Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Immersive Inscribed Spaces:

Spatial Interaction Techniques for Understanding Written Artefacts

Dissertation to achieve a doctoral degree at the Faculty of Mathematics, Informatics and Natural Sciences at the University of Hamburg

Jenny Gabel

Human-Computer Interaction
Department of Informatics
Faculty of Mathematics, Informatics and Natural Sciences
University of Hamburg

2025

Supervisor & First Reviewer:	Prof. Dr. Frank Steinicke
Supervisor:	Dr. Susanne Schmidt
Second Reviewer:	Prof. Dr. Anil Ufuk Batmaz
Head of Examination Commission:	Prof. Dr. Ralf Möller
Deputy Head of Examination Commission:	Prof. Dr. Eva Bittner
Date of the Thesis Defense:	28.04.2025

ACKNOWLEDGMENTS

The research for this dissertation was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2176 "Understanding Written Artefacts: Material, Interaction and Transmission in Manuscript Cultures" (UWA), project no. 390893796. The research was conducted within the scope of the Centre for the Study of Manuscript Cultures (CSMC) and at the Human-Computer Interaction research group at the University of Hamburg.

I want to thank all individuals and institutions whose support and contributions made this dissertation possible. First, I would like to express my sincerest gratitude to my supervisors Prof. Dr. Frank Steinicke and Dr. Susanne Schmidt. They encouraged me to start my research and embark on this academic journey. Their guidance and insightful feedback helped me to explore new ideas and approach challenges from different perspectives, which has been vital for shaping my research.

I would also like to thank Prof. Dr. Anil Ufuk Batmaz for being the second reviewer of my dissertation and for his encouragement and support during the final phase of my doctoral procedures and the oral defense. Furthermore, I want to thank Prof. Dr. Ralf Möller and Prof. Dr. Eva Bittner for being members of my examination committee and for chairing my oral defense.

I am truly grateful for my amazing colleagues at the Human-Computer Interaction research group and for their support, constructive feedback, and helpful discussions during my work on this dissertation. I feel particularly connected with this research group and its members, as I started my academic journey here as a bachelor student. Being part of this group with its diverse research projects and dedicated and driven members has allowed me to grow personally and professionally throughout my academic career.

Next, I would like to acknowledge the UWA Graduate School and my fellow doctoral students. I really appreciated the exchange and the interesting and enriching insights into diverse manuscript cultures and a wide range of research fields at the UWA Cluster of Excellence.

Moreover, I am thankful to have had the opportunity to be part of two research projects within research field B Inscribing Spaces at the UWA Cluster: RFB02: Immersive City Scripts, involving the ancient theater of Miletus, and RFB05: The Interior of the Church in Lucklum. I would like to thank everyone involved in both projects for the great collaboration: Prof. Dr. Christof Berns, Dr. Sebastian Bosch, Dr. Jost Eickmeyer, Prof. Dr. Kaja Harter, Nathalie Martin, Dr. Ann Lauren Osthof, Prof. Dr. Johann Anselm Steiger, Prof. Dr. Frank Steinicke, and Eric Werner. I am particularly thankful to Dr. Ann Lauren Osthof for the wonderful and close collaboration, for introducing me to archaeology, and for taking me with her on her field research in Miletus. My thanks also extends to our external cooperation partners who carried out the 3D digitization of the two historical sites for the projects: Prof. Michael Breuer, Jens Rothe, Felix Neupert, and Johannes Neis from the laboratory of photogrammetry at the Berliner Hochschule für Technik (BHT Berlin) as well as Martin Schaich and his team from ArcTron 3D.

Special thanks go to the Miletus Excavation / Miletus Kazısı, Prof. Dr. Christof Berns, and Dr. Lisa Steinmann for their support and hospitality during my stay at the Miletus excavation house, in Balat, Türkiye. Thanks to them and Dr. Ann Lauren Osthof, I had the opportunity to gain first-hand experience at an archaeological excavation. Additionally, we were able to carry out the 3D digitization and our research work at the church in Lucklum thanks to the kind permission and support of the Rittergut Lucklum.

Finally, I would like to express my heartfelt gratitude to my friends and family for their continuous support and encouragement throughout my dissertation journey. I am especially grateful to my husband, whose unwavering support in all areas of my life gave me the strength to complete this dissertation.

