial Interaction for 3D Data Visualization on Mobile Devices.” In: *Proceedings of the ACM Conference on Interactive Surfaces and Spaces*. ISS ’17. New York, NY, USA: ACM, 2017, pp. 62–71. isbn: 978-1-4503-4691-7. doi: 10.1145/3132272.3134125 (cit. on pp. 112, 114, 116).

[63] Peter W. S. Butcher, Andrea Batch, Panagiotis D. Ritsos, and Niklas Elmqvist. “Don’t Pull the Balrog — Lessons Learned from Designing Wizualization: a Magic-inspired Data Analytics System in XR.” In: *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Vol. 30. Piscataway, NJ, USA: IEEE, 2023, pp. 234–237. doi: 10.1109/ismar-adjunct60411.2023.00054 (cit. on pp. 1, 182).

[64] Peter W. S. Butcher, Nigel W. John, and Panagiotis D. Ritsos. “VRIA - A Framework for Immersive Analytics on the Web.” In: *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems*. CHI EA ’19. New York, NY, USA: ACM, 2019, pp. 1–6. isbn: 978-1-4503-5971-9. doi: 10.1145/3290607.3312798 (cit. on p. 24).

[65] Simon Butscher. “Reality-based Idioms : Designing Interfaces for Visual Data Analysis that Provide the Means for Familiar Interaction.” PhD thesis. Konstanz: Universität Konstanz, 2018 (cit. on p. 194).

[66] Simon Butscher, **Sebastian Hubenschmid**, Jens Müller, Johannes Fuchs, and Harald Reiterer. “Clusters, Trends, and Outliers: How Immersive Technologies Can Facilitate the Collaborative Analysis of Multidimensional Data.” In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI ’18. Montreal, Canada: ACM, 2018, pp. 1–12. isbn: 978-1-4503-5620-6. doi: 10.1145/3173574.3173664 (cit. on pp. x, 24, 40, 42–44, 48, 59, 61, 62, 78, 181).

[67] Simon Butscher and Harald Reiterer. “Applying Guidelines for the Design of Distortions on Focus+Context Interfaces.” In: *Proceedings of the ACM Conference on Advanced Visual Interfaces*. AVI ’16. New York, NY, USA: ACM, 2016, pp. 244–247. isbn: 978-1-4503-4131-8. doi: 10.1145/2909132.2909284 (cit. on p. 86).

[68] Andreas Butz, Tobias Höllerer, Steven Feiner, Blair MacIntyre, and Clifford Beshers. “Enveloping Users and Computers in a Collaborative 3D Augmented Reality.” In: *Proceedings of the IEEE and ACM International Workshop on Augmented Reality*. IWAR-99. Los Alamitos, CA, USA: IEEE Computer Society, 1999, pp. 35–44. isbn: 0-7695-0359-4. doi: 10.1109/IWAR.1999.803804 (cit. on pp. 21, 32, 49).

[69] Giuseppe Caggianese, Luigi Gallo, and Pietro Neroni. “The Vive Controllers vs. Leap Motion for Interactions in Virtual Environments: A Comparative Evaluation.” In: *Intelligent Interactive Multimedia Systems and Services*. Ed. by Giuseppe De Pietro, Luigi Gallo, Robert J. Howlett, Lakhmi C. Jain, and Ljubo Vlacic. Vol. 98. New York, NY, USA: Springer Publishing, 2019, pp. 24–33. isbn: 9783319922317. doi: 10.1007/978-3-319-92231-7_3 (cit. on pp. 147, 150).

[70] Nicko R. Caluya, Alexander Plopski, Christian Sandor, Yuichiro Fujimoto, Masayuki Kanbara, and Hirokazu Kato. “Does Overlay Field of View in Head-Mounted Displays Affect Spatial Memorization?” In: *Computers & Graphics* 102 (2022), pp. 554–565. issn: 00978493. doi: 10.1016/j.cag.2021.09.004 (cit. on pp. 86, 87, 102, 104).

[71] Felipe G. Carvalho, Daniela G. Trevisan, and Alberto Raposo. “Toward the design of transitional interfaces: an exploratory study on a semi-immersive hybrid user interface.” In: *Virtual Reality* 16.4 (2011), pp. 271–288. issn: 1434-9957. doi: 10.1007/s10055-011-0205-y (cit. on pp. 20–22, 31, 124, 125, 150).

[72] Jessica R. Cauchard, Markus Löchtefeld, Pourang Irani, Johannes Schoening, Antonio Krüger, Mike Fraser, and Sriram Subramanian. “Visual separation in mobile multi-display environments.” In: *Proceedings of the ACM Symposium on User Interface Software and Technology*. UIST ’11. New York, NY, USA: ACM, 2011, pp. 451–460. ISBN: 978-1-4503-0716-1. DOI: 10.1145/2047196.2047256 (cit. on pp. 45, 87).

[73] Marco Cavallo, Mishal Dolakia, Matous Havlena, Kenneth Ocheltree, and Mark Podlaseck. “Immersive Insights: A Hybrid Analytics System for Collaborative Exploratory Data Analysis.” In: *Proceedings of the IEEE Symposium on Virtual Reality Software & Technology*. VRST ’19. New York, NY, USA: ACM, 2019, pp. 30–52. ISBN: 978-1-4503-7001-1. DOI: 10.1145/3359996.3364242 (cit. on pp. 1, 60, 70, 73, 116).

[74] Sze-Wing Chan, Nanjia Wang, and Frank Maurer. “Single User Cross Reality Workflow for Reservoir Engineering Work in Progress.” In: *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2023, pp. 211–214. ISBN: 9798350328912. DOI: 10.1109/ismar-adjunct60411.2023.00049 (cit. on p. 42).

[75] Li-Wei Chan, Hui-Shan Kao, Mike Y. Chen, Ming-Sui Lee, Jane Hsu, and Yi-Ping Hung. “Touching the Void: Direct-Touch Interaction for Intangible Displays.” In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI ’10. ACM, 2010, pp. 2625–2634. DOI: 10.1145/1753326.1753725 (cit. on p. 3).

[76] Tom Chandler, Maxime Cordeil, Tobias Czauderna, Tim Dwyer, Jaroslaw Glowacki, Cagatay Goncu, Matthias Klapperstueck, Karsten Klein, Kim Marriott, Falk Schreiber, and Elliot Wilson. “Immersive Analytics.” In: *Proceedings of the International Symposium on Big Data Visual Analytics*. Piscataway, NJ, USA: IEEE, 2015, pp. 1–8. ISBN: 978-1-4673-7343-2. DOI: 10.1109/BDVA.2015.7314296 (cit. on pp. 7, 23, 24, 48, 112).

[77] Henry Chen, Austin S. Lee, Mark Swift, and John C. Tang. “3D Collaboration Method over HoloLens™ and Skype™ End Points.” In: *Proceedings of the 3rd International Workshop on Immersive Media Experiences - ImmersiveME ’15*. MM ’15. ACM, 2015, pp. 27–30. DOI: 10.1145/2814347.2814350 (cit. on pp. 113, 142).

[78] Yi Fei Cheng, Tiffany Luong, Andreas Rene Fender, Paul Streli, and Christian Holz. “ComforTable User Interfaces: Surfaces Reduce Input Error, Time, and Exertion for Tabletop and Mid-air User Interfaces.” In: *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2022, pp. 150–159. ISBN: 978-1-6654-5325-7. DOI: 10.1109/ISMAR55827.2022.00029 (cit. on pp. 3, 150, 152).

[79] EunJeong Cheon, Eike Schneiders, Kristina Diekjobst, and Mikael B. Skov. “Robots as a Place for Socializing: Influences of Collaborative Robots on Social Dynamics In- and Outside the Production Cells.” In: *Proceedings of the ACM on Human-Computer Interaction* 6.CSCW2 (2022), pp. 1–26. ISSN: 2573-0142. DOI: 10.1145/3555558 (cit. on p. 142).

[80] L. Chittaro, R. Ranon, and L. Ieronutti. “VU-Flow: A Visualization Tool for Analyzing Navigation in Virtual Environments.” In: *IEEE Transactions on Visualization and Computer Graphics* 12.6 (2006), pp. 1475–1485. ISSN: 1077-2626. DOI: 10.1109/TVCG.2006.109 (cit. on p. 115).

[81] Hyunwoo Cho, Bowen Yuan, Jonathon Derek Hart, Zhuang Chang, Jiashuo Cao, Eunhee Chang, and Mark Billinghurst. "Time Travellers: An Asynchronous Cross Reality Collaborative System." In: *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2023, pp. 848–853. ISBN: 979-8-3503-2891-2. DOI: 10.1109/ISMAR-Adjunct60411.2023.00186 (cit. on p. 37).

[82] Neil Chulpongsatorn, Wesley Willett, and Ryo Suzuki. "HoloTouch: Interacting with Mixed Reality Visualizations Through Smartphone Proxies." In: *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems*. CHI '23. New York, NY, USA: ACM, 2023, 156:1–156:8. ISBN: 978-1-4503-9422-2. DOI: 10.1145/3544549.3585738 (cit. on pp. 41, 42, 44, 48).

[83] Jinsung Chun, Byeonguk Bae, and Sungho Jo. "BCI Based Hybrid Interface for 3D Object Control in Virtual Reality." In: *2016 4th International Winter Conference on Brain-Computer Interface (BCI)*. Piscataway, NJ, USA: IEEE, 2016, pp. 101–104. ISBN: 978-1-4673-7841-3. DOI: 10.1109/IWW-BCI.2016.7457461 (cit. on p. 33).

[84] Andy Cockburn and Bruce McKenzie. "Evaluating the Effectiveness of Spatial Memory in 2D and 3D Physical and Virtual Environments." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '02. New York, NY, USA: ACM, 2002, pp. 203–210. ISBN: 978-1-58113-453-7. DOI: 10.1145/503376.503413 (cit. on p. 87).

[85] Dane Coffey, Nicholas Malbraaten, Trung Le, Iman Borazjani, Fotis Sotiropoulos, and Daniel F. Keefe. "Slice WIM: A Multi-Surface, Multi-Touch Interface for Overview+Detail Exploration of Volume Datasets in Virtual Reality." In: *Proceeding of the Symposium on Interactive 3D Graphics and Games (I3D '11)*. Vol. D. I3D '11. New York, NY, USA: ACM, 2011, pp. 191–198. ISBN: 978-1-4503-0565-5. DOI: 10.1145/1944745.1944777 (cit. on p. 150).

[86] Robbe Cools, Matt Gottsacker, Adalberto Simeone, Gerd Bruder, Greg Welch, and Steven Feiner. "Towards a Desktop-AR Prototyping Framework: Prototyping Cross-Reality between Desktops and Augmented Reality." In: *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2022, pp. 175–182. DOI: 10.1109/ISMAR-Adjunct57072.2022.00040 (cit. on pp. 147, 150, 166).

[87] Robbe Cools, Inne Maerevoet, Matt Gottsacker, and Adalberto L. Simeone. "Comparison of Cross-Reality Transition Techniques between 3D and 2D Display Spaces in Desktop–AR Systems." In: *IEEE Transactions on Visualization and Computer Graphics* 31.5 (2025), pp. 2798–2805. ISSN: 2160-9306. DOI: 10.1109/TVCG.2025.3549907 (cit. on pp. 147, 150, 156, 162, 164).

[88] Maxime Cordeil, Andrew Cunningham, Benjamin Bach, Christophe Hurter, Bruce H. Thomas, Kim Marriott, and Tim Dwyer. "IATK: An Immersive Analytics Toolkit." In: *Proceedings of the IEEE Conference on Virtual Reality*. Piscataway, NJ, USA: IEEE, 2019, pp. 200–209. ISBN: 978-1-7281-1377-7. DOI: 10.1109/VR.2019.8797978 (cit. on pp. 24, 59, 61, 183).

[89] Maxime Cordeil, Andrew Cunningham, Tim Dwyer, Bruce H. Thomas, and Kim Marriott. "ImAxes: Immersive Axes as Embodied Affordances for Interactive Multivariate Data Visualisation." In: *Proceedings of the ACM Symposium on User Interface Software and Technology*. UIST '17. New York, NY, USA: ACM, 2017, pp. 71–83.

ISBN: 978-1-4503-4981-9. DOI: 10.1145/3126594.3126613 (cit. on pp. 24, 59, 61).

[90] Tarik Crnovrsanin, Sara Di Bartolomeo, Connor Wilson, and Cody Dunne. *Indy Survey Tool: A Framework to Unearth Correlations in Survey Data*. 2023. DOI: 10.31219/osf.io/um9gs (cit. on p. 38).

[91] Carolina Cruz-Neira, Dj Daniel J Sandin, and Ta Thomas A DeFanti. "Surround-Screen Projection-Based Virtual Reality : The Design and Implementation of the CAVE." In: *Proceedings of the Annual Conference on Computer Graphics and Interactive Techniques*. SIGGRAPH93. New York, NY, USA: ACM, 1993, pp. 135–142. ISBN: 0-89791-601-8. DOI: 10.1145/166117.166134 (cit. on p. 105).

[92] Mihaly Csikszentmihalyi. *Beyond Boredom and Anxiety: Experiencing Flow in Work and Play*. 1. ed., 25. anniversary ed. San Francisco: Jossey-Bass, 2000. ISBN: 978-0-7879-5140-5 (cit. on p. 26).

[93] Parisa Daeijavad. "Investigating the Impact of Multiple View Layouts on Users' Visual Task Performance in Extended Reality." In: *Proceedings of the ACM Conference on Advanced Visual Interfaces*. AVI 2024. New York, NY, USA: ACM, 2024, pp. 1–3. ISBN: 979-8-4007-1764-2. DOI: 10.1145/3656650.3656756 (cit. on p. 107).

[94] Parisa Daeijavad and Frank Maurer. "Hybrid User Interfaces for Multiple Views: why designer intuition is not enough." In: *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2023, pp. 196–198. DOI: 10.1109/ISMAR-Adjunct60411.2023.00045 (cit. on p. 182).

[95] Florian Daiber, Eric Falk, and Antonio Krüger. "Balloon Selection Revisited: Multi-Touch Selection Techniques for Stereoscopic Data." In: *Proceedings of the ACM Conference on Advanced Visual Interfaces*. AVI'12. New York, NY, USA: ACM, 2012, pp. 441–444. ISBN: 978-1-4503-1287-5. DOI: 10.1145/2254556.2254641 (cit. on p. 150).

[96] Satabdi Das, Arshad Nasser, and Khalad Hasan. "FingerButton: Enabling Controller-Free Transitions between Real and Virtual Environments." In: *2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. Piscataway, NJ, USA: IEEE, 2023, pp. 533–542. ISBN: 979-8-3503-2838-7. DOI: 10.1109/ISMAR59233.2023.00068 (cit. on p. 37).

[97] Felipe Gomes de Carvalho, Alberto Raposo, and Marcelo Gattass. "Designing a Hybrid User Interface: A Case Study on an Oil and Gas Application." In: *Proceedings of the 8th International Conference on Virtual Reality Continuum and Its Applications in Industry*. VRCAI '09. New York, NY, USA: ACM, 2009, pp. 191–196. ISBN: 978-1-60558-912-1. DOI: 10.1145/1670252.1670293 (cit. on p. 31).

[98] Philip DeCamp, George Shaw, Rony Kubat, and Deb Roy. "An Immersive System for Browsing and Visualizing Surveillance Video." In: *Proceedings of the 18th ACM International Conference on Multimedia*. MM '10. New York, NY, USA: ACM, 2010, pp. 371–380. ISBN: 978-1-60558-933-6. DOI: 10.1145/1873951.1874002 (cit. on p. 115).

[99] Nicolas J. Dedual, Ohan Oda, and Steven K. Feiner. "Creating Hybrid User Interfaces with a 2D Multi-Touch Tabletop and a 3D See-through Head-Worn Display." In: *2011 10th IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2011, pp. 231–232. ISBN: 978-1-4577-2185-4. DOI: 10.1109/ISMAR.2011.6092391 (cit. on p. 41).

[100] Joachim Deisinger, Roland Blach, Gerold Wesche, Ralf Breining, and Andreas Simon. "Towards Immersive Modeling — Challenges and Recommendations: A Workshop Analyzing the Needs of Designers." In: *Virtual Environments 2000*. Ed. by W. Hansmann, W. Purgathofer, F. Sillion, Jurriaan Mulder, and Robert Van Liere. New York, NY, USA: Springer Publishing, 2000, pp. 145–156. isbn: 978-3-211-83516-6. doi: 10.1007/978-3-7091-6785-4_16 (cit. on pp. 3, 150, 152).

[101] Dazhen Deng, Weiwei Cui, Xiyu Meng, Mengye Xu, Yu Liao, Haidong Zhang, and Yingcai Wu. "Revisiting the Design Patterns of Composite Visualizations." In: *IEEE Transactions on Visualization and Computer Graphics* 29.12 (2023), pp. 5406–5421. issn: 2160-9306. doi: 10.1109/TVCG.2022.3213565 (cit. on p. 25).

[102] Francesco Riccardo Di Gioia, Eugenie Brasier, Emmanuel Pietriga, and Caroline Appert. "Investigating the Use of AR Glasses for Content Annotation on Mobile Devices." In: *Proceedings of the ACM on Human-Computer Interaction* 6.ISS (2022), pp. 430–447. issn: 2573-0142. doi: 10.1145/3567727 (cit. on p. 105).

[103] Nikos Dimitropoulos, Theodoros Togias, George Michalos, and Sotiris Makris. "Operator Support in Human–Robot Collaborative Environments Using AI Enhanced Wearable Devices." In: *Procedia CIRP* 97 (2021), pp. 464–469. issn: 22128271. doi: 10.1016/j.procir.2020.07.006 (cit. on p. 142).

[104] Tobias Drey, Jan Gugenheimer, Julian Karlbauer, Maximilian Milo, and Enrico Rukzio. "VRSketchIn: Exploring the Design Space of Pen and Tablet Interaction for 3D Sketching in Virtual Reality." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '20. New York, NY, USA: ACM, 2020, pp. 1–14. isbn: 978-1-4503-6708-0. doi: 10.1145/3313831.3376628 (cit. on pp. 40–42, 44, 45, 47, 49).

[105] Andrew T. Duchowski, Krzysztof Krejtz, Izabela Krejtz, Cezary Biele, Anna Niedzielska, Peter Kiefer, Martin Raubal, and Ioannis Giannopoulos. "The Index of Pupillary Activity: Measuring Cognitive Load *Vis-à-Vis* Task Difficulty with Pupil Oscillation." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '18. New York, NY, USA: ACM, 2018, pp. 1–13. isbn: 978-1-4503-5620-6. doi: 10.1145/3173574.3173856 (cit. on pp. 90, 96).

[106] Anna Eiberger, Per Ola Kristensson, Susanne Mayr, Matthias Kranz, and Jens Grubert. "Effects of Depth Layer Switching between an Optical See-Through Head-Mounted Display and a Body-Proximate Display." In: *Proceedings of the ACM Symposium on Spatial User Interfaces*. SUI '19. New York, NY, USA: ACM, 2019, 15:1–15:9. isbn: 978-1-4503-6975-6. doi: 10.1145/3357251.3357588 (cit. on pp. 87, 91, 102, 104, 175).

[107] Christian Eichhorn, David A. Plecher, Tobias Mesmer, Lucas Leder, Tim Simecek, Nassim Boukadida, and Gudrun Klinker. "Shopping in between Realities-Using an Augmented Virtuality Smartphone in a Virtual Supermarket." In: *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2023, pp. 1161–1170. isbn: 9798350328387. doi: 10.1109/ismar59233.2023.00133 (cit. on p. 47).

[108] Mats Ole Ellenberg, Marc Satkowski, Weizhou Luo, and Raimund Dachselt. "Spatiality and Semantics - Towards Understanding Content Placement in Mixed Reality." In: *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems*. CHI '23. New York, NY, USA: ACM, 2023, 254:1–254:8. doi: 10.1145/3544549.3585853 (cit. on p. 45).

[109] Niklas Elmqvist. "Data Analytics Anywhere and Everywhere." In: *Communications of the ACM* 66.12 (2023), pp. 52–63. issn: 1557-7317. doi: 10.1145/3584858 (cit. on pp. 112, 197).

[110] Niklas Elmqvist. "Distributed User Interfaces: State of the Art." In: *Distributed User Interfaces*. Ed. by José A. Gallud, Ricardo Tesoriero, and Victor M.R. Penichet. New York, NY, USA: Springer Publishing, 2011, pp. 1–12. isbn: 9781447122715. doi: 10.1007/978-1-4471-2271-5_1 (cit. on pp. 4, 8, 17, 33).

[111] Niklas Elmqvist, Andrew Vande Moere, Hans-Christian Jetter, Daniel Cernea, Harald Reiterer, and T J Jankun-Kelly. "Fluid Interaction for Information Visualization." In: *Information Visualization - Special Issue on State of the Field and New Research Directions* 10.4 (2011), pp. 327–340. issn: 1473-8724. doi: 10.1177/1473871611413180 (cit. on pp. 7, 25, 26, 59, 152, 164, 193).

[112] Neven ElSayed, Bruce Thomas, Kim Marriott, Julia Piantadosi, and Ross Smith. "Situated Analytics." In: *2015 Big Data Visual Analytics (BDVA)*. Piscataway, NJ, USA: IEEE, 2015, pp. 101–108. isbn: 978-1-4673-7343-2. doi: 10.1109/BDVA.2015.7314302 (cit. on pp. 7, 24, 151).

[113] Carmine Elvezio, Pierre Amelot, Robert Boyle, Catherine Ilona Wes, and Steven Feiner. "Hybrid UIs for Music Exploration in AR and VR." In: *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2018, pp. 411–412. isbn: 978-1-5386-7592-2. doi: 10.1109/ISMAR-Adjunct.2018.00121 (cit. on p. 49).

[114] Barrett Ens, Benjamin Bach, Maxime Cordeil, Ulrich Engelke, Marcos Serrano, Wesley Willett, Arnaud Prouzeau, Christoph Anthes, Wolfgang Büschel, Cody Dunne, Tim Dwyer, Jens Grubert, Jason H. Haga, Nurit Kirshenbaum, Dylan Kobayashi, Tica Lin, Monsurat Olaosebikan, Fabian Pointecker, David Saffo, Nazmus Saquib, Dieter Schmalstieg, Danielle Albers Szafir, Matt Whitlock, and Yalong Yang. "Grand Challenges in Immersive Analytics." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '21. New York, NY, USA: ACM, 2021, pp. 1–17. isbn: 978-1-4503-8096-6. doi: 10.1145/3411764.3446866 (cit. on pp. 24–26, 112–114).

[115] Barrett Ens, Juan David Hincapié-Ramos, and Pourang Irani. "Ethereal Planes: A Design Framework for 2D Information Space in 3D Mixed Reality Environments." In: *Proceedings of the ACM Symposium on Spatial User Interfaces*. SUI '14. New York, NY, USA: ACM, 2014, pp. 2–12. isbn: 978-1-4503-2820-3. doi: 10.1145/2659766.2659769 (cit. on p. 4).

[116] Nuno Estalagem and Augusto Esteves. "Between Wearable and Spatial Computing: Exploring Four Interaction Techniques at the Intersection of Smartwatches and Head-mounted Displays." In: *Proceedings of the ACM Symposium on Eye Tracking Research and Applications*. ETRA '24. New York, NY, USA: ACM, 2024, 57:1–57:7. doi: 10.1145/3649902.3656365 (cit. on pp. 41, 43).

[117] João Marcelo Evangelista Belo, Anna Maria Feit, Tiare Feuchtner, and Kaj Grønbæk. "XRgonomics: Facilitating the Creation of Ergonomic 3D Interfaces." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '21. Yokohama, Japan: ACM, 2021, pp. 1–11. ISBN: 9781450380966. DOI: 10.1145/3411764.3445349 (cit. on pp. 112, 114, 115, 182).

[118] Steven Feiner. "Environment management for hybrid user interfaces." In: *IEEE Personal Communications* 7.5 (2000), pp. 50–53. ISSN: 1070-9916. DOI: 10.1109/98.878539 (cit. on pp. 32, 37).

[119] Steven Feiner. "Redefining the User Interface: Augmented Reality." In: *ACM SIGGRAPH* (1994) (cit. on p. 195).

[120] Steven Feiner. "The Importance of Being Mobile: Some Social Consequences of Wearable Augmented Reality Systems." In: *Proceedings 2nd IEEE and ACM International Workshop on Augmented Reality (IWAR'99)*. IWAR-99. Los Alamitos, CA, USA: IEEE Computer Society, 1999, pp. 145–148. ISBN: 978-0-7695-0359-2. DOI: 10.1109/IWAR.1999.803815 (cit. on p. 37).

[121] Steven Feiner, Blair MacIntyre, Tobias Höllerer, and Anthony Webster. "A touring machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment." In: *Personal Technologies* 1.4 (1997), pp. 208–217. ISSN: 1617-4917. DOI: 10.1007/bf01682023 (cit. on pp. 42, 44).

[122] Steven Feiner and Ari Shamash. "Hybrid user interfaces: breeding virtually bigger interfaces for physically smaller computers." In: *Proceedings of the ACM Symposium on User Interface Software and Technology*. UIST '91. New York, NY, USA: ACM, 1991, pp. 9–17. ISBN: 0-89791-451-1. DOI: 10.1145/120782.120783 (cit. on pp. 4–7, 21, 23, 32–34, 36, 41, 43, 44, 48, 49, 85, 191, 195, 196).

[123] Stephen Few. *Information Dashboard Design: The Effective Visual Communication of Data*. O'Reilly Media, Inc., 2006. ISBN: 978-0-596-10016-2 (cit. on p. 79).

[124] Janis Fiedler*, Kathrin Wunsch*, **Sebastian Hubenschmid**, Harald Reiterer, Britta Renner, and Alexander Woll. "SMARTFAMILY2.0: A refined mobile health intervention to promote physical activity and healthy eating in a family setting – A randomized-controlled trial." In: *JMIR mHealth and uHealth* (2025). DOI: 10.2196/65558 (cit. on pp. x, 79).

[125] Daniel Immanuel Fink, Johannes Zagermann, Harald Reiterer, and Hans-Christian Jetter. "Re-locations: Augmenting Personal and Shared Workspaces to Support Remote Collaboration in Incongruent Spaces." In: *Proceedings of the ACM on Human-Computer Interaction* 6.ISS (2022), pp. 1–30. ISSN: 2573-0142. DOI: 10.1145/3567709 (cit. on pp. 117, 181).

[126] M. Fiorentino, R. De Amicis, G. Monno, and A. Stork. "Spacedesign: A Mixed Reality Workspace for Aesthetic Industrial Design." In: *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. ISMAR-02. Los Alamitos, CA, USA: IEEE Computer Society, 2002, pp. 86–318. ISBN: 978-0-7695-1781-0. DOI: 10.1109/ISMAR.2002.1115077 (cit. on p. 152).

[127] Paul M. Fitts and James R. Peterson. "Information Capacity of Discrete Motor Responses." In: *Journal of Experimental Psychology* 67.2 (1964), pp. 103–112. ISSN: 0022-1015. DOI: 10.1037/h0045689 (cit. on p. 167).

[128] Philipp Fleck, Aimée Sousa Calepso, **Sebastian Hubenschmid**, Michael Sedlmair, and Dieter Schmalstieg. "RagRug: A Toolkit for Situated Analytics." In: *IEEE Transactions on Visualization and Computer Graphics* 29.7 (2023), pp. 3281–3297. ISSN: 2160-9306. DOI: 10.1109/TVCG.2022.3157058 (cit. on pp. x, 24, 183).

[129] Philipp Fleck, Fernando Reyes-Aviles, Christian Pirchheim, Clemens Arth, and Dieter Schmalstieg. "MAUI: Tele-assistance for Maintenance of Cyber-physical Systems." In: *15th International Conference on Computer Vision Theory and Applications*. SCITEPRESS - Science and Technology Publications, 2020, pp. 800–812. DOI: 10.5220/0009093708000812 (cit. on pp. 113, 142).

[130] R Foundation. *R Project for Statistical Computing*. 2021. URL: https://www.r-project.org/ (cit. on p. 115).

[131] Judith Friedl-Knirsch, Christian Stach, and Christoph Anthes. "Exploring Collaboration for Data Analysis in Augmented Reality for Multiple Devices." In: *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2023, pp. 65–69. ISBN: 9798350328912. DOI: 10.1109/ismar-adjunct60411.2023.00021 (cit. on pp. 49, 181).

[132] Thibault Friedrich, Arnaud Prouzeau, and Michael McGuffin. "The Effect of Increased Body Motion in Virtual Reality on a Placement-Retrieval Task." In: *Proceedings of the IEEE Symposium on Virtual Reality Software & Technology*. VRST '21. New York, NY, USA: ACM, 2021, pp. 1–5. ISBN: 978-1-4503-9092-7. DOI: 10.1145/3489849.3489888 (cit. on p. 86).

[133] Bernd Froehlich, Jan Hochstrate, Verena Skuk, and Anke Huckauf. "The GlobeFish and the GlobeMouse: Two New Six Degree of Freedom Input Devices for Graphics Applications." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI06. New York, NY, USA: ACM, 2006, pp. 191–199. ISBN: 978-1-59593-372-0. DOI: 10.1145/1124772.1124802 (cit. on p. 151).

[134] Bernd Fröhlich and John Plate. "The Cubic Mouse: A New Device for Three-Dimensional Input." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI00. New York, NY, USA: ACM, 2000, pp. 526–531. ISBN: 978-1-58113-216-8. DOI: 10.1145/332040.332491 (cit. on p. 151).

[135] Bruno Fruchard, Eric Lecolinet, and Olivier Chapuis. "Impact of Semantic Aids on Command Memorization for On-Body Interaction and Directional Gestures." In: *Proceedings of the ACM Conference on Advanced Visual Interfaces*. AVI '18. New York, NY, USA: ACM, 2018, pp. 1–9. ISBN: 978-1-4503-5616-9. DOI: 10.1145/3206505.3206524 (cit. on p. 86).

[136] BoYu Gao, Zhiguo Chen, Xian Chen, Huawei Tu, and Feiran Huang. "The Effects of Audiovisual Landmarks on Spatial Learning and Recalling for Image Browsing Interface in Virtual Environments." In: *Journal of Systems Architecture* 117 (2021), p. 102096. ISSN: 1383-7621. DOI: 10.1016/j.sysarc.2021.102096 (cit. on p. 86).

[137] BoYu Gao, Byungmoon Kim, Jee-In Kim, and HyungSeok Kim. "Amphitheater Layout with Egocentric Distance-Based Item Sizing and Landmarks for Browsing in Virtual Reality." In: *International Journal of Human–Computer Interaction* 35.10 (2019), pp. 831–845. ISSN: 1532-7590. DOI: 10.1080/10447318.2018.1498654 (cit. on pp. 101, 105).

[138] BoYu Gao, HyungSeok Kim, Byungmoon Kim, and Jee-In Kim. “Artificial Landmarks to Facilitate Spatial Learning and Recalling for Curved Visual Wall Layout in Virtual Reality.” In: *2018 IEEE International Conference on Big Data and Smart Computing (BigComp)*. Piscataway, NJ, USA: IEEE, 2018, pp. 475–482. ISBN: 978-1-5386-3649-7. DOI: 10.1109/BigComp.2018.00076 (cit. on pp. 86, 101, 105).

[139] Varun Gaur, Md. Sami Uddin, and Carl Gutwin. “Multiplexing Spatial Memory: Increasing the Capacity of FastTap Menus with Multiple Tabs.” In: *Proceedings of the ACM Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI ’18. New York, NY, USA: ACM, 2018, pp. 1–13. ISBN: 978-1-4503-5898-9. DOI: 10.1145/3229434.3229482 (cit. on p. 86).

[140] Christian Geiger, Robin Fritze, Anke Lehmann, and Jörg Stöcklein. “HYUI: a visual framework for prototyping hybrid user interfaces.” In: *Proceedings of the ACM conference on Tangible and Embedded Interaction*. TEI08. New York, NY, USA: ACM, 2008, pp. 63–70. ISBN: 978-1-60558-004-3. DOI: 10.1145/1347390.1347406 (cit. on pp. 21, 32).

[141] Paul Ginns. “Integrating Information: A Meta-Analysis of the Spatial Contiguity and Temporal Contiguity Effects.” In: *Learning and Instruction* 16.6 (2006), pp. 511–525. ISSN: 0959-4752. DOI: 10.1016/j.learninstruc.2006.10.001 (cit. on pp. 87, 102).

[142] Eric J. Gonzalez, Khushman Patel, Karan Ahuja, and Mar Gonzalez-Franco. “XDTK: A Cross-Device Toolkit for Input & Interaction in XR.” In: *IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops*. Piscataway, NJ, USA: IEEE, 2024, pp. 467–470. DOI: 10.1109/vrw62533.2024.00092 (cit. on pp. 49, 182).

[143] Google. *Colaboratory*. 2021. URL: https://colab.research.google.com (cit. on p. 115).

[144] Matt Gottsacker, Mengyu Chen, David Saffo, Feiyu Lu, and Blair MacIntyre. “Hybrid User Interface for Audience Feedback Guided Asymmetric Immersive Presentation of Financial Data.” In: *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2023, pp. 199–204. ISBN: 9798350328912. DOI: 10.1109/ismar-adjunct60411.2023.00046 (cit. on p. 47).

[145] John W. Graham. “Missing Data Analysis: Making It Work in the Real World.” In: *Annual Review of Psychology* 60.1 (2009), pp. 549–576. ISSN: 0066-4308, 1545-2085. DOI: 10.1146/annurev.psych.58.110405.085530 (cit. on p. 159).

[146] Raphael Grasset, Julian Looser, and Mark Billinghurst. “Transitional interface: concept, issues and framework.” In: *Proceedings of the IEEE/ACM International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2006, pp. 231–232. DOI: 10.1109/ismar.2006.297819 (cit. on p. 21).

[147] Saul Greenberg and Bill Buxton. “Usability Evaluation Considered Harmful (Some of the Time).” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’08. New York, NY, USA: ACM, 2008, pp. 111–120. ISBN: 978-1-60558-011-1. DOI: 10.1145/1357054.1357074 (cit. on p. 9).

[148] Jens Emil Grønbæk. “Transitional Blended Realities: Linking between Video and Mixed Reality for Distributed Team Collaboration.” In: *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2023, pp. 220–222. isbn: 9798350328912. doi: 10.1109/ismar-adjunct60411.2023.00051 (cit. on p. 181).

[149] Jens Emil Sloth Grønbæk, Juan Sánchez Esquivel, Germán Leiva, Eduardo Velloso, Hans Gellersen, and Ken Pfeuffer. “Blended Whiteboard: Physicality and Reconfigurability in Remote Mixed Reality Collaboration.” In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI ’24. New York, NY, USA: ACM, 2024, 798:1–798:16. isbn: 9798400703300. doi: 10.1145/3613904.3642293 (cit. on p. 181).

[150] Jens Grubert, Matthias Heinisch, Aaron Quigley, and Dieter Schmalstieg. “MultiFi: Multi Fidelity Interaction with Displays On and Around the Body.” In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI ’15. New York, NY, USA: ACM, 2015, pp. 3933–3942. isbn: 978-1-4503-3145-6. doi: 10.1145/2702123.2702331 (cit. on pp. 43, 86, 87, 91, 102, 104, 175).

[151] Jens Grubert, Matthias Kranz, and Aaron Quigley. “Challenges in Mobile Multi-Device Ecosystems.” In: *mUX: The Journal of Mobile User Experience* 5.1 (2016). issn: 2196-873X. doi: 10.1186/s13678-016-0007-y (cit. on pp. 59, 60, 65).

[152] Jens Grubert, Eyal Ofek, Michel Pahud, and Per Ola Kristensson. “The Office of the Future: Virtual, Portable, and Global.” In: *IEEE Computer Graphics and Applications* 38.6 (2018), pp. 125–133. issn: 1558-1756. doi: 10.1109/MCG.2018.2875609 (cit. on pp. 4, 149, 165).

[153] Jens Grubert, Lukas Witzani, Eyal Ofek, Michel Pahud, Matthias Kranz, and Per Ola Kristensson. “Text Entry in Immersive Head-Mounted Display-Based Virtual Reality Using Standard Keyboards.” In: *Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces*. Piscataway, NJ, USA: IEEE, 2018, pp. 159–166. isbn: 978-1-5386-3365-6. doi: 10.1109/VR.2018.8446059 (cit. on pp. 59, 150).