COPYRIGHT NOTICE

ABSTRACT

Written Artefacts (WA), a cover term for artifacts created by humans, such as written manuscripts and inscriptions [Bau+23], are an important part of our cultural heritage. In this dissertation, we introduce the concept of Immersive Inscribed Spaces (IIS), which are immersive, interactive virtual reality (VR) applications based on real historical sites that contain WA. The research for this dissertation was carried out in the scope of two cross-disciplinary projects at the Cluster of Excellence "Understanding Written Artefacts", where we explored the application of immersive technologies for the study of WA.

Our work addresses the user requirement analysis, feature implementation, the development of spatial interaction techniques, and iterative empirical evaluations: (i) We investigated how interactive, immersive applications for this context can be designed and how they can provide new perspectives for the spatial understanding of WA. We describe our technical setup and implementation of the IIS, including the integration of research data for the hundreds of WA. Since the WA are spatially located inside true-to-scale 3D representations of the historical sites, we explored spatial interaction techniques for interacting with them. (ii) We designed and evaluated four variants of spatial UI panels for reading and interacting with long texts, based on two established desktop UI patterns, continuous scrolling and discrete pagination. Our evaluation showed that desktop UI patterns can be transferred and used for spatial text display in VR applications. (iii) Afterward, we worked on facilitating precise pointer-based selection for small and distant objects. We implemented an assistive raycasting interaction technique that gradually redirects the raycast of the pointer towards the target center to support users with spatial selection. We designed three techniques with one straight and two curved pointers for visualizing the raycast. The findings of our standardized evaluation with a Fitts' Law selection task suggest that the concept of redirection can be successfully applied to raycasting for spatial interactions. Our results revealed that redirected raycasting was faster, achieved higher effective throughput, and had decreased perceived workload compared to raycasting without assistance. The sense of agency was not negatively affected by the redirection, but individual user preferences seem to influence the preferred type of raycast visualization. (iv) We further extended redirection to hand-anchored raycasting, so-called HandRays, and implemented two techniques, (a) based on the previous redirection and (b) an extension using gaze for redirecting the raycast. Our Fitts' Law evaluation confirmed that both techniques improved selection performance and reduced perceived workload in comparison to selection without redirection. However, our findings indicate that performance for technique (b) deteriorated with increased target depth. The subjective feedback revealed different selection strategies that might have caused mismatched timings between gaze-targeting and selection confirmation with the hand gesture. (v) Finally, we focused on eye-tracking-based spatial interaction. We developed Gaze&Blink, a hands-free technique to support discrete selection and continuous interaction tasks, such as scrolling and drag-and-drop. Blinking is used for selections and winking (closing one

eye) in combination with head rotations for performing scroll and drag interactions. We evaluated Gaze&Blink against the established Gaze&Pinch technique [Pfe+17], which uses gaze for targeting and a pinch gesture for selection. Our technique had comparable performance and subjective measures and could be a viable hands-free alternative interaction technique.

A final pilot project demonstrated that the concept of IIS can be transferred to other WA locations. Our work serves as an initial starting point for creating immersive applications for this novel context. Our findings provide learnings, insights, and recommendations for the design and implementation of spatial interaction techniques and the application of redirection for facilitating spatial interaction. We provide an outlook on future work for the application of immersive technologies and their potential for promoting interest and understanding of WA, both in an academic context and beyond.

ZUSAMMENFASSUNG

Written Artefacts (WA), ein Oberbegriff für von Menschen geschaffene, schriftliche Artefakte wie Manuskripte und Inschriften [Bau+23], sind ein wichtiger Teil unseres kulturellen Erbes. In dieser Dissertation stellen wir das Konzept der Immersive Inscribed Spaces (IIS) vor, bei denen es sich um immersive, interaktive Virtual Reality VR Anwendungen handelt, die auf realen historischen Stätten basieren, welche schriftliche Artefakte enthalten. Die Forschung für diese Dissertation wurde im Rahmen von zwei interdisziplinären Projekten am Exzellenzcluster „Understanding Written Artefacts“ durchgeführt, in denen wir die Anwendung immersiver Technologien für die Untersuchung von WA erforschten.