[154] Jens Grubert, Lukas Witzani, Alexander Otte, Travis Gesslein, Matthias Kranz, and Per Ola Kristensson. “Text Entry Performance and Situation Awareness of a Joint Optical See-Through Head-Mounted Display and Smartphone System.” In: *IEEE Transactions on Visualization and Computer Graphics* 30.8 (2024), pp. 5830–5846. issn: 2160-9306. doi: 10.1109/tvcg.2023.3309316 (cit. on pp. 41, 47, 49, 59, 150).

[155] Uwe Gruenefeld, Jonas Auda, Florian Mathis, Stefan Schneegass, Mohamed Khamis, Jan Gugenheimer, and Sven Mayer. “VRception: Rapid Prototyping of Cross-Reality Systems in Virtual Reality.” In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI ’22. New York, NY, USA: ACM, 2022, pp. 1–15. isbn: 978-1-4503-9157-3. doi: 10.1145/3491102.3501821 (cit. on p. 147).

[156] Elisa Gusai, Chiara Bassano, Fabio Solari, and Manuela Chessa. “Interaction in an Immersive Collaborative Virtual Reality Environment: A Comparison Between Leap Motion and HTC Controllers.” In: *New Trends in Image Analysis and Processing – ICIAP 2017*. Ed. by Sebastiano Battiato, Giovanni Maria Farinella, Marco Leo, and Giovanni Gallo. Vol. 10590. New York, NY, USA: Springer Publishing, 2017,

pp. 290–300. isbn: 978-3-319-70741-9. doi: 10.1007/978-3-319-70742-6_27 (cit. on pp. 147, 150).

[157] Carl Gutwin and Saul Greenberg. “Design for individuals, design for groups: tradeoffs between power and workspace awareness.” In: *Proceedings of the ACM Conference on Computer-Supported Cooperative Work and Social Computing*. CSCW98. New York, NY, USA: ACM, 1998, pp. 207–216. isbn: 1-58113-009-0. doi: 10.1145/289444.289495 (cit. on p. 181).

[158] Edward T. Hall. *The Hidden Dimension*. Doubleday, 1966 (cit. on p. 18).

[159] Mark Hancock, Otmar Hilliges, Christopher Collins, Dominikus Baur, and Sheelagh Carpendale. “Exploring Tangible and Direct Touch Interfaces for Manipulating 2D and 3D Information on a Digital Table.” In: *Proceedings of the ACM Conference on Interactive Tabletops and Surfaces*. ITS’09. New York, NY, USA: ACM, 2009, p. 77. isbn: 978-1-60558-733-2. doi: 10.1145/1731903.1731921 (cit. on p. 75).

[160] Chris Hand. “A Survey of 3D Interaction Techniques.” In: *Computer Graphics Forum* 16.5 (1997), pp. 269–281. issn: 0167-7055, 1467-8659. doi: 10.1111/1467-8659.00194 (cit. on p. 150).

[161] Jeffrey T. Hansberger, Chao Peng, Shannon L. Mathis, Vaidyanath Areyur Shanthakumar, Sarah C. Meacham, Lizhou Cao, and Victoria R. Blakely. “Dispelling the Gorilla Arm Syndrome: The Viability of Prolonged Gesture Interactions.” In: *Virtual, Augmented and Mixed Reality*. Ed. by Stephanie Lackey and Jessie Chen. Vol. 10280. New York, NY, USA: Springer Publishing, 2017, pp. 505–520. isbn: 9783319579870. doi: 10.1007/978-3-319-57987-0_41 (cit. on pp. 3, 4).

[162] Johannes Harms, Martina Kratky, Christoph Wimmer, Karin Kappel, and Thomas Grechenig. “Navigation in Long Forms on Smartphones: Scrolling Worse than Tabs, Menus, and Collapsible Fieldsets.” In: *Human-Computer Interaction – INTERACT 2015*. Ed. by Julio Abascal, Simone Barbosa, Mirko Fetter, Tom Gross, Philippe Palanque, and Marco Winckler. Vol. 9298. New York, NY, USA: Springer Publishing, 2015, pp. 333–340. isbn: 9783319226989. doi: 10.1007/978-3-319-22698-9_21 (cit. on p. 106).

[163] Sandra G. Hart. “Nasa-Task Load Index (NASA-TLX); 20 Years Later.” In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50.9 (2006), pp. 904–908. issn: 1071-1813. doi: 10.1177/154193120605000909 (cit. on p. 182).

[164] Sandra G. Hart and Lowell E. Staveland. “Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research.” In: *Human Mental Workload*. Advances in psychology, 52. Oxford, UK: Elsevier, 1988, pp. 139–183. isbn: 978-0-444-70388-0. doi: 10.1016/s0166-4115(08)62386-9 (cit. on pp. 90, 92, 98, 158, 159, 182).

[165] Rex Hartson and Pardha Pyla. *The UX Book*. Elsevier, 2018. isbn: 978-0-12-805342-3. doi: 10.1016/C2013-0-19285-9 (cit. on pp. 4, 129).

[166] Anne-Wil Harzing. *Publish or Perish*. 2007. url: http://www.harzing.com/pop.htm (visited on 2025-05-12) (cit. on p. 36).

[167] Antti Hietanen, Roel Pieters, Minna Lanz, Jyrki Latokartano, and Joni-Kristian Kämäräinen. “AR-based Interaction for Human-Robot Collaborative Manufacturing.” In: *Robotics and Computer-Integrated Manufacturing* 63 (2020), pp. 101–125. issn: 07365845. doi: 10.1016/j.rcim.2019.101891 (cit. on p. 142).

[168] Juan David Hincapié-Ramos, Xiang Guo, Paymahn Moghadasian, and Pourang Irani. "Consumed Endurance: A Metric to Quantify Arm Fatigue of Mid-Air Interactions." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '14. New York, NY, USA: ACM, 2014, pp. 1063–1072. isbn: 978-1-4503-2473-1. doi: 10.1145/2556288.2557130 (cit. on pp. 2–4, 114, 152, 168).

[169] Ken Hinckley, Randy Pausch, John C. Goble, and Neal F. Kassell. "A Survey of Design Issues in Spatial Input." In: *Proceedings of the ACM Symposium on User Interface Software and Technology*. UIST '94. New York, NY, USA: ACM, 1994, pp. 213–222. isbn: 978-0-89791-657-8. doi: 10.1145/192426.192501 (cit. on pp. 3, 33, 152, 166).

[170] Ken Hinckley, Mike Sinclair, Erik Hanson, Richard Szeliski, and Matt Conway. "The VideoMouse: A Camera-Based Multi-Degree-of-Freedom Input Device." In: *Proceedings of the ACM Symposium on User Interface Software and Technology*. UIST99. New York, NY, USA: ACM, 1999, pp. 103–112. isbn: 978-1-58113-075-1. doi: 10.1145/320719.322591 (cit. on p. 151).

[171] Uta Hinrichs, Simon Butscher, Jens Müller, and Harald Reiterer. "Diving in at the Deep End: The Value of Alternative In-Situ Approaches for Systematic Library Search." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI'16. New York, NY, USA: ACM, 2016, pp. 4634–4646. isbn: 978-1-4503-3362-7. doi: 10.1145/2858036.2858549 (cit. on p. 26).

[172] Tobias Höllerer, Steven Feiner, Tachio Terauchi, Gus Rashid, and Drexel Hallaway. "Exploring MARS: developing indoor and outdoor user interfaces to a mobile augmented reality system." In: *Computers & Graphics* 23.6 (1999), pp. 779–785. issn: 0097-8493. doi: 10.1016/s0097-8493(99)00103-x (cit. on p. 177).

[173] Tom Horak, Sriram Karthik Badam, Niklas Elmqvist, and Raimund Dachselt. "When David Meets Goliath: Combining Smartwatches with a Large Vertical Display for Visual Data Exploration." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '18. New York, NY, USA: ACM, 2018, pp. 102–113. isbn: 978-1-4503-5620-6. doi: 10.1145/3173574.3173593 (cit. on pp. 46, 61, 76).

[174] Kasper Hornbæk, Benjamin B. Bederson, and Catherine Plaisant. "Navigation Patterns and Usability of Zoomable User Interfaces with and without an Overview." In: *ACM Transactions on Computer-Human Interaction* 9.4 (2002), pp. 362–389. issn: 1073-0516, 1557-7325. doi: 10.1145/586081.586086 (cit. on p. 86).

[175] Kasper Hornbaek. "Some Whys and Hows of Experiments in Human–Computer Interaction." In: *Foundations and Trends® in Human–Computer Interaction* 5.4 (2011), pp. 299–373. issn: 1551-3963. doi: 10.1561/1100000043 (cit. on pp. 9, 181).

[176] Steven Houben, Nicolai Marquardt, Jo Vermeulen, Clemens Klokmose, Johannes Schöning, Harald Reiterer, and Christian Holz. "Opportunities and Challenges for Cross-Device Interactions in the Wild." In: *Interactions* 24.5 (2017), pp. 58–63. issn: 1072-5520, 1558-3449. doi: 10.1145/3121348 (cit. on p. 22).

[177] Yi-Jheng Huang, Kang-Yi Liu, Suiang-Shyan Lee, and I-Cheng Yeh. "Evaluation of a Hybrid of Hand Gesture and Controller Inputs in Virtual Reality." In: *International Journal of Human–Computer Interaction* 37.2 (2021), pp. 169–180. ISSN: 1044-7318, 1532-7590. DOI: 10.1080/10447318.2020.1809248 (cit. on p. 150).

[178] **Sebastian Hubenschmid**, Simon Butscher, Johannes Zagermann, and Harald Reiterer. "Employing Tangible Visualisations in Augmented Reality with Mobile Devices." In: *MultimodalVis '18 Workshop*. 2018, p. 4 (cit. on p. x).

[179] **Sebastian Hubenschmid**, Johannes Zagermann, Simon Butscher, and Harald Reiterer. "STREAM: Exploring the Combination of Spatially-Aware Tablets with Augmented Reality Head-Mounted Displays for Immersive Analytics." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '21. New York, NY, USA: Association for Computing Machinery, 2021, pp. 1–14. ISBN: 978-1-4503-8096-6. DOI: 10.1145/3411764.3445298 (cit. on pp. ix, 6, 40, 42–45, 48, 58, 114, 117).

[180] **Sebastian Hubenschmid**, Johannes Zagermann, Raimund Dachselt, Niklas Elmqvist, Steven Feiner, Tiare Feuchtner, Benjamin Lee, Harald Reiterer, and Dieter Schmalstieg. "Hybrid User Interfaces: Complementary Interfaces for Mixed Reality Interaction." In: Workshop at IEEE ISMAR'23 (2023). DOI: 10.48787/KOPS/352-2-6B4C33KEJAWW2 (cit. on pp. 33, 36).

[181] **Sebastian Hubenschmid**, Johannes Zagermann, Niklas Elmqvist, Tiare Feuchtner, Jens Grubert, Markus Tatzgern, Dieter Schmalstieg, and Harald Reiterer. "Revisiting Hybrid Input Devices for Immersive Analytics." In: *2025 IEEE Conference on Human Factors in Immersive Analytics (HFIA)*. IEEE, 2025. DOI: 10.1109/HFIA68651.2025.00011 (cit. on pp. ix, 146).

[182] **Sebastian Hubenschmid**, Johannes Zagermann, Robin Erb, Tiare Feuchtner, Jens Grubert, Markus Tatzgern, Dieter Schmalstieg, and Harald Reiterer. "SpatialMouse: A Hybrid Pointing Device for Seamless Interaction Across 2D and 3D Spaces." In: *Proceedings of the 2025 31st ACM Symposium on Virtual Reality Software and Technology*. New York, NY, USA: ACM, 2025, pp. 1–13. ISBN: 979-8-4007-2118-2. DOI: 10.1145/3756884.3766047 (cit. on pp. viii, 6, 146).

[183] **Sebastian Hubenschmid**, Johannes Zagermann, Daniel Leicht, Harald Reiterer, and Tiare Feuchtner. "ARound the Smartphone: Investigating the Effects of Virtually-Extended Display Size on Spatial Memory." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '23. Hamburg Germany: ACM, 2023, pp. 1–15. ISBN: 978-1-4503-9421-5. DOI: 10.1145/3544548.3581438 (cit. on pp. viii, 41–43, 84).

[184] **Sebastian Hubenschmid***, Daniel Immanuel Fink*, Johannes Zagermann, Jonathan Wieland, Harald Reiterer, and Tiare Feuchtner. "Colibri: A Toolkit for Rapid Prototyping of Networking Across Realities." In: *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Sydney, Australia: IEEE, 2023, pp. 9–13. ISBN: 9798350328912. DOI: 10.1109/ISMAR-Adjunct60411.2023.00010 (cit. on pp. ix, 37, 172, 182).

[185] **Sebastian Hubenschmid***, Marc Satkowski*, Johannes Zagermann*, Julián Méndez*, Niklas Elmqvist, Steven Feiner, Tiare Feuchtner, Jens Emil Sloth Grønbæk, Benjamin Lee, Dieter Schmalstieg, Raimund Dachselt, and Harald Reiterer. "Hybrid User Interfaces: Past, Present, and Future of Complementary Cross-Device

Interaction in Mixed Reality.” Submitted to: *IEEE Transactions on Visualization and Computer Graphics*. 2025 (cit. on pp. vii, 16, 30, 172, 188).

[186] **Sebastian Hubenschmid***, Jonathan Wieland*, Daniel Immanuel Fink*, Andrea Batch, Johannes Zagermann, Niklas Elmqvist, and Harald Reiterer. “ReLive: Bridging In-Situ and Ex-Situ Visual Analytics for Analyzing Mixed Reality User Studies.” In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI ’22. New Orleans LA USA: ACM, 2022, pp. 1–20. isbn: 978-1-4503-9157-3. doi: 10.1145/3491102.3517550 (cit. on pp. viii, 25, 42, 43, 48, 110, 149, 151, 152, 182, 183).

[187] **Sebastian Hubenschmid***, Johannes Zagermann*, Daniel Fink*, Jonathan Wieland*, Tiare Feuchtner*, and Harald Reiterer*. “Towards Asynchronous Hybrid User Interfaces for Cross-Reality Interaction.” In: *ISS’21 Workshop Proceedings: ”Transitional Interfaces in Mixed and Cross-Reality: A New Frontier?”* 2021. doi: 10.18148/kops/352-2-84mm0sggczq02 (cit. on pp. x, 6, 20, 34, 42, 110, 113).

[188] Christophe Hurter, Nathalie Henry Riche, Steven M. Drucker, Maxime Cordeil, Richard Alligier, and Romain Vuillemot. “FiberClay: Sculpting Three Dimensional Trajectories to Reveal Structural Insights.” In: *IEEE Transactions on Visualization and Computer Graphics* 25.1 (2019), pp. 704–714. issn: 1941-0506. doi: 10.1109/TVCG.2018.2865191 (cit. on p. 24).

[189] T. Igarashi, J.D. Mackinlay, Bay-Wei Chang, and P.T. Zellweger. “Fluid Visualization of Spreadsheet Structures.” In: *Proceedings of the IEEE Symposium on Visual Languages*. VL-98. Los Alamitos, CA, USA: IEEE Computer Society, 1998, pp. 118–125. isbn: 978-0-8186-8712-9. doi: 10.1109/VL.1998.706154 (cit. on p. 25).

[190] Adrian Ilie, Kok-Lim Low, Greg Welch, Anselmo Lastra, Henry Fuchs, and Bruce Cairns. “Combining Head-Mounted and Projector-Based Displays for Surgical Training.” In: *Presence: Teleoperators and Virtual Environments* 13.2 (2004), pp. 128–145. issn: 1054-7460. doi: 10.1162/1054746041382375 (cit. on pp. 44, 49).

[191] Yhonatan Iquiapaza, Jorge Wagner, and Luciana Nedel. “DeAR: Combining Desktop and Augmented Reality for Visual Data Analysis.” In: *Symposium on Virtual and Augmented Reality*. SVR ’23. New York, NY, USA: ACM, 2023, pp. 233–237. isbn: 979-8-4007-0943-2. doi: 10.1145/3625008.3625021 (cit. on pp. 40, 177).

[192] Hiroshi Ishii. “Tangible bits: beyond pixels.” In: *Proceedings of the ACM Conference on Tangible and Embedded Interaction*. TEI08. New York, NY, USA: ACM, 2008, p. xv. isbn: 978-1-60558-004-3. doi: 10.1145/1347390.1347392 (cit. on pp. 33, 48).

[193] Robert J. K. Jacob. “What You Look at Is What You Get: Eye Movement-Based Interaction Techniques.” In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. Vol. 23. CHI ’90. New York, NY, USA: ACM, 1990, pp. 11–18. isbn: 0-201-50932-6. doi: 10.1145/97243.97246 (cit. on pp. 64, 77).

[194] Mikkel R. Jakobsen, Yvonne Jansen, Sebastian Boring, and Kasper Hornbæk. “Should I Stay or Should I Go? Selecting Between Touch and Mid-Air Gestures for Large-Display Interaction.” In: *Human-Computer Interaction – INTERACT 2015*. Ed. by Julio Abascal, Simone Barbosa, Mirko Fetter, Tom Gross, Philippe Palanque, and Marco Winckler. Vol. 9298. New York, NY, USA: Springer Publishing, 2015, pp. 455–473. isbn: 9783319226989. doi: 10.1007/978-3-319-22698-9_31 (cit. on p. 150).

[195] Raphaël James, Anastasia Bezerianos, and Olivier Chapuis. "Evaluating the Extension of Wall Displays with AR for Collaborative Work." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '23. New York, NY, USA: ACM, 2023, pp. 1–17. doi: 10.1145/3544548.3580752 (cit. on pp. 41, 43, 44).

[196] Marium-E Jannat and Khalad Hasan. "Exploring the Effects of Virtually-Augmented Display Sizes on Users' Spatial Memory in Smartwatches." In: *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2023, pp. 553–562. isbn: 979-8-3503-2838-7. doi: 10.1109/ISMAR59233.2023.00070 (cit. on p. 107).

[197] Pascal Jansen, Julian Britten, Alexander Häusele, Thilo Segschneider, Mark Colley, and Enrico Rukzio. "AutoVis: Enabling Mixed-Immersive Analysis of Automotive User Interface Interaction Studies." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '23. New York, NY, USA: ACM, 2023, pp. 1–23. isbn: 978-1-4503-9421-5. doi: 10.1145/3544548.3580760 (cit. on pp. 25, 142).

[198] Yvonne Jansen, Jonas Schjerlund, and Kasper Hornbæk. "Effects of Locomotion and Visual Overview on Spatial Memory When Interacting with Wall Displays." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '19. New York, NY, USA: ACM, 2019, pp. 1–12. isbn: 978-1-4503-5970-2. doi: 10.1145/3290605.3300521 (cit. on p. 86).

[199] Waqas Javed and Niklas Elmqvist. "Exploring the Design Space of Composite Visualization." In: *2012 IEEE Pacific Visualization Symposium*. Piscataway, NJ, USA: IEEE, 2012, pp. 1–8. isbn: 978-1-4673-0866-3. doi: 10.1109/PacificVis.2012.6183556 (cit. on p. 25).

[200] Hans-Christian Jetter, Jens Gerken, Michael Zöllner, Harald Reiterer, and Natasa Milic-Frayling. "Materializing the Query with Facet-Streams: A Hybrid Surface for Collaborative Search on Tabletops." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '11. ACM, 2011, pp. 3013–3022. isbn: 978-1-4503-0228-9. doi: 10.1145/1978942.1979390 (cit. on p. 26).

[201] Hans-Christian Jetter, Svenja Leifert, Jens Gerken, Sören Schubert, and Harald Reiterer. "Does (Multi-)Touch Aid Users' Spatial Memory and Navigation in 'panning' and in 'Zooming & Panning' UIs?" In: *Proceedings of the ACM Conference on Advanced Visual Interfaces*. AVI'12. New York, NY, USA: ACM, 2012, p. 83. isbn: 978-1-4503-1287-5. doi: 10.1145/2254556.2254575 (cit. on pp. 86, 89–91, 93).

[202] Hans-Christian Jetter, Roman Rädle, Tiare Feuchtner, Christoph Anthes, Judith Friedl, and Clemens Nylandsted Klokmose. ""In VR, everything is possible!": Sketching and Simulating Spatially-Aware Interactive Spaces in Virtual Reality." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '20. New York, NY, USA: ACM, 2020, pp. 1–16. isbn: 978-1-4503-6708-0. doi: 10.1145/3313831.3376652 (cit. on pp. 47, 113, 147, 156).

[203] Peter E Jones. "Three-Dimensional Input Device with Six Degrees of Freedom." In: *Mechatronics* 9.7 (1999), pp. 717–729. issn: 0957-4158. doi: 10.1016/S0957-4158(99)00032-X (cit. on p. 150).

[204] Ramsundar Kalpagam Ganesan, Yash K. Rathore, Heather M. Ross, and Heni Ben Amor. “Better Teaming Through Visual Cues: How Projecting Imagery in a Workspace Can Improve Human-Robot Collaboration.” In: *IEEE Robotics & Automation Magazine* 25.2 (2018), pp. 59–71. issn: 1558-223X. doi: 10.1109/MRA.2018.2815655 (cit. on p. 142).

[205] Shaun K. Kane, Daniel Avrahami, Jacob O. Wobbrock, Beverly Harrison, Adam D. Rea, Matthai Philipose, and Anthony LaMarca. “Bonfire: A Nomadic System for Hybrid Laptop-Tabletop Interaction.” In: *Proceedings of the ACM Symposium on User Interface Software and Technology*. UIST ’09. New York, NY, USA: ACM, 2009, pp. 129–138. isbn: 978-1-60558-745-5. doi: 10.1145/1622176.1622202 (cit. on p. 49).

[206] Bonifaz Kaufmann and David Ahlström. “Studying Spatial Memory and Map Navigation Performance on Projector Phones with Peephole Interaction.” In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI ’13. New York, NY, USA: ACM, 2013, pp. 3173–3176. isbn: 978-1-4503-1899-0. doi: 10.1145/2470654.2466434 (cit. on p. 86).

[207] Hiroyuki Kawakita and Toshio Nakagawa. “Augmented TV: An Augmented Reality System for TV Programs beyond the TV Screen.” In: *Proceedings of the IEEE International Conference on Multimedia Computing and Systems*. Piscataway, NJ, USA: IEEE, 2014, pp. 955–960. isbn: 978-1-4799-3824-7. doi: 10.1109/ICMCS.2014.6911158 (cit. on p. 49).

[208] Daniel A. Keim. “Information Visualization and Visual Data Mining.” In: *IEEE Transactions on Visualization and Computer Graphics* 8.1 (2002), pp. 1–8. issn: 1077-2626. doi: 10.1109/2945.981847 (cit. on pp. 62, 125, 152).

[209] Robert S. Kennedy, Norman E. Lane, Kevin S. Berbaum, and Michael G. Lilienthal. “Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness.” In: *The International Journal of Aviation Psychology* 3.3 (1993), pp. 203–220. issn: 1050-8414, 1532-7108. doi: 10.1207/s15327108ijap0303_3 (cit. on pp. 1, 116).

[210] R. Kijima and T. Ojika. “Transition between Virtual Environment and Workstation Environment with Projective Head Mounted Display.” In: *Proceedings of the IEEE Conference on Virtual Reality*. VRAIS-97. Los Alamitos, CA, USA: IEEE Computer Society, 1997, pp. 130–137. isbn: 978-0-8186-7843-1. doi: 10.1109/VRAIS.1997.583062 (cit. on p. 22).

[211] Daekun Kim, Nikhita Joshi, and Daniel Vogel. “Perspective and Geometry Approaches to Mouse Cursor Control in Spatial Augmented Reality.” In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI ’23. New York, NY, USA: ACM, 2023, pp. 1–19. isbn: 978-1-4503-9421-5. doi: 10.1145/3544548.3580849 (cit. on pp. 44, 49, 150).

[212] David Kirby. “The Future Is Now: Diegetic Prototypes and the Role of Popular Films in Generating Real-world Technological Development.” In: *Social Studies of Science* 40.1 (2009), pp. 41–70. issn: 1460-3659. doi: 10.1177/0306312709338325 (cit. on pp. 2, 197).

[213] Akihiro Kiuchi, Taishi Eguchi, Jun Rekimoto, Manabu Tsukada, and Hiroshi Esaki. “ZIGEN: A Windowing System Enabling Multitasking Among 3D and 2D Applications in Immersive Environments.” In: *ACM SIGGRAPH 2022 Posters*. SIGGRAPH ’22. New York, NY, USA: ACM, 2022, pp. 1–2. isbn: 978-1-4503-9361-4. doi: 10.1145/3532719.3543200 (cit. on pp. 149, 152).

[214] Daniel Klinkhammer, Jan Oke Tennie, Paula Erdoes, and Harald Reiterer. “Body Panning: A Movement-Based Navigation Technique for Large Interactive Surfaces.” In: *Proceedings of the ACM Conference on Interactive Tabletops and Surfaces*. ITS ’13. New York, NY, USA: ACM, 2013, pp. 37–40. isbn: 978-1-4503-2271-3. doi: 10.1145/2512349.2512822 (cit. on pp. 86, 89).

[215] Simon Kloiber, Volker Settgast, Christoph Schinko, Martin Weinzerl, Johannes Fritz, Tobias Schreck, and Reinhold Preiner. “Immersive Analysis of User Motion in VR Applications.” In: *The Visual Computer* 36.10–12 (2020), pp. 1937–1949. issn: 1432-2315. doi: 10.1007/s00371-020-01942-1 (cit. on pp. 112–114, 116).

[216] Clemens N. Klokmose, Christian Remy, Janus Bager Kristensen, Rolf Bagge, Michel Beaudouin-Lafon, and Wendy Mackay. “Videostrates: Collaborative, Distributed and Programmable Video Manipulation.” In: *Proceedings of the ACM Symposium on User Interface Software and Technology*. UIST ’19. New York, NY, USA: ACM, 2019, pp. 233–247. isbn: 978-1-4503-6816-2. doi: 10.1145/3332165.3347912 (cit. on p. 141).

[217] Pascal Knierim, Dimitri Hein, Albrecht Schmidt, and Thomas Kosch. “The SmARtphone Controller: Leveraging Smartphones as Input and Output Modality for Improved Interaction within Mobile Augmented Reality Environments.” In: *i-com* 20.1 (2021), pp. 49–61. issn: 1618-162X. doi: 10.1515/icom-2021-0003 (cit. on pp. 31, 49).

[218] Pascal Knierim and Thomas Kosch. “Towards Universal Interaction for Extended Reality.” In: *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2023, pp. 205–207. isbn: 9798350328912. doi: 10.1109/ismar-adjunct60411.2023.00047 (cit. on pp. 4, 178).

[219] Shalva Kohen, Carmine Elvezio, and Steven Feiner. “MiXR: A Hybrid AR Sheet Music Interface for Live Performance.” In: *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2020, pp. 76–77. isbn: 978-1-72817-675-8. doi: 10.1109/ISMAR-Adjunct51615.2020.00035 (cit. on p. 49).

[220] Thomas Kosch, Jakob Karolus, Johannes Zagermann, Harald Reiterer, Albrecht Schmidt, and Paweł W. Woźniak. “A Survey on Measuring Cognitive Workload in Human-Computer Interaction.” In: *ACM Computing Surveys* 55.13s (2023), pp. 1–39. issn: 1557-7341. doi: 10.1145/3582272 (cit. on pp. 162, 182).

[221] Gregory Kramida. “Resolving the Vergence-Accommodation Conflict in Head-Mounted Displays.” In: *IEEE Transactions on Visualization and Computer Graphics* 22.7 (2016), pp. 1912–1931. issn: 1077-2626. doi: 10.1109/TVCG.2015.2473855 (cit. on p. 1).

[222] M. Kraus, N. Weiler, D. Oelke, J. Kehrer, D. A. Keim, and J. Fuchs. "The Impact of Immersion on Cluster Identification Tasks." In: *IEEE Transactions on Visualization and Computer Graphics* 26.1 (2019), pp. 1–1. issn: 2160-9306. doi: 10.1109/TVCG.2019.2934395 (cit. on pp. 24, 62, 112, 116, 140).

[223] Matthias Kraus, Katrin Angerbauer, Juri Buchmüller, Daniel Schweitzer, Daniel A. Keim, Michael Sedlmair, and Johannes Fuchs. "Assessing 2D and 3D Heatmaps for Comparative Analysis: An Empirical Study." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '20. New York, NY, USA: ACM, 2020, pp. 1–14. isbn: 978-1-4503-6708-0. doi: 10.1145/3313831.3376675 (cit. on pp. 24, 112, 140).

[224] Max Krichenbauer, Goshiro Yamamoto, Takafumi Taketomi, Christian Sandor, and Hirokazu Kato. "Towards Augmented Reality User Interfaces in 3D Media Production." In: *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2014, pp. 23–28. isbn: 978-1-4799-6184-9. doi: 10.1109/ISMAR.2014.6948405 (cit. on pp. 2, 152).

[225] Katja Krug, Wolfgang Buschel, Konstantin Klamka, and Raimund Dachselt. "CleAR Sight: Exploring the Potential of Interacting with Transparent Tablets in Augmented Reality." In: *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2022, pp. 196–205. doi: 10.1109/ismar55827.2022.00034 (cit. on pp. 41, 43–45, 47, 49, 80).

[226] Katja Krug, Ricardo Langner, and Konstantin Klamka. "Discussing Facets of Hybrid User Interfaces for the Medical Domain." In: *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2023, pp. 257–260. isbn: 9798350328912. doi: 10.1109/ismar-adjunct60411.2023.00059 (cit. on pp. 36, 180).

[227] Andrew Kurauchi, Wenxin Feng, Carlos Morimoto, and Margrit Betke. "HMAGIC: Head Movement and Gaze Input Cascaded Pointing." In: *Proceedings of the 8th ACM International Conference on PErvasive Technologies Related to Assistive Environments*. PETRA '15. New York, NY, USA: ACM, 2015, pp. 1–4. isbn: 978-1-4503-3452-5. doi: 10.1145/2769493.2769550 (cit. on p. 167).

[228] Ricardo Langner, Tom Horak, and Raimund Dachselt. "VisTiles: Coordinating and Combining Co-Located Mobile Devices for Visual Data Exploration." In: *IEEE Transactions on Visualization and Computer Graphics* 24.1 (2018), pp. 626–636. issn: 10772626. doi: 10.1109/TVCG.2017.2744019 (cit. on pp. 60, 67, 72, 77).

[229] Ricardo Langner, Marc Satkowski, Wolfgang Büschel, and Raimund Dachselt. "MARVIS: Combining Mobile Devices and Augmented Reality for Visual Data Analysis." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '21. New York, NY, USA: ACM, 2021, 468:1–468:17. isbn: 978-1-4503-8096-6. doi: 10.1145/3411764.3445593 (cit. on pp. 41–44, 48, 50, 105).

[230] Joseph J. LaViola, Ernst Kruijff, Ryan P. McMahan, Doug A. Bowman, and Ivan Poupyrev. *3D User Interfaces: Theory and Practice*. Second edition. Addison-Wesley Usability and HCI Series. Boston, USA: Addison-Wesley, 2017. isbn: 978-0-13-403432-4 (cit. on pp. 60, 150).

[231] Huy Viet Le, Sven Mayer, and Niels Henze. "InfiniTouch: Finger-Aware Interaction on Fully Touch Sensitive Smartphones." In: *Proceedings of the ACM Symposium on User Interface Software and Technology*. UIST '18. New York, NY, USA: ACM, 2018, pp. 779–792. ISBN: 978-1-4503-5948-1. DOI: 10.1145/3242587.3242605 (cit. on p. 65).

[232] David Ledo, Steven Houben, Jo Vermeulen, Nicolai Marquardt, Lora Oehlberg, and Saul Greenberg. "Evaluation Strategies for HCI Toolkit Research." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '18. New York, NY, USA: ACM, 2018, 36:1–36:17. ISBN: 978-1-4503-5620-6. DOI: 10.1145/3173574.3173610 (cit. on pp. 37, 51).

[233] B. Lee, D. Brown, B. Lee, C. Hurter, S. Drucker, and T. Dwyer. "Data Visceralization: Enabling Deeper Understanding of Data Using Virtual Reality." In: *IEEE Transactions on Visualization and Computer Graphics* 27.2 (2020), pp. 1095–1105. ISSN: 2160-9306. DOI: 10.1109/TVCG.2020.3030435 (cit. on p. 116).

[234] Benjamin Lee, Maxime Cordeil, Arnaud Prouzeau, Bernhard Jenny, and Tim Dwyer. "A Design Space For Data Visualisation Transformations Between 2D And 3D In Mixed-Reality Environments." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '22. New York, NY, USA: ACM, 2022, pp. 1–14. ISBN: 978-1-4503-9157-3. DOI: 10.1145/3491102.3501859 (cit. on pp. 21, 152, 177, 183, 184).

[235] Benjamin Lee, Xiaoyun Hu, Maxime Cordeil, Arnaud Prouzeau, Bernhard Jenny, and Tim Dwyer. "Shared Surfaces and Spaces: Collaborative Data Visualisation in a Co-located Immersive Environment." In: *IEEE Transactions on Visualization and Computer Graphics* 27.2 (2021), pp. 1171–1181. ISSN: 2160-9306. DOI: 10.1109/TVCG.2020.3030450. arXiv: 2009.00050 (cit. on p. 78).

[236] Benjamin Lee, Arvind Satyanarayan, Maxime Cordeil, Arnaud Prouzeau, Bernhard Jenny, and Tim Dwyer. "Deimos: A Grammar of Dynamic Embodied Immersive Visualisation Morphs and Transitions." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '23. New York, NY, USA: ACM, 2023, pp. 1–18. ISBN: 978-1-4503-9421-5. DOI: 10.1145/3544548.3580754 (cit. on pp. 177, 183, 184).

[237] Bongshin Lee, Eun Kyoung Choe, Petra Isenberg, Kim Marriott, and John Stasko. "Reaching Broader Audiences With Data Visualization." In: *IEEE Computer Graphics and Applications* 40.2 (2020), pp. 82–90. ISSN: 1558-1756. DOI: 10.1109/MCG.2020.2968244 (cit. on p. 60).

[238] Bongshin Lee, Arjun Srinivasan, John Stasko, Melanie Tory, and Vidya Setlur. "Multimodal Interaction for Data Visualization." In: *Proceedings of the ACM Conference on Advanced Visual Interfaces*. Proceedings of the 2018 International Conference on Advanced Visual Interfaces - AVI '18. ACM, 2018, pp. 1–3. DOI: 10.1145/3206505.3206602 (cit. on p. 60).

[239] Byungjoo Lee, Qiao Deng, Eve Hoggan, and Antti Oulasvirta. "Boxer: A Multimodal Collision Technique for Virtual Objects." In: *Proceedings of the 19th ACM International Conference on Multimodal Interaction - ICMI 2017*. ICMI '17. New York, NY, USA: ACM, 2017, pp. 252–260. ISBN: 978-1-4503-5543-8. DOI: 10.1145/3136755.3136761 (cit. on p. 3).

[240] Sangyoon Lee, Sunghoon Yim, Gerard Jounghyun Kim, Ungyeon Yang, and Chang-Hun Kim. “Design and Evaluation of a Hybrid Display System for Motion-Following Tasks.” In: *Proceedings of the ACM Conference on Human-Computer Interaction with Mobile Devices and Services*. Vol. 4563. Lecture Notes in Computer Science. New York, NY, USA: Springer Publishing, 2007, pp. 280–289. isbn: 9783540733355. doi: 10.1007/978-3-540-73335-5_31 (cit. on p. 44).

[241] Svenja Leifert. “The Influence of Grids on Spatial and Content Memory.” In: *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems*. CHI ’11. New York, NY, USA: ACM, 2011, p. 941. isbn: 978-1-4503-0268-5. doi: 10.1145/1979742.1979522 (cit. on pp. 86, 89, 91).

[242] Yi Li, Robert Crowther, Jim Smiley, Tim Dwyer, Benjamin Tag, Pourang Irani, and Barrett Ens. “Revisiting Consumed Endurance: A NICE Way to Quantify Shoulder Fatigue in Virtual Reality.” In: *Proceedings of the IEEE Symposium on Virtual Reality Software & Technology*. VRST 2023. New York, NY, USA: ACM, 2023, pp. 101–125. isbn: 979-8-4007-0328-7. doi: 10.1145/3611659.3615702 (cit. on p. 3).

[243] Yi Li, Benjamin Tag, Shaozhang Dai, Robert Crowther, Tim Dwyer, Pourang Irani, and Barrett Ens. “NICER: A New and Improved Consumed Endurance and Recovery Metric to Quantify Muscle Fatigue of Mid-Air Interactions.” In: *ACM Transactions on Graphics* 43.4 (2024), pp. 304–324. issn: 0730-0301, 1557-7368. doi: 10.1145/3658230 (cit. on p. 3).

[244] Shuqi Liao, Sparsh Chaudhri, Maanas K. Karwa, and Voicu Popescu. “SeamlessVR: Bridging the Immersive to Non-Immersive Visualization Divide.” In: *IEEE Transactions on Visualization and Computer Graphics* 31.5 (2025), pp. 2806–2816. issn: 1077-2626, 1941-0506, 2160-9306. doi: 10.1109/TVCG.2025.3549564 (cit. on pp. 21, 152, 179).

[245] Klemen Lilija, Henning Pohl, and Kasper Hornbæk. “Who Put That There? Temporal Navigation of Spatial Recordings by Direct Manipulation.” In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI ’20. New York, NY, USA: ACM, 2020, pp. 1–11. isbn: 978-1-4503-6708-0. doi: 10.1145/3313831.3376604 (cit. on pp. 113, 115, 116, 120, 124).

[246] Jiazhou Liu, Arnaud Prouzeau, Barrett Ens, and Tim Dwyer. “Effects of Display Layout on Spatial Memory for Immersive Environments.” In: *Proceedings of the ACM on Human-Computer Interaction* 6.ISS (2022), pp. 468–488. issn: 2573-0142. doi: 10.1145/3567729 (cit. on p. 105).

[247] Jiazhou Liu, Kadek Ananta Satriadi, Barrett Ens, and Tim Dwyer. “Investigating the Effects of Physical Landmarks on Spatial Memory for Information Visualisation in Augmented Reality.” In: *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2024, pp. 289–298. isbn: 979-8-3315-1647-5. doi: 10.1109/ISMAR62088.2024.00043 (cit. on p. 86).

[248] Xinze Liu, Yubo Zhang, Ziqian Yan, and Yan Ge. “Defining ‘seamlessly connected’: user perceptions of operation latency in cross-device interaction.” In: *International Journal of Human-Computer Studies* 177 (2023), p. 103068. issn: 1071-5819. doi: 10.1016/j.ijhcs.2023.103068 (cit. on p. 177).

[249] Yunxing Liu and Jean-Bernard Martens. “Conversation-Based Hybrid User Interface for Structured Qualitative Survey: A Pilot Study Using Repertory Grid.” In: *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems*. CHI ’22. New York, NY, USA: ACM, 2022, pp. 1–8. isbn: 978-1-4503-9156-6. doi: 10.1145/3491101.3519690 (cit. on p. 33).

[250] Pedro Lopes, Alexandra Ion, Sijing You, and Patrick Baudisch. “Adding Force Feedback to Mixed Reality Experiences and Games Using Electrical Muscle Stimulation.” In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI ’18. ACM, 2018, pp. 201–213. doi: 10.1145/3173574.3174020 (cit. on p. 3).

[251] Miguel L. Lourenço, Rui A. Pitarma, and Denis A. Coelho. “A Design Contribution to Ergonomic PC Mice Development.” In: *International Journal of Environmental Research and Public Health* 19.13 (2022), p. 8126. issn: 1660-4601. doi: 10.3390/ijerph19138126 (cit. on pp. 149, 155, 163).

[252] Yiqin Lu, Chun Yu, Xin Yi, Yuanchun Shi, and Shengdong Zhao. “BlindType: Eyes-Free Text Entry on Handheld Touchpad by Leveraging Thumb’s Muscle Memory.” In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1.2 (2017), pp. 1–24. issn: 24749567. doi: 10.1145/3090083 (cit. on p. 61).

[253] Mille Skovhus Lunding, Jens Emil Sloth Grønbæk, Karl-Emil Kjær Bilstrup, Marie-Louise Stisen Kjerstein Sørensen, and Marianne Graves Petersen. “ExposAR: Bringing Augmented Reality to the Computational Thinking Agenda through a Collaborative Authoring Tool.” In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI ’22. New York, NY, USA: ACM, 2022, pp. 1–14. isbn: 978-1-4503-9157-3. doi: 10.1145/3491102.3517636 (cit. on pp. 37, 47).

[254] Rasmus Lunding, **Sebastian Hubenschmid**, and Tiare Feuchtner. “Proposing a Hybrid Authoring Interface for AR-Supported Human-Robot Collaboration.” In: *7th International Workshop on Virtual, Augmented, and Mixed-Reality for Human-Robot Interactions*. 2024 (cit. on pp. x, 180).

[255] Rasmus Lunding, **Sebastian Hubenschmid**, Tiare Feuchtner, and Kaj Grønbæk. “ARTHUR: Authoring Human–Robot Collaboration Processes with Augmented Reality Using Hybrid User Interfaces.” In: *Virtual Reality* 29.2 (2025), p. 73. issn: 1434-9957. doi: 10.1007/s10055-025-01149-6 (cit. on pp. vii, 110, 142).

[256] Rasmus S. Lunding, Mathias N. Lystbæk, Tiare Feuchtner, and Kaj Grønbæk. “AR-supported Human-Robot Collaboration: Facilitating Workspace Awareness and Parallelized Assembly Tasks.” In: *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. New York, NY, USA: IEEE, 2023, pp. 1064–1073. isbn: 979-8-3503-2838-7. doi: 10.1109/ISMAR59233.2023.00123 (cit. on p. 142).

[257] Rasmus Skovhus Lunding, Tiare Feuchtner, and Kaj Grønbæk. “Investigating AR Assistance for Human-Robot Collaboration in Mould Assembly “in the Wild”.” In: *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction*. Piscataway, NJ, USA: IEEE, 2025, pp. 540–549. isbn: 979-8-3503-7893-1. doi: 10.1109/hri61500.2025.10973920 (cit. on p. 143).

[258] Rasmus Skovhus Lunding, Mille Skovhus Lunding, Tiare Feuchtner, Marianne Graves Petersen, Kaj Grønbæk, and Ryo Suzuki. "RoboVisAR: Immersive Authoring of Condition-based AR Robot Visualisations." In: *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction*. HRI '24. New York, NY, USA: ACM, 2024, pp. 462–471. ISBN: 979-8-4007-0322-5. DOI: 10.1145/3610977.3634972 (cit. on p. 142).

[259] Weizhou Luo, Eva Goebel, Patrick Reipschlager, Mats Ole Ellenberg, and Raimund Dachselt. "Exploring and Slicing Volumetric Medical Data in Augmented Reality Using a Spatially-Aware Mobile Device." In: *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2021, pp. 334–339. ISBN: 978-1-66541-298-8. DOI: 10.1109/ismar-adjunct54149.2021.00076 (cit. on pp. 41, 44, 45, 48, 49).

[260] Tiffany Luong, Yi Fei Cheng, Max Möbus, Andreas Fender, and Christian Holz. "Controllers or Bare Hands? A Controlled Evaluation of Input Techniques on Interaction Performance and Exertion in Virtual Reality." In: *IEEE Transactions on Visualization and Computer Graphics* 29.11 (2023), pp. 4633–4643. ISSN: 2160-9306. DOI: 10.1109/TVCG.2023.3320211 (cit. on pp. 147, 148, 150).

[261] Wendy E. Mackay and Anne-Laure Fayard. "HCI, Natural Science and Design: A Framework for Triangulation across Disciplines." In: *Proceedings of the Conference on Designing Interactive Systems Processes, Practices, Methods, and Techniques - DIS '97*. DIS '97. New York, NY, USA: ACM, 1997, pp. 223–234. ISBN: 978-0-89791-863-3. DOI: 10.1145/263552.263612 (cit. on p. 9).

[262] I. Scott MacKenzie. "Fitts' Law as a Research and Design Tool in Human-Computer Interaction." In: *Human–Computer Interaction* 7.1 (1992), pp. 91–139. ISSN: 0737-0024, 1532-7051. DOI: 10.1207/s15327051hci0701_3 (cit. on p. 167).

[263] I. Scott MacKenzie, R. William Soukoreff, and Chris Pal. "A Two-Ball Mouse Affords Three Degrees of Freedom." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '97. New York, NY, USA: ACM, 1997, p. 303. ISBN: 978-0-89791-926-5. DOI: 10.1145/1120212.1120405 (cit. on p. 151).

[264] Ali Ahmad Malik and Vijitashwa Pandey. "Drive the Cobots Aright: Guidelines for Industrial Application of Cobots." In: *27th Design for Manufacturing and the Life Cycle Conference (DFMLC)*. IDETC-CIE2022. St. Louis, Missouri, USA: American Society of Mechanical Engineers, 2022, V005T05A015. ISBN: 978-0-7918-8625-0. DOI: 10.1115/DETC2022-90777 (cit. on p. 142).

[265] Veera Bhadra Harish Mandalika, Alexander I. Chernoglazov, Mark Billinghurst, Christoph Bartneck, Michael A. Hurrell, Niels De Ruiter, Anthony P. H. Butler, and Philip H. Butler. "A Hybrid 2D/3D User Interface for Radiological Diagnosis." In: *Journal of Digital Imaging* 31.1 (2018), pp. 56–73. ISSN: 1618-727X. DOI: 10.1007/s10278-017-0002-6 (cit. on p. 150).

[266] Federico Manuri and Giovanni Piumatti. "A Preliminary Study of a Hybrid User Interface for Augmented Reality Applications." In: *Proceedings of the 7th International Conference on Intelligent Technologies for Interactive Entertainment*. INTETAIN. Piscataway, NJ, USA: IEEE, 2015. ISBN: 978-1-63190-061-7. DOI: 10.4108/icst.intetain.2015.259629 (cit. on p. 37).

[267] Michael R. Marner, Ross T. Smith, and Bruce H. Thomas. "Mapping 2D Input to 3D Immersive Spatial Augmented Reality." In: *Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces*. Piscataway, NJ, USA: IEEE, 2015, pp. 171–172. ISBN: 978-1-4673-6886-5. DOI: 10.1109/3DUI.2015.7131755 (cit. on p. 150).

[268] Nicolai Marquardt, Frederico Schardong, and Anthony Tang. "EXCITE: EXploring Collaborative Interaction in Tracked Environments." In: *Human-Computer Interaction – INTERACT 2015*. Ed. by Julio Abascal, Simone Barbosa, Mirko Fetter, Tom Gross, Philippe Palanque, and Marco Winckler. Vol. 9297. New York, NY, USA: Springer Publishing, 2015, pp. 89–97. ISBN: 9783319226682. DOI: 10.1007/978-3-319-22668-2_8 (cit. on p. 115).

[269] Elisabeth Mayer, Francesco Chiossi, and Sven Mayer. "Crossing Mixed Realities: A Review for Transitional Interfaces Design." In: *Proceedings of Mensch und Computer*. MuC '24. New York, NY, USA: ACM, 2024, pp. 629–634. ISBN: 9798400709982. DOI: 10.1145/3670653.3677481 (cit. on pp. 20–22).

[270] Mark McGill, Daniel Boland, Roderick Murray-Smith, and Stephen Brewster. "A Dose of Reality: Overcoming Usability Challenges in VR Head-Mounted Displays." In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI '15. New York, NY, USA: ACM, 2015, pp. 2143–2152. ISBN: 978-1-4503-3145-6. DOI: 10.1145/2702123.2702382 (cit. on pp. 20, 150).

[271] Sumit Mehra, Peter Werkhoven, and Marcel Worring. "Navigating on handheld displays: Dynamic versus static peephole navigation." In: *ACM Transactions on Computer-Human Interaction* 13.4 (2006), pp. 448–457. ISSN: 1557-7325. DOI: 10.1145/1188816.1188818 (cit. on pp. 41, 86, 89, 104).

[272] D. Mendes, F. M. Caputo, A. Giachetti, A. Ferreira, and J. Jorge. "A Survey on 3D Virtual Object Manipulation: From the Desktop to Immersive Virtual Environments." In: *Computer Graphics Forum* 38.1 (2019), pp. 21–45. ISSN: 1467-8659. DOI: 10.1111/cgf.13390 (cit. on p. 150).

[273] Julián Méndez, Marc Satkowski, and Rufat Rzayev. "How Does Explainability Look in Hybrid User Interfaces?" In: *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2023, pp. 251–256. DOI: 10.1109/ismar-adjunct60411.2023.00058 (cit. on p. 180).

[274] Leonel Merino, Magdalena Schwarzl, Matthias Kraus, Michael Sedlmair, Dieter Schmalstieg, and Daniel Weiskopf. "Evaluating Mixed and Augmented Reality: A Systematic Literature Review (2009-2019)." In: *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. IEEE, 2020, pp. 438–451. DOI: 10.1109/ISMAR50242.2020.00069 (cit. on pp. 9, 114).

[275] Adriano Viegas Milani, Mario Gutierrez, Ronan Boulic, Mark McGill, and Aidan Kehoe. "Comparison of Mouse, 6DoF Controller and Pen Interactions on a Glasses-Free 3D Display." In: *2024 IEEE Gaming, Entertainment, and Media Conference (GEM)*. Piscataway, NJ, USA: IEEE, 2024, pp. 1–6. ISBN: 979-8-3503-7453-7. DOI: 10.1109/GEM61861.2024.10585578 (cit. on p. 151).

[276] Paul Milgram and Fumio Kishino. "A Taxonomy of Mixed Reality Visual Displays." In: *IEICE TRANSACTIONS on Information and Systems* 77.12 (1994), pp. 1321–1329. ISSN: 0916-8532 (cit. on pp. 1, 19, 111, 113).

[277] K.Jarrod Millman and Fernando Pérez. "Developing Open-Source Scientific Practice." In: *Implementing Reproducible Research*. Ed. by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. 1st ed. Chapman and Hall/CRC, 2018, pp. 149–183. ISBN: 978-1-315-37346-1. DOI: 10.1201/9781315373461-6 (cit. on pp. 115, 121).

[278] Mark Mine. "Working in a virtual world: Interaction techniques used in the chapel hill immersive modeling program." In: *University of North Carolina* 1.1 (1996), p. 1 (cit. on p. 152).

[279] Mark Mine, Arun Yoganandan, and Dane Coffey. "Making VR Work: Building a Real-World Immersive Modeling Application in the Virtual World." In: *Proceedings of the ACM Symposium on Spatial User Interfaces*. SUI '14. New York, NY, USA: ACM, 2014, pp. 80–89. ISBN: 978-1-4503-2820-3. DOI: 10.1145/2659766.2659780 (cit. on p. 152).

[280] Meredith Ringel Morris, Andreea Danielescu, Steven Drucker, Danyel Fisher, Bongshin Lee, M. C. Schraefel, and Jacob O. Wobbrock. "Reducing Legacy Bias in Gesture Elicitation Studies." In: *Interactions* 21.3 (2014), pp. 40–45. ISSN: 1072-5520, 1558-3449. DOI: 10.1145/2591689 (cit. on pp. 4, 26).

[281] Suzanne Mueller, Andreas Dippon, and Gudrun Klinker. "Capture The Flag Demo: Engaging In A Multi-Device Augmented Reality Game." In: *Proceedings of the ACM Conference on Interactive Tabletops and Surfaces*. ITS '15. New York, NY, USA: ACM, 2015, pp. 393–396. ISBN: 978-1-4503-3899-8. DOI: 10.1145/2817721.2823475 (cit. on pp. 43, 49).

[282] Florian Müller, Martin Schmitz, Daniel Schmitt, Sebastian Günther, Markus Funk, and Max Mühlhäuser. "Walk The Line: Leveraging Lateral Shifts of the Walking Path as an Input Modality for Head-Mounted Displays." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '20. New York, NY, USA: ACM, 2020, pp. 1–15. ISBN: 978-1-4503-6708-0. DOI: 10.1145/3313831.3376852 (cit. on p. 114).

[283] Jens Müller, Roman Rädle, Hans-Christian Jetter, and Harald Reiterer. "An Experimental Comparison of Vertical and Horizontal Dynamic Peephole Navigation." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '15. New York, NY, USA: ACM, 2015, pp. 1523–1526. ISBN: 978-1-4503-3145-6. DOI: 10.1145/2702123.2702227 (cit. on pp. 86, 89).

[284] Jens Müller, Roman Rädle, and Harald Reiterer. "Remote Collaboration With Mixed Reality Displays: How Shared Virtual Landmarks Facilitate Spatial Referencing." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '17. New York, NY, USA: ACM, 2017, pp. 6481–6486. ISBN: 978-1-4503-4655-9. DOI: 10.1145/3025453.3025717 (cit. on p. 139).

[285] Jens Müller, Roman Rädle, and Harald Reiterer. "Virtual Objects as Spatial Cues in Collaborative Mixed Reality Environments: How They Shape Communication Behavior and User Task Load." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI'16. New York, NY, USA: ACM, 2016, pp. 1245–1249. ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858043 (cit. on p. 89).

[286] Jens Müller, Johannes Zagermann, Jonathan Wieland, Ulrike Pfeil, and Harald Reiterer. "A Qualitative Comparison Between Augmented and Virtual Reality Collaboration with Handheld Devices." In: *Proceedings of Mensch Und Computer*. MuC'19. New York, NY, USA: ACM, 2019, pp. 399–410. ISBN: 978-1-4503-7198-8. DOI: 10.1145/3340764.3340773 (cit. on pp. 86, 89, 117).

[287] Tamara Munzner. *Visualization Analysis and Design*. A K Peters/CRC Press, 2014. ISBN: 9780429088902. DOI: 10.1201/b17511 (cit. on pp. 25, 79).

[288] Miguel A. Nacenta, Regan L. Mandryk, and Carl Gutwin. "Targeting across Displayless Space." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '08. New York, NY, USA: ACM, 2008, pp. 777–786. ISBN: 978-1-60558-011-1. DOI: 10.1145/1357054.1357178 (cit. on pp. 87, 102).

[289] Michael Nebeling, Maximilian Speicher, Xizi Wang, Shwetha Rajaram, Brian D. Hall, Zijian Xie, Alexander R. E. Raistrick, Michelle Aebersold, Edward G. Happ, Jiayin Wang, Yanan Sun, Lotus Zhang, Leah E. Ramsier, and Rhea Kulkarni. "MRAT: The Mixed Reality Analytics Toolkit." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '20. New York, NY, USA: ACM, 2020, pp. 1–12. ISBN: 978-1-4503-6708-0. DOI: 10.1145/3313831.3376330 (cit. on pp. 25, 41, 44, 45, 48, 49, 112–116, 119, 122, 124, 126, 182, 183).

[290] Gregory M. Nielson and Dan R. Olsen. "Direct Manipulation Techniques for 3D Objects Using 2D Locator Devices." In: *Proceedings of the 1986 Workshop on Interactive 3D Graphics - SI3D '86*. SI3D '86. New York, NY, USA: ACM, 1987, pp. 175–182. ISBN: 978-0-89791-228-0. DOI: 10.1145/319120.319134 (cit. on p. 150).

[291] Donald A. Norman and Stephen W. Draper. *User Centered System Design; New Perspectives on Human-Computer Interaction*. USA: L. Erlbaum Associates Inc., 1986. ISBN: 0898597811 (cit. on p. 26).

[292] Erwan Normand and Michael J. McGuffin. "Enlarging a Smartphone with AR to Create a Handheld VESAD (Virtually Extended Screen-Aligned Display)." In: *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2018, pp. 123–133. ISBN: 978-1-5386-7459-8. DOI: 10.1109/ismar.2018.00043 (cit. on pp. 41, 47, 50–52, 87, 91, 102, 104, 105, 175).

[293] Kenton O'Hara, Abigail Sellen, and Richard Bentley. "Supporting Memory for Spatial Location While Reading from Small Displays." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '99. New York, NY, USA: ACM, 1999, p. 220. ISBN: 978-1-58113-158-1. DOI: 10.1145/632716.632853 (cit. on p. 86).

[294] Ian Oakley and Junseok Park. "Motion Marking Menus: An Eyes-Free Approach to Motion Input for Handheld Devices." In: *International Journal of Human-Computer Studies* 67.6 (2009), pp. 515–532. ISSN: 10715819. DOI: 10.1016/j.ijhcs.2009.02.002 (cit. on pp. 61, 65, 75).

[295] Dan Odell and Peter Johnson. "Evaluation of Flat, Angled, and Vertical Computer Mice and Their Effects on Wrist Posture, Pointing Performance, and Preference." In: *Work* 52.2 (2015). Ed. by Thomas J. Albin, pp. 245–253. ISSN: 1875-9270. DOI: 10.3233/WOR-152167 (cit. on pp. 149, 155).

[296] Eyal Ofek, Jens Grubert, Michel Pahud, Mark Phillips, and Per Ola Kristensson. "Towards a Practical Virtual Office for Mobile Knowledge Workers." In: (2020). DOI: 10.48550/ARXIV.2009.02947 (cit. on p. 156).

[297] Patrick Aggergaard Olin, Ahmad Mohammad Issa, Tiare Feuchtner, and Kaj Grønbæk. "Designing for Heterogeneous Cross-Device Collaboration and Social Interaction in Virtual Reality." In: *32nd Australian Conference on Human-Computer Interaction*. OzCHI '20. New York, NY, USA: ACM, 2020, pp. 112–127. ISBN: 978-1-4503-8975-4. DOI: 10.1145/3441000.3441070 (cit. on pp. 5, 20, 37, 47).

[298] Antti Oulasvirta and Kasper Hornbæk. "Counterfactual Thinking: What Theories Do in Design." In: *International Journal of Human–Computer Interaction* 38.1 (2022), pp. 78–92. ISSN: 1532-7590. DOI: 10.1080/10447318.2021.1925436 (cit. on pp. 9, 181).

[299] Antti Oulasvirta and Kasper Hornbæk. "HCI Research as Problem-Solving." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI'16. New York, NY, USA: ACM, 2016, pp. 4956–4967. ISBN: 9781450333627. DOI: 10.1145/2858036.2858283 (cit. on p. 181).

[300] Sharon Oviatt. "Ten Myths of Multimodal Interaction." In: *Communications of the ACM* 42.11 (1999), pp. 74–81. ISSN: 1557-7317. DOI: 10.1145/319382.319398 (cit. on p. 60).

[301] Sharon Oviatt and Philip Cohen. "Perceptual User Interfaces: Multimodal Interfaces That Process What Comes Naturally." In: *Communications of the ACM* 43.3 (2000), pp. 45–53. ISSN: 1557-7317. DOI: 10.1145/330534.330538 (cit. on p. 60).

[302] Sharon Oviatt, Rebecca Lunsford, and Rachel Coulston. "Individual Differences in Multimodal Integration Patterns: What Are They and Why Do They Exist?" In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '05. ACM, 2005, pp. 241–249. DOI: 10.1145/1054972.1055006 (cit. on p. 60).

[303] Jamal R Paden, Arpit Narechania, and Alex Endert. "BiasBuzz: Combining Visual Guidance with Haptic Feedback to Increase Awareness of Analytic Behavior during Visual Data Analysis." In: *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems*. CHI '24. New York, NY, USA: ACM, 2024, pp. 1–7. ISBN: 979-8-4007-0331-7. DOI: 10.1145/3613905.3651064 (cit. on p. 168).

[304] Matthew J. Page, Joanne E. McKenzie, Patrick M. Bossuyt, Isabelle Boutron, Tammy C. Hoffmann, Cynthia D. Mulrow, Larissa Shamseer, Jennifer M. Tetzlaff, Elie A. Akl, Sue E. Brennan, Roger Chou, Julie Glanville, Jeremy M. Grimshaw, Asbjørn Hróbjartsson, Manoj M. Lalu, Tianjing Li, Elizabeth W. Loder, Evan Mayo-Wilson, Steve McDonald, Luke A. McGuinness, Lesley A. Stewart, James Thomas, Andrea C. Tricco, Vivian A. Welch, Penny Whiting, and David Moher. "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews." In: *Systematic Reviews* 10.1 (2021), p. 89. ISSN: 2046-4053. DOI: 10.1186/s13643-021-01626-4 (cit. on p. 34).

[305] Michel Pahud, Ken Hinckley, Shamsi Iqbal, Abigail Sellen, and Bill Buxton. "Toward Compound Navigation Tasks on Mobiles via Spatial Manipulation." In: *Proceedings of the ACM Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI '13. New York, NY, USA: ACM, 2013, pp. 113–122. isbn: 978-1-4503-2273-7. doi: 10.1145/2493190.2493210 (cit. on pp. 86, 104).

[306] Federica Pallavicini and Alessandro Pepe. "Comparing Player Experience in Video Games Played in Virtual Reality or on Desktop Displays: Immersion, Flow, and Positive Emotions." In: *Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*. CHI PLAY '19. New York, NY, USA: ACM, 2019, pp. 195–210. isbn: 978-1-4503-6871-1. doi: 10.1145/3341215.3355736 (cit. on p. 112).

[307] Eduardo G. Q. Palmeira, Alexandre Campos, Ígor A. Moraes, Alexandre G. De Siqueira, and Marcelo G. G. Ferreira. "Quantifying the 'Gorilla Arm' Effect in a Virtual Reality Text Entry Task via Ray-Casting: A Preliminary Single-Subject Study." In: *Symposium on Virtual and Augmented Reality*. SVR '23. New York, NY, USA: ACM, 2023, pp. 274–278. isbn: 979-8-4007-0943-2. doi: 10.1145/3625008.3625046 (cit. on pp. 3, 4, 152, 168).

[308] Max Pascher, Uwe Gruenefeld, Stefan Schneegass, and Jens Gerken. "How to Communicate Robot Motion Intent: A Scoping Review." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '23. New York, NY, USA: ACM, 2023, pp. 1–17. isbn: 978-1-4503-9421-5. doi: 10.1145/3544548.3580857 (cit. on p. 142).

[309] Leonardo Pavanatto, Feiyu Lu, Chris North, and Doug A. Bowman. "Multiple Monitors or Single Canvas? Evaluating Window Management and Layout Strategies on Virtual Displays." In: *IEEE Transactions on Visualization and Computer Graphics* 31.3 (2025), pp. 1713–1730. issn: 2160-9306. doi: 10.1109/tvcg.2024.3368930 (cit. on pp. 41, 49).

[310] Leonardo Pavanatto, Chris North, Doug A. Bowman, Carmen Badea, and Richard Stoakley. "Do we still need physical monitors? An evaluation of the usability of AR virtual monitors for productivity work." In: *Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces*. Piscataway, NJ, USA: IEEE, 2021, pp. 759–767. isbn: 978-1-66541-838-6. doi: 10.1109/vr50410.2021.00103 (cit. on pp. 1, 41, 44, 49, 116, 147, 149, 165).

[311] Fernando Perez and Brian E. Granger. "IPython: A System for Interactive Scientific Computing." In: *Computing in Science & Engineering* 9.3 (2007), pp. 21–29. issn: 1521-9615. doi: 10.1109/MCSE.2007.53 (cit. on pp. 115, 116, 121, 140).

[312] Simon T. Perrault, Eric Lecolinet, James Eagan, and Yves Guiard. "Watchit: Simple Gestures and Eyes-Free Interaction for Wristwatches and Bracelets." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '13. New York, NY, USA: ACM, 2013, pp. 1451–1460. isbn: 978-1-4503-1899-0. doi: 10.1145/2470654.2466192 (cit. on p. 61).

[313] Ken Pfeuffer, Benedikt Mayer, Diako Mardanbegi, and Hans Gellersen. "Gaze + Pinch Interaction in Virtual Reality." In: *Proceedings of the ACM Symposium on Spatial User Interfaces*. SUI '17. New York, NY, USA: ACM, 2017, pp. 99–108. isbn: 978-1-4503-5486-8. doi: 10.1145/3131277.3132180 (cit. on pp. 2, 150).

[314] Thomas Plank, Hans-christian Jetter, Roman Rädle, Clemens N Klokmose, Thomas Luger, and Harald Reiterer. "Is Two Enough?! Studying Benefits, Barriers, and Biases of Multi-Tablet Use for Collaborative Visualization." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '17. New York, NY, USA: ACM, 2017, pp. 4548–4560. ISBN: 978-1-4503-4655-9. DOI: 10.1145/3025453.3025537 (cit. on pp. 4, 22, 23, 103).

[315] Carole Plasson, Renaud Blanch, and Laurence Nigay. "Selection Techniques for 3D Extended Desktop Workstation with AR HMD." In: *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2022, pp. 460–469. ISBN: 978-1-66545-325-7. DOI: 10.1109/ISMAR55827.2022.00062 (cit. on pp. 150, 156, 162, 164).

[316] Michael Prilla and Lisa M. Rühmann. "An Analysis Tool for Cooperative Mixed Reality Scenarios." In: *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. IEEE, 2018, pp. 31–35. DOI: 10.1109/ISMAR-Adjunct.2018.00026 (cit. on pp. 115, 116, 122, 124, 126).

[317] Kathrin Probst. "Peripheral Interaction in Desktop Computing: Why It's Worth Stepping Beyond Traditional Mouse and Keyboard." In: *Peripheral Interaction*. Ed. by Saskia Bakker, Doris Hausen, and Ted Selker. New York, NY, USA: Springer Publishing, 2016, pp. 183–205. ISBN: 9783319295237. DOI: 10.1007/978-3-319-29523-7_9 (cit. on p. 86).

[318] Arnaud Prouzeau, Antoine Lhuillier, Barrett Ens, Daniel Weiskopf, and Tim Dwyer. "Visual Link Routing in Immersive Visualisations." In: *Proceedings of the ACM Conference on Interactive Surfaces and Spaces*. ISS '19. New York, NY, USA: ACM, 2019, pp. 241–253. ISBN: 978-1-4503-6891-9. DOI: 10.1145/3343055.3359709 (cit. on p. 24).

[319] Paulo R.V. Quemelo and Edgar Ramos Vieira. "Biomechanics and Performance When Using a Standard and a Vertical Computer Mouse." In: *Ergonomics* 56.8 (2013), pp. 1336–1344. ISSN: 1366-5847. DOI: 10.1080/00140139.2013.805251 (cit. on pp. 149, 155).

[320] Roman Rädle, Hans-Christian Jetter, Simon Butscher, and Harald Reiterer. "The Effect of Egocentric Body Movements on Users' Navigation Performance and Spatial Memory in Zoomable User Interfaces." In: *Proceedings of the ACM Conference on Interactive Tabletops and Surfaces*. ITS '13. New York, NY, USA: ACM, 2013, pp. 23–32. ISBN: 978-1-4503-2271-3. DOI: 10.1145/2512349.2512811 (cit. on pp. 86, 89).

[321] Roman Rädle, Hans-Christian Jetter, Jens Müller, and Harald Reiterer. "Bigger Is Not Always Better: Display Size, Performance, and Task Load during Peephole Map Navigation." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '14. New York, NY, USA: ACM, 2014, pp. 4127–4136. ISBN: 978-1-4503-2473-1. DOI: 10.1145/2556288.2557071 (cit. on pp. 85–88, 101–104).

[322] Roman Rädle, Hans-Christian Jetter, Mario Schreiner, Zhihao Lu, Harald Reiterer, and Yvonne Rogers. "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, NY, USA: ACM, 2015, pp. 3913–3922. isbn: 978-1-4503-3145-6. doi: 10.1145/2702123.2702287 (cit. on p. 22).

[323] Roman Rädle, Midas Nouwens, Kristian Antonsen, James R. Eagan, and Clemens N. Klokmose. “Codestrates: Literate Computing with Webstrates.” In: *Proceedings of the ACM Symposium on User Interface Software and Technology*. UIST ’17. New York, NY, USA: ACM, 2017, pp. 715–725. isbn: 978-1-4503-4981-9. doi: 10.1145/3126594.3126642 (cit. on pp. 115, 116, 119, 121).

[324] Hanna-Riikka Rantamaa, Jari Kangas, Sriram Kishore Kumar, Helena Mehtonen, Jorma Järnstedt, and Roope Raisamo. “Comparison of a VR Stylus with a Controller, Hand Tracking, and a Mouse for Object Manipulation and Medical Marking Tasks in Virtual Reality.” In: *Applied Sciences* 13.4 (2023), p. 2251. issn: 2076-3417. doi: 10.3390/app13042251 (cit. on p. 150).

[325] Umar Rashid, Miguel A. Nacenta, and Aaron Quigley. “Factors influencing visual attention switch in multi-display user interfaces: a survey.” In: *Proceedings of the ACM Symposium on Pervasive Displays*. PerDis 2012. New York, NY, USA: ACM, 2012, pp. 1–6. doi: 10.1145/2307798.2307799 (cit. on pp. 45, 87, 102, 105, 174, 175).

[326] Umar Rashid, Miguel A. Nacenta, and Aaron Quigley. “The cost of display switching: a comparison of mobile, large display and hybrid UI configurations.” In: *Proceedings of the ACM Conference on Advanced Visual Interfaces*. AVI’12. New York, NY, USA: ACM, 2012, pp. 99–106. doi: 10.1145/2254556.2254577 (cit. on pp. 86, 87, 175).

[327] Tobias Rau, Tobias Isenberg, Andreas Koehn, Michael Sedlmair, and Benjamin Lee. “Traversing Dual Realities: Investigating Techniques for Transitioning 3D Objects between Desktop and Augmented Reality Environments.” In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI ’25. New York, NY, USA: ACM, 2025, pp. 1–16. isbn: 979-8-4007-1394-1. doi: 10.1145/3706598.3713949 (cit. on pp. 152, 164, 166).

[328] Carolin Reichherzer, Andrew Cunningham, Tracey Coleman, Ruochen Cao, Kurt McManus, Dion Sheppard, Mark Kohler, Mark Billinghurst, and Bruce H Thomas. “Bringing the Jury to the Scene of the Crime: Memory and Decision-Making in a Simulated Crime Scene.” In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI ’21. New York, NY, USA: ACM, 2021, pp. 1–12. isbn: 978-1-4503-8096-6. doi: 10.1145/3411764.3445464 (cit. on p. 116).

[329] Carolin Reichherzer, Jack Fraser, and Mark Billinghurst. “SecondSight: Demonstrating Cross-Device Augmented Reality.” In: *ACM SIGGRAPH 2021 Immersive Pavilion*. SIGGRAPH ’21. New York, NY, USA: ACM, 2021, pp. 51–52. isbn: 978-1-4503-8368-4. doi: 10.1145/3450615.3464545 (cit. on p. 36).

[330] Carolin Reichherzer, Jack Fraser, Damien Constantine Rompapas, and Mark Billinghurst. “SecondSight: A Framework for Cross-Device Augmented Reality Interfaces.” In: *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems*. CHI ’21. New York, NY, USA: ACM, 2021, pp. 1–6. isbn: 978-1-4503-8368-4. doi: 10.1145/3411763.3451839 (cit. on pp. 36, 45, 49, 51).

[331] Anke V. Reinschluessel and Johannes Zagermann. "Exploring Hybrid User Interfaces for Surgery Planning." In: *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2023, pp. 208–210. isbn: 9798350328912. doi: 10.1109/ismar-adjunct60411.2023.00048 (cit. on pp. 36, 180).

[332] Patrick Reipschläger, Frederik Brudy, Raimund Dachselt, Justin Matejka, George Fitzmaurice, and Fraser Anderson. "AvatAR: An Immersive Analysis Environment for Human Motion Data Combining Interactive 3D Avatars and Trajectories." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '22. New York, NY, USA: ACM, 2022, pp. 1–15. isbn: 978-1-4503-9157-3. doi: 10.1145/3491102.3517676 (cit. on p. 182).

[333] Patrick Reipschläger and Raimund Dachselt. "DesignAR: Immersive 3D-Modeling Combining Augmented Reality with Interactive Displays." In: *Proceedings of the ACM Conference on Interactive Surfaces and Spaces*. ISS '19. New York, NY, USA: ACM, 2019, pp. 29–41. isbn: 978-1-4503-6891-9. doi: 10.1145/3343055.3359718 (cit. on pp. 41, 49, 105, 149, 152).

[334] Patrick Reipschläger, Severin Engert, and Raimund Dachselt. "Augmented Displays: Seamlessly Extending Interactive Surfaces With Head-Mounted Augmented Reality." In: *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems*. CHI '20. New York, NY, USA: ACM, 2020, pp. 1–4. doi: 10.1145/3334480.3383138 (cit. on pp. 41, 48, 49, 51, 149).

[335] Patrick Reipschläger, Tamara Flemisch, and Raimund Dachselt. "Personal Augmented Reality for Information Visualization on Large Interactive Displays." In: *IEEE Transactions on Visualization and Computer Graphics* 27.2 (2021), pp. 1182–1192. issn: 2160-9306. doi: 10.1109/tvcg.2020.3030460 (cit. on pp. 41–43, 48, 78, 183).

[336] Patrick Reipschläger, Burcu Kulahcioglu Ozkan, Aman Shankar Mathur, Stefan Gumhold, Rupak Majumdar, and Raimund Dachselt. "DebugAR: Mixed Dimensional Displays for Immersive Debugging of Distributed Systems." In: *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems*. CHI '18. New York, NY, USA: ACM, 2018, pp. 1–6. isbn: 978-1-4503-5621-3. doi: 10.1145/3170427.3188679 (cit. on pp. 41, 49, 51).

[337] Jun Rekimoto and Masanori Saitoh. "Augmented surfaces: a spatially continuous work space for hybrid computing environments." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '99. New York, NY, USA: ACM, 1999, pp. 378–385. doi: 10.1145/302979.303113 (cit. on pp. 41, 43, 49, 150).

[338] Jonathan C. Roberts. "State of the Art: Coordinated & Multiple Views in Exploratory Visualization." In: *Fifth International Conference on Coordinated and Multiple Views in Exploratory Visualization (CMV 2007)*. Piscataway, NJ, USA: IEEE, 2007, pp. 61–71. isbn: 978-0-7695-2903-5. doi: 10.1109/CMV.2007.20 (cit. on p. 25).

[339] Yvonne Rogers. "HCI Theory: Classical, Modern, and Contemporary." In: *Synthesis Lectures on Human-Centered Informatics* 5.2 (2012), pp. 1–129. issn: 1946-7699. doi: 10.2200/S00418ED1V01Y201205HCI014 (cit. on pp. 113, 116).

[340] Ayshwarya Saktheeswaran, Arjun Srinivasan, and John Stasko. "Touch? Speech? Or Touch and Speech? Investigating Multimodal Interaction for Visual Network Exploration and Analysis." In: *IEEE Transactions on Visualization and Computer Graphics* 26.6 (2020), pp. 2168–2179. ISSN: 2160-9306. DOI: 10.1109/TVCG.2020.2970512 (cit. on p. 60).

[341] Lawrence Sambrooks and Brett Wilkinson. "Comparison of Gestural, Touch, and Mouse Interaction with Fitts' Law." In: *Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration*. OzCHI '13. New York, NY, USA: ACM, 2013, pp. 119–122. ISBN: 978-1-4503-2525-7. DOI: 10.1145/2541016.2541066 (cit. on p. 149).

[342] Audrey Sanctorum and Beat Signer. "A Unifying Reference Framework and Model for Adaptive Distributed Hybrid User Interfaces." In: *2019 13th International Conference on Research Challenges in Information Science (RCIS)*. Piscataway, NJ, USA: IEEE, 2019, pp. 101–106. ISBN: 978-1-7281-4844-1. DOI: 10.1109/RCIS.2019.8877048 (cit. on pp. 33, 37).

[343] C. Sandor, A. Olwal, B. Bell, and S. Feiner. "Immersive mixed-reality configuration of hybrid user interfaces." In: *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Vol. 2005. Piscataway, NJ, USA: IEEE, 2005, pp. 110–113. ISBN: 0-7695-2459-1. DOI: 10.1109/ismar.2005.37 (cit. on pp. 32, 49, 182).

[344] Prasanth Sasikumar, Max Collins, Huidong Bai, and Mark Billinghurst. "XRTB: A Cross Reality Teleconference Bridge to Incorporate 3D Interactivity to 2D Teleconferencing." In: *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems*. CHI '21. New York, NY, USA: ACM, 2021, pp. 1–4. ISBN: 978-1-4503-8095-9. DOI: 10.1145/3411763.3451546 (cit. on p. 113).

[345] Marc Satkowski, Weizhou Luo, and Raimund Dachselt. "Towards In-situ Authoring of AR Visualizations with Mobile Devices." In: *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2021, pp. 324–325. ISBN: 978-1-66541-298-8. DOI: 10.1109/ISMAR-Adjunct54149.2021.00073 (cit. on pp. 41, 43, 48).

[346] Marc Satkowski and Julián Méndez. "Fantastic Hybrid User Interfaces and How to Define Them." In: *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2023, pp. 247–250. DOI: 10.1109/ismar-adjunct60411.2023.00057 (cit. on pp. 31, 33).

[347] Marc Satkowski, Rufat Rzayev, Eva Goebel, and Raimund Dachselt. "ABOVE & BELOW: Investigating Ceiling and Floor for Augmented Reality Content Placement." In: *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2022, pp. 518–527. DOI: 10.1109/ismar55827.2022.00068 (cit. on p. 44).

[348] Arvind Satyanarayan, Dominik Moritz, Kanit Wongsuphasawat, and Jeffrey Heer. "Vega-Lite: A Grammar of Interactive Graphics." In: *IEEE Transactions on Visualization and Computer Graphics* 23.1 (2017), pp. 341–350. ISSN: 1077-2626. DOI: 10.1109/tvcg.2016.2599030 (cit. on pp. 115, 120, 183).

[349] Arvind Satyanarayan, Ryan Russell, Jane Hoffswell, and Jeffrey Heer. "Reactive Vega: A Streaming Dataflow Architecture for Declarative Interactive Visualization." In: *IEEE Transactions on Visualization and Computer Graphics* 22.1 (2016), pp. 659–668. issn: 1077-2626. doi: 10.1109/tvcg.2015.2467091 (cit. on p. 183).

[350] Joey Scarr, Andy Cockburn, and Carl Gutwin. "Supporting and Exploiting Spatial Memory in User Interfaces." In: *Foundations and Trends® in Human–Computer Interaction* 6.1 (2013), pp. 1–84. issn: 1551-3963. doi: 10.1561/1100000046 (cit. on p. 86).

[351] Dieter Schmalstieg and Tobias Höllerer. *Augmented Reality: Principles and Practice*. Addison-Wesley Usability and HCI Series. Boston: Addison-Wesley, 2016. isbn: 978-0-321-88357-5 (cit. on pp. 5, 150).

[352] Martin Schrepp, Andreas Hinderks, and Jörg Thomaschewski. "Construction of a Benchmark for the User Experience Questionnaire (UEQ)." In: *International Journal of Interactive Multimedia and Artificial Intelligence* 4.4 (2017), p. 40. issn: 1989-1660. doi: 10.9781/ijimai.2017.445 (cit. on pp. 70, 90, 92, 100).

[353] Jan-Henrik Schröder and Hans-Christian Jetter. "Towards a Model for Space and Time in Transitional Collaboration." In: *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2023, pp. 223–227. isbn: 9798350328912. doi: 10.1109/ismar-adjunct60411.2023.00052 (cit. on p. 182).

[354] Jan-Henrik Schröder, Daniel Schacht, Niklas Peper, Anita Marie Hamurculu, and Hans-Christian Jetter. "Collaborating Across Realities: Analytical Lenses for Understanding Dyadic Collaboration in Transitional Interfaces." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '23. New York, NY, USA: ACM, 2023, pp. 1–16. isbn: 978-1-4503-9421-5. doi: 10.1145/3544548.3580879 (cit. on pp. 42, 43, 49, 181).

[355] Daniel Schwajda, Judith Friedl, Fabian Pointecker, Hans-Christian Jetter, and Christoph Anthes. "Transforming graph data visualisations from 2D displays into augmented reality 3D space: A quantitative study." In: *Frontiers in Virtual Reality* 4 (2023). issn: 2673-4192. doi: 10.3389/frvir.2023.1155628 (cit. on pp. 21, 43–45, 152, 177, 179).

[356] Mohammad Rajabi Seraji and Wolfgang Stuerzlinger. "HybridAxes: An Immersive Analytics Tool With Interoperability Between 2D and Immersive Reality Modes." In: *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2022, pp. 155–160. isbn: 978-1-6654-5365-3. doi: 10.1109/ISMAR-Adjunct57072.2022.00036 (cit. on pp. 151, 152).

[357] Mickael Sereno, Stéphane Gosset, Lonni Besançon, and Tobias Isenberg. "Hybrid Touch/Tangible Spatial Selection in Augmented Reality." In: *Computer Graphics Forum* 41.3 (2022), pp. 403–415. issn: 1467-8659. doi: 10.1111/cgf.14550 (cit. on pp. 40, 177).

[358] Marcos Serrano, Barrett Ens, Xing-Dong Yang, and Pourang Irani. "Gluey: Developing a Head-Worn Display Interface to Unify the Interaction Experience in Distributed Display Environments." In: *Proceedings of the ACM Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI '15. New York, NY, USA: ACM, 2015, pp. 161–171. isbn: 978-1-4503-3652-9. doi: 10.1145/2785830.2785838 (cit. on pp. 42, 44, 49).

[359] Jenna M. Shanis and Alan Hedge. "Comparison of Mouse, Touchpad and Multitouch Input Technologies." In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 47.4 (2003), pp. 746–750. ISSN: 2169-5067, 1071-1813. DOI: 10.1177/154193120304700418 (cit. on pp. 147, 149).

[360] B. Shneiderman. "Why Not Make Interfaces Better than 3D Reality?" In: *IEEE Computer Graphics and Applications* 23.6 (2003), pp. 12–15. ISSN: 0272-1716. DOI: 10.1109/MCG.2003.1242376 (cit. on pp. 7, 24, 151).

[361] Ronell Sicat, Jiabao Li, Junyoung Choi, Maxime Cordeil, Won-Ki Jeong, Benjamin Bach, and Hanspeter Pfister. "DXR: A Toolkit for Building Immersive Data Visualizations." In: *IEEE Transactions on Visualization and Computer Graphics* 25.1 (2019), pp. 715–725. ISSN: 1077-2626, 1941-0506, 2160-9306. DOI: 10.1109/TVCG.2018.2865152 (cit. on pp. 24, 183).

[362] Adalberto L. Simeone, Mohamed Khamis, Augusto Esteves, Florian Daiber, Matjaž Kljun, Klen Čopič Pucihar, Poika Isokoski, and Jan Gugenheimer. "International Workshop on Cross-Reality (XR) Interaction." In: *Companion Proceedings of the ACM Conference on Interactive Surfaces and Spaces*. ISS '20. New York, NY, USA: ACM, 2020, pp. 111–114. ISBN: 978-1-4503-7526-9. DOI: 10.1145/3380867.3424551 (cit. on pp. 8, 20).

[363] Richard Skarbez, Missie Smith, and Mary C. Whitton. "Revisiting Milgram and Kishino's Reality-Virtuality Continuum." In: *Frontiers in Virtual Reality* 2 (2021), p. 647997. ISSN: 2673-4192. DOI: 10.3389/frvir.2021.647997 (cit. on p. 114).

[364] Amy Skopik and Carl Gutwin. "Improving Revisitation in Fisheye Views with Visit Wear." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI05. New York, NY, USA: ACM, 2005, pp. 771–780. ISBN: 978-1-58113-998-3. DOI: 10.1145/1054972.1055079 (cit. on p. 86).

[365] Anthony Smith, Kristy de Salas, Benjamin Schüz, Stuart G Ferguson, and Ian Lewis. "mHealth Intervention Design: Creating mHealth Interventions for Behaviour Change." In: *Proceedings of the 28th Australian Conference on Computer-Human Interaction*. OzCHI '16. New York, NY, USA: ACM, 2016, pp. 531–536. ISBN: 978-1-4503-4618-4. DOI: 10.1145/3010915.3010986 (cit. on p. 79).

[366] Shamus P. Smith and Dilip Gahankari. "Integrating View Magnification into an Augmented Reality Head-Mounted Display to Support Surgery." In: *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2023, pp. 238–241. ISBN: 9798350328912. DOI: 10.1109/ismar-adjunct60411.2023.00055 (cit. on p. 180).

[367] Maurício Sousa, Daniel Mendes, Soraia Paulo, Nuno Matela, Joaquim Jorge, and Daniel Simões Lopes. "VRRRRoom: Virtual Reality for Radiologists in the Reading Room." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '17. New York, NY, USA: ACM, 2017, pp. 4057–4062. ISBN: 978-1-4503-4655-9. DOI: 10.1145/3025453.3025566 (cit. on p. 149).

[368] Maximilian Speicher, Brian D. Hall, and Michael Nebeling. "What Is Mixed Reality?" In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '19. New York, NY, USA: ACM, 2019, pp. 101–115. ISBN: 978-1-4503-5970-2. DOI: 10.1145/3290605.3300767 (cit. on p. 1).

[369] Arjun Srinivasan, Bongshin Lee, Nathalie Henry Riche, Steven M. Drucker, and Ken Hinckley. “InChorus: Designing Consistent Multimodal Interactions for Data Visualization on Tablet Devices.” In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI ’20. New York, NY, USA: ACM, 2020, pp. 1–13. isbn: 978-1-4503-6708-0. doi: 10.1145/3313831.3376782 (cit. on pp. 60, 66).

[370] Arjun Srinivasan and John Stasko. “Orko: Facilitating Multimodal Interaction for Visual Exploration and Analysis of Networks.” In: *IEEE Transactions on Visualization and Computer Graphics* 24.1 (2018), pp. 511–521. issn: 1077-2626. doi: 10.1109/TVCG.2017.2745219 (cit. on pp. 60, 66, 73).

[371] Seung Youb Ssin, James A. Walsh, Ross T. Smith, Andrew Cunningham, and Bruce H. Thomas. “GeoGate: Correlating Geo-Temporal Datasets Using an Augmented Reality Space-Time Cube and Tangible Interactions.” In: *Proceedings of the IEEE Conference on Virtual Reality*. New York, NY, USA: IEEE, 2019, pp. 210–219. isbn: 978-1-7281-1377-7. doi: 10.1109/VR.2019.8797812 (cit. on p. 24).

[372] Carolin Stellmacher, Florian Mathis, Yannick Weiss, Meagan B. Loerakker, Nadine Wagener, and Johannes Schöning. “Exploring Mobile Devices as Haptic Interfaces for Mixed Reality.” In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI ’24. New York, NY, USA: ACM, 2024, pp. 1–17. isbn: 9798400703300. doi: 10.1145/3613904.3642176 (cit. on p. 48).

[373] Leon Straker, Clare Pollock, Anthony Frosh, Arne Aarås, and Marvin Dainoff. “An Ergonomic Field Comparison of a Traditional Computer Mouse and a Vertical Computer Mouse in Uninjured Office Workers.” In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 44.33 (2000), pp. 6-356-6–359. issn: 2169-5067. doi: 10.1177/154193120004403376 (cit. on pp. 149, 155).

[374] Amanda Strawhacker and Marina U. Bers. ““I want my robot to look for food”: Comparing Kindergartner’s programming comprehension using tangible, graphic, and hybrid user interfaces.” In: *International Journal of Technology and Design Education* 25.3 (2014), pp. 293–319. issn: 1573-1804. doi: 10.1007/s10798-014-9287-7 (cit. on pp. 32, 33).

[375] Hemant Bhaskar Surale, Aakar Gupta, Mark Hancock, and Daniel Vogel. “TabletInVR: Exploring the Design Space for Using a Multi-Touch Tablet in Virtual Reality.” In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI ’19. New York, NY, USA: ACM, 2019, 13:1–13:13. isbn: 978-1-4503-5970-2. doi: 10.1145/3290605.3300243 (cit. on pp. 47, 49, 75–77).

[376] Ivan E. Sutherland. “A Head-Mounted Three Dimensional Display.” In: *Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I on - AFIPS ’68 (Fall, Part I)*. AFIPS ’68 (Fall, part I). New York, NY, USA: ACM, 1968, p. 757. doi: 10.1145/1476589.1476686 (cit. on pp. 1, 2).

[377] Ivan E. Sutherland. “The Ultimate Display.” In: *Proceedings of the IFIP Congress*. Vol. 2. New York, NY, USA, 1965, pp. 506–508 (cit. on pp. 12, 168, 189, 197).

[378] Ryo Suzuki, Adnan Karim, Tian Xia, Hooman Hedayati, and Nicolai Marquardt. “Augmented Reality and Robotics: A Survey and Taxonomy for AR-enhanced Human-Robot Interaction and Robotic Interfaces.” In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI ’22. New York, NY, USA: ACM, 2022, pp. 1–33. isbn: 978-1-4503-9157-3. doi: 10.1145/3491102.3517719 (cit. on p. 142).

[379] Danielle Albers Szafir. “The Good, the Bad, and the Biased: Five Ways Visualizations Can Mislead (and How to Fix Them).” In: *Interactions* 25.4 (2018), pp. 26–33. ISSN: 1558-3449. DOI: 10.1145/3231772 (cit. on pp. 7, 24, 151).

[380] Zsolt Szalavári and Michael Gervautz. “The Personal Interaction Panel – a Two-Handed Interface for Augmented Reality.” In: *Computer Graphics Forum* 16.3 (1997), pp. C335–C346. ISSN: 1467-8659. DOI: 10.1111/1467-8659.00137 (cit. on pp. 40, 41, 47, 49, 76).

[381] Slawomir Konrad Tadeja, Timoleon Kipouros, and Per Ola Kristensson. “Exploring Parallel Coordinates Plots in Virtual Reality.” In: *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems*. CHI ’19. New York, NY, USA: ACM, 2019, pp. 1–6. ISBN: 978-1-4503-5971-9. DOI: 10.1145/3290607.3313068 (cit. on p. 24).

[382] Desney S. Tan, Randy Pausch, Jeanine K. Stefanucci, and Dennis R. Proffitt. “Kinesthetic Cues Aid Spatial Memory.” In: *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems*. CHI02. New York, NY, USA: ACM, 2002, p. 806. ISBN: 978-1-58113-454-4. DOI: 10.1145/506443.506607 (cit. on p. 86).

[383] Anthony Tang, Michel Pahud, Sheelagh Carpendale, and Bill Buxton. “VisTACO: Visualizing Tabletop Collaboration.” In: *Proceedings of the ACM Conference on Interactive Surfaces and Spaces*. ITS’10. New York, NY, USA: ACM, 2010, pp. 29–38. ISBN: 978-1-4503-0399-6. DOI: 10.1145/1936652.1936659 (cit. on p. 115).

[384] Anthony Tang, Melanie Tory, Barry Po, Petra Neumann, and Sheelagh Carpendale. “Collaborative coupling over tabletop displays.” In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI06. New York, NY, USA: ACM, 2006, pp. 1181–1190. DOI: 10.1145/1124772.1124950 (cit. on p. 181).

[385] Robert J. Teather and Wolfgang Stuerzlinger. “Visual Aids in 3D Point Selection Experiments.” In: *Proceedings of the ACM Symposium on Spatial User Interfaces*. SUI ’14. New York, NY, USA: ACM, 2014, pp. 127–136. ISBN: 978-1-4503-2820-3. DOI: 10.1145/2659766.2659770 (cit. on p. 167).

[386] Paul Tolstoi and Andreas Dippon. “Towering Defense: An Augmented Reality Multi-Device Game.” In: *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems*. CHI ’15. New York, NY, USA: ACM, 2015, pp. 89–92. ISBN: 978-1-4503-3146-3. DOI: 10.1145/2702613.2728659 (cit. on p. 49).

[387] Wai Tong, Haobo Li, Meng Xia, Wong Kam-Kwai, Ting-Chuen Pong, Huamin Qu, and Yalong Yang. “Exploring Spatial Hybrid User Interface for Visual Sensemaking.” In: *IEEE Transactions on Visualization and Computer Graphics* (2025), pp. 1–16. ISSN: 2160-9306. DOI: 10.1109/TVCG.2025.3538771 (cit. on pp. 25, 142, 144, 149).

[388] M. Tory, A.E. Kirkpatrick, M.S. Atkins, and T. Moller. “Visualization Task Performance with 2D, 3D, and Combination Displays.” In: *IEEE Transactions on Visualization and Computer Graphics* 12.1 (2006), pp. 2–13. ISSN: 1077-2626. DOI: 10.1109/TVCG.2006.17 (cit. on pp. 7, 24, 151).

[389] Md. Sami Uddin and Carl Gutwin. “The Image of the Interface: How People Use Landmarks to Develop Spatial Memory of Commands in Graphical Interfaces.” In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI ’21. New York, NY, USA: ACM, 2021, pp. 1–17. ISBN: 978-1-4503-8096-6. DOI: 10.1145/3411764.3445050 (cit. on p. 86).

[390] Md. Sami Uddin, Carl Gutwin, and Andy Cockburn. "The Effects of Artificial Landmarks on Learning and Performance in Spatial-Memory Interfaces." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '17. New York, NY, USA: ACM, 2017, pp. 3843–3855. ISBN: 978-1-4503-4655-9. DOI: 10.1145/3025453.3025497 (cit. on pp. 86, 101).

[391] Andries Van Dam. "Post-WIMP User Interfaces." In: *Communications of the ACM* 40.2 (1997), pp. 63–67. ISSN: 1557-7317. DOI: 10.1145/253671.253708 (cit. on p. 22).

[392] Koen van Turnhout, Arthur Bennis, Sabine Craenmehr, Robert Holwerda, Marjolein Jacobs, Ralph Niels, Lambert Zaad, Stijn Hoppenbrouwers, Dick Lenior, and René Bakker. "Design Patterns for Mixed-Method Research in HCI." In: *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational*. NordiCHI '14. New York, NY, USA: ACM, 2014, pp. 361–370. ISBN: 978-1-4503-2542-4. DOI: 10.1145/2639189.2639220 (cit. on p. 9).

[393] Dan Venolia. "Facile 3D Direct Manipulation." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '93. New York, NY, USA: ACM, 1993, pp. 31–36. ISBN: 978-0-89791-575-5. DOI: 10.1145/169059.169065 (cit. on p. 151).

[394] Katja Vock, **Sebastian Hubenschmid**, Johannes Zagermann, Simon Butscher, and Harald Reiterer. "IDIAR: Augmented Reality Dashboards to Supervise Mobile Intervention Studies." In: *Mensch und Computer 2021*. MuC '21. New York, NY: ACM, 2021, pp. 248–259. ISBN: 978-1-4503-8645-6. DOI: 10.1145/3473856.3473876 (cit. on pp. ix, 45, 48, 58, 114, 140, 165, 167).

[395] Ulrich von Zadow and Raimund Dachselt. "GIAnT: Visualizing Group Interaction at Large Wall Displays." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '17. New York, NY, USA: ACM, 2017, pp. 2639–2647. ISBN: 978-1-4503-4655-9. DOI: 10.1145/3025453.3026006 (cit. on p. 115).

[396] Jia Wang and Robert Lindeman. "Coordinated 3D interaction in tablet- and HMD-based hybrid virtual environments." In: *Proceedings of the ACM Symposium on Spatial User Interfaces*. SUI '14. New York, NY, USA: ACM, 2014, pp. 70–79. ISBN: 978-1-4503-2820-3. DOI: 10.1145/2659766.2659777 (cit. on pp. 31, 43, 49).

[397] Mengcheng Wang, Guangchuan Li, Federico Arippa, Alan Barr, Yanmin Xue, and Carisa Harris-Adamson. "The Effects of Mouse Weight and Connection Type on Performance, Muscle Activity, and Preferences among Professional Gamers." In: *International Journal of Industrial Ergonomics* 97 (2023), p. 103493. ISSN: 0169-8141. DOI: 10.1016/j.ergon.2023.103493 (cit. on p. 149).

[398] Nanjia Wang, Sze-Wing Chan, David Aigner, Omar Addam, Christoph Anthes, and Frank Maurer. "Serious Cross Reality - Using CR to Enhance Analytics Workflow." In: *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2023, pp. 44–49. ISBN: 9798350328912. DOI: 10.1109/ismar-adjunct60411.2023.00017 (cit. on p. 178).

[399] Nanjia Wang and Frank Maurer. "A Design Space for Single-User Cross-Reality Applications." In: *Proceedings of the ACM Conference on Advanced Visual Interfaces*. AVI 2022. New York, NY, USA: ACM, 2022, 31:1–31:5. ISBN: 978-1-4503-9719-3. DOI: 10.1145/3531073.3531116 (cit. on p. 20).

[400] Nanjia Wang, Daniel Zielasko, and Frank Maurer. "User Preferences for Interactive 3D Object Transitions in Cross Reality - An Elicitation Study." In: *Proceedings of the ACM Conference on Advanced Visual Interfaces*. AVI 2024. New York, NY, USA: ACM, 2024, pp. 1–9. ISBN: 9798400717642. DOI: 10.1145/3656650.3656698 (cit. on pp. 42, 150, 177).

[401] Yun Wang, Xiaojuan Ma, Qiong Luo, and Huamin Qu. "Data Edibilization: Representing Data with Food." In: *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems*. CHI'16. New York, NY, USA: ACM, 2016, pp. 409–422. ISBN: 978-1-4503-4082-3. DOI: 10.1145/2851581.2892570 (cit. on p. 115).

[402] Michelle Q. Wang Baldonado, Allison Woodruff, and Allan Kuchinsky. "Guidelines for Using Multiple Views in Information Visualization." In: *Proceedings of the ACM Conference on Advanced Visual Interfaces*. AVI00. New York, NY, USA: ACM, 2000, pp. 110–119. ISBN: 978-1-58113-252-6. DOI: 10.1145/345513.345271 (cit. on p. 25).

[403] C. Ware and D.R. Jessome. "Using the Bat: A Six-Dimensional Mouse for Object Placement." In: *IEEE Computer Graphics and Applications* 8.6 (1988), pp. 65–70. ISSN: 0272-1716. DOI: 10.1109/38.20319 (cit. on p. 151).

[404] Colin Ware and Leonard Slipp. "Using Velocity Control to Navigate 3D Graphical Environments: A Comparison of Three Interfaces." In: *Proceedings of the Human Factors Society Annual Meeting* 35.5 (1991), pp. 300–304. ISSN: 0163-5182. DOI: 10.1177/154193129103500513 (cit. on p. 151).

[405] Mark Weiser. "The computer for the 21st Century." In: *ACM SIGMOBILE Mobile Computing and Communications Review* 3.3 (1999), pp. 3–11. ISSN: 1931-1222. DOI: 10.1145/329124.329126 (cit. on pp. 1, 17, 22, 34, 178, 189, 196).

[406] Johann Wentzel, Fraser Anderson, George Fitzmaurice, Tovi Grossman, and Daniel Vogel. "SwitchSpace: Understanding Context-Aware Peeking Between VR and Desktop Interfaces." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '24. New York, NY, USA: ACM, 2024, pp. 1–16. ISBN: 9798400703300. DOI: 10.1145/3613904.3642358 (cit. on pp. 42, 149).

[407] Jonathan Wieland, Hyunsung Cho, **Sebastian Hubenschmid**, Akihiro Kiuchi, Harald Reiterer, and David Lindlbauer. "Push2AR: Enhancing Mobile List Interactions Using Augmented Reality." In: *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2024, pp. 671–680. ISBN: 979-8-3315-1647-5. DOI: 10.1109/ISMAR62088.2024.00082 (cit. on pp. viii, 6, 47, 84, 106, 166, 175).

[408] Jonathan Wieland, Johannes Zagermann, Jens Muller, and Harald Reiterer. "Separation, Composition, or Hybrid? – Comparing Collaborative 3D Object Manipulation Techniques for Handheld Augmented Reality." In: *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2021, pp. 403–412. ISBN: 978-1-6654-0158-6. DOI: 10.1109/ISMAR52148.2021.00057 (cit. on pp. 117, 184).

[409] R. Wilton and A. J. Pennisi. "Evaluating the Accuracy of Transcribed Clinical Data." In: *Proceedings of the Annual Symposium on Computer Application in Medical Care* (1993), pp. 279–283. ISSN: 0195-4210 (cit. on p. 79).

[410] Jacob O. Wobbrock and Julie A. Kientz. "Research contributions in human-computer interaction." In: *Interactions* 23.3 (2016), pp. 38–44. issn: 1558-3449. doi: 10.1145/2907069 (cit. on pp. 10, 37, 50, 191).

[411] Dennis Wolf, Jan Gugenheimer, Marco Combosch, and Enrico Rukzio. "Understanding the Heisenberg Effect of Spatial Interaction: A Selection Induced Error for Spatially Tracked Input Devices." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '20. New York, NY, USA: ACM, 2020, pp. 101–120. isbn: 978-1-4503-6708-0. doi: 10.1145/3313831.3376876 (cit. on pp. 3, 147).

[412] Pui Chung Wong, Kening Zhu, Xing-Dong Yang, and Hongbo Fu. "Exploring Eyes-free Bezel-initiated Swipe on Round Smartwatches." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '20. New York, NY, USA: ACM, 2020, pp. 1–11. isbn: 978-1-4503-6708-0. doi: 10.1145/3313831.3376393 (cit. on p. 61).

[413] Shengzhi Wu, Daragh Byrne, and Molly Wright Steenson. ""Mergereality": Leveraging Physical Affordances for Multi-Device Gestural Interaction in Augmented Reality." In: *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems*. CHI '20. New York, NY, USA: ACM, 2020, pp. 1–4. isbn: 978-1-4503-6819-3. doi: 10.1145/3334480.3383170 (cit. on pp. 42–45, 177).

[414] Kathrin Wunsch, Janis Fiedler, **Sebastian Hubenschmid**, Harald Reiterer, Britta Renner, and Alexander Woll. "An mHealth Intervention Promoting Physical Activity and Healthy Eating in a Family Setting (SMARTFAMILY): Randomized Controlled Trial." In: *JMIR mHealth and uHealth* 12 (2024). doi: 10.2196/51201 (cit. on pp. x, 79).

[415] Wanze Xie, Yining Liang, Janet Johnson, Andrea Mower, Samuel Burns, Colleen Chelini, Paul D'Alessandro, Nadir Weibel, and Jürgen P. Schulze. "Interactive Multi-User 3D Visual Analytics in Augmented Reality." In: *Electronic Imaging* 32.13 (2020), pp. 363-1-363–6. issn: 2470-1173. doi: 10.2352/ISSN.2470-1173.2020.13.ERVR-363 (cit. on p. 114).

[416] Yan Yan, Ke Chen, Yu Xie, Yiming Song, and Yonghong Liu. "The Effects of Weight on Comfort of Virtual Reality Devices." In: *Advances in Ergonomics in Design*. Ed. by Francisco Rebelo and Marcelo M. Soares. Vol. 777. New York, NY, USA: Springer Publishing, 2019, pp. 239–248. isbn: 9783319947068. doi: 10.1007/978-3-319-94706-8_27 (cit. on pp. 1, 116).

[417] Yishu Yan, Ketki Joshi, Alan Barr, and Carisa Harris Adamson. "The Impact of Computer Mice Weight on Muscle Activity, Performance, and User Preferences While Gaming." In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 66.1 (2022), pp. 868–870. issn: 2169-5067. doi: 10.1177/1071181322661516 (cit. on p. 149).

[418] Yalong Yang, Tim Dwyer, Bernhard Jenny, Kim Marriott, Maxime Cordeil, and Haohui Chen. "Origin-Destination Flow Maps in Immersive Environments." In: *IEEE Transactions on Visualization and Computer Graphics* 25.1 (2018), pp. 693–703. issn: 2160-9306. doi: 10.1109/TVCG.2018.2865192 (cit. on p. 24).

[419] Bo Yi, Xiang Cao, Morten Fjeld, and Shengdong Zhao. "Exploring User Motivations for Eyes-Free Interaction on Mobile Devices." In: *Proceedings of the ACM Conference on Advanced Visual Interfaces*. CHI '12. New York, NY, USA: ACM, 2012, p. 2789. ISBN: 978-1-4503-1015-4. DOI: 10.1145/2207676.2208678 (cit. on pp. 61, 65).

[420] Johannes Zagermann. "Evaluating Complementary Interfaces as a Design-Informing Activity." PhD thesis. Konstanz: Universität Konstanz, 2024 (cit. on p. 23).

[421] Johannes Zagermann, **Sebastian Hubenschmid**, Daniel Immanuel Fink, Jonathan Wieland, Harald Reiterer, and Tiare Feuchtner. "Challenges and Opportunities for Collaborative Immersive Analytics with Hybrid User Interfaces." In: *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Sydney, Australia: IEEE, 2023. DOI: 10.1109/ismar-adjunct60411.2023.00044 (cit. on pp. x, 36, 181).

[422] Johannes Zagermann, Ulrike Pfeil, Philipp von Bauer, Daniel Fink, and Harald Reiterer. ""It's in my other hand!" – Studying the Interplay of Interaction Techniques and Multi-Tablet Activities." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '20. New York, NY, USA: ACM, 2020, pp. 1–13. ISBN: 9781450367080. DOI: 10.1145/3313831.3376540 (cit. on pp. 8, 181).

[423] Johannes Zagermann, Ulrike Pfeil, Daniel Fink, Philipp von Bauer, and Harald Reiterer. "Memory in Motion: The Influence of Gesture- and Touch-Based Input Modalities on Spatial Memory." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '17. New York, NY, USA: ACM, 2017, pp. 1899–1910. ISBN: 978-1-4503-4655-9. DOI: 10.1145/3025453.3026001 (cit. on pp. 22, 85–87, 89, 91, 101, 103, 104).

[424] Johannes Zagermann*, **Sebastian Hubenschmid***, Priscilla Balestrucci, Tiare Feuchtner, Sven Mayer, Marc O. Ernst, Albrecht Schmidt, and Harald Reiterer. "Complementary Interfaces for Visual Computing." In: *it - Information Technology* 64.4–5 (2022), pp. 145–154. ISSN: 2196-7032. DOI: 10.1515/itit-2022-0031 (cit. on pp. vii, 5, 6, 10, 16, 112, 147, 197).

[425] Abdelrahman Zaky, Johannes Zagermann, Harald Reiterer, and Tiare Feuchtner. "Opportunities and Challenges of Hybrid User Interfaces for Optimization of Mixed Reality Interfaces." In: *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Piscataway, NJ, USA: IEEE, 2023, pp. 215–219. ISBN: 9798350328912. DOI: 10.1109/ismar-adjunct60411.2023.00050 (cit. on pp. 36, 182).

[426] André Zenner and Antonio Krüger. "Drag:On: A Virtual Reality Controller Providing Haptic Feedback Based on Drag and Weight Shift." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '19. New York, NY, USA: ACM, 2019, pp. 1–12. ISBN: 978-1-4503-5970-2. DOI: 10.1145/3290605.3300441 (cit. on p. 167).

[427] Shumin Zhai. "User Performance in Relation to 3D Input Device Design." In: *ACM SIGGRAPH Computer Graphics* 32.4 (1998), pp. 50–54. ISSN: 0097-8930. DOI: 10.1145/307710.307728 (cit. on p. 151).

[428] Shuhao Zhang, Yue Li, Ka Lok Man, Yong Yue, and Jeremy Smith. "Towards Cross-Reality Interaction and Collaboration: A Comparative Study of Object Selection and Manipulation in Reality and Virtuality." In: *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 2023, pp. 330–337. DOI: 10.1109/VRW58643.2023.00075 (cit. on pp. 150, 166).

[429] Lixiang Zhao, Tobias Isenberg, Fuqi Xie, Hai-Ning Liang, and Lingyun Yu. "SpatialTouch: Exploring Spatial Data Visualizations in Cross-Reality." In: *IEEE Transactions on Visualization and Computer Graphics* 31.1 (2025), pp. 897–907. ISSN: 1077-2626, 1941-0506, 2160-9306. DOI: 10.1109/TVCG.2024.3456368 (cit. on p. 150).

[430] Qian Zhou, George Fitzmaurice, and Fraser Anderson. "In-Depth Mouse: Integrating Desktop Mouse into Virtual Reality." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '22. New York, NY, USA: ACM, 2022, pp. 1–17. ISBN: 978-1-4503-9157-3. DOI: 10.1145/3491102.3501884 (cit. on p. 150).

[431] Fengyuan Zhu and Tovi Grossman. "BISHARE: Exploring Bidirectional Interactions Between Smartphones and Head-Mounted Augmented Reality." In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. CHI '20. New York, NY, USA: ACM, 2020, pp. 1–14. DOI: 10.1145/3313831.3376233 (cit. on pp. 41, 44, 45, 50, 51, 105).

[432] Qian Zhu, Tao Lu, Shunan Guo, Xiaojuan Ma, and Yalong Yang. "CompositingVis: Exploring Interactions for Creating Composite Visualizations in Immersive Environments." In: *IEEE Transactions on Visualization and Computer Graphics* 31.1 (2025), pp. 591–601. ISSN: 1077-2626, 1941-0506, 2160-9306. DOI: 10.1109/TVCG.2024.3456210 (cit. on p. 25).

[433] Chris Zimmerer, Philipp Krop, Martin Fischbach, and Marc Erich Latoschik. "Reducing the Cognitive Load of Playing a Digital Tabletop Game with a Multimodal Interface." In: *CHI Conference on Human Factors in Computing Systems*. CHI '22. New York, NY, USA: ACM, 2022, pp. 1–13. ISBN: 978-1-4503-9157-3. DOI: 10.1145/3491102.3502062 (cit. on p. 22).

List of Figures

List of Tables