Unsere Arbeit befasst sich mit der Analyse der Benutzeranforderungen, der Implementierung von Funktionen, der Entwicklung räumlicher Interaktionstechniken und iterativen empirischen Evaluationen: (i) Wir haben untersucht, wie interaktive, immersive Anwendungen für diesen Kontext gestaltet werden können und wie sie neue Perspektiven für das räumliche Verständnis von WA bieten können. Wir beschreiben unseren technischen Aufbau und die Umsetzung der IIS, einschließlich der Integration von Forschungsdaten für hunderte von WA. Da sich die WA räumlich innerhalb maßstabsgetreuer 3D-Repräsentationen der historischen Stätten befinden, haben wir räumliche Interaktionstechniken für die Interaktion mit ihnen erforscht. (ii) Wir entwarfen und evaluierten vier Varianten von räumlichen UI-Tafeln zum Lesen und Interagieren mit langen Texten, basierend auf zwei etablierten Desktop-UI-Pattern, kontinuierlichem Scrollen und diskreter Paginierung. Unsere Studie zeigte, dass Desktop-UI-Pattern für die räumliche Textanzeige in VR-Anwendungen übertragen und verwendet werden können. (iii) Anschließend arbeiteten wir daran, die präzise Zeiger-basierte Selektion für kleine und weit entfernte Objekte zu erleichtern. Wir implementierten eine assistive Raycasting-Interaktionstechnik, die den Raycast des Zeigers graduell auf das Zielzentrum umlenkt, um Benutzer bei der räumlichen Selektion zu unterstützen. Wir entwickelten drei Techniken mit einem geraden und zwei gebogenen Zeigern zur Visualisierung des Raycasts. Die Ergebnisse unserer standardisierten Auswertung mit einer Selektionaufgabe nach Fitts’ Gesetz deuten darauf hin, dass das Konzept der Umlenkung (redirection) erfolgreich auf Raycasting für räumliche Interaktionen angewendet werden kann. Unsere Ergebnisse zeigten, dass umgelenktes Raycasting schneller war, einen höheren effektiven Durchsatz (effective throughput) erzielte und die wahrgenommene Arbeitsbelastung im Vergleich zu Raycasting ohne Unterstützung verringerte. Dabei wurde das Gefühl der Handlungsfähigkeit (sense of agency) durch die Umlenkung nicht negativ beeinflusst. Jedoch scheinen individuelle Benutzerpräferenzen einen Einfluss auf die bevorzugte Art der Raycast-Visualisierung zu haben. (iv) Wir weiteten den Einsatz von Umlenkung weiter auf handverankertes Raycasting, sogenannte Hand-Rays, aus und implementierten zwei Techniken, (a) basierend auf der vorherigen Umlenkung und (b) eine Erweiterung, die den Blick zur Umlenkung des Raycasts verwendet. Unsere Auswertung nach Fitts’ Gesetz bestätigte, dass beide Techniken die Selektionsleistung verbesserten und die wahrgenommene

Arbeitsbelastung im Vergleich zur Selektion ohne Umlenkung verringerten. Unsere Ergebnisse deuten jedoch darauf hin, dass sich die Leistung bei Technik (b) mit zunehmender Distanz zum Ziel verschlechtert. Das subjektive Feedback zeigte, dass unterschiedliche Selektionstrategien möglicherweise zu zeitlichen Diskrepanzen zwischen dem Anvisieren des Ziels mit den Augen und der Bestätigung der Selektion durch die Handgeste geführt haben könnten. (v) Schließlich konzentrierten wir uns auf Eye-Tracking basierte räumliche Interaktion. Wir entwickelten Gaze&Blink, eine freihändige Technik für diskrete Selektion und kontinuierliche Interaktionsaufgaben, wie Scrollen und Drag-and-Drop. Blinzeln wird für die Ausführung von Selektionen und Zwinkern (Schließen eines Auges) in Kombination mit Kopfdrehungen für Scroll- und Drag-Interaktionen verwendet. Wir haben Gaze&Blink mit der etablierten Gaze&Pinch [Pfe+17] Technik verglichen, die den Blick zum Zielen und eine Pinch-Geste zur Selektion verwendet. Unsere Technik wies eine vergleichbare Leistung und subjektive Bewertungen auf und könnte eine geeignete freihändige alternative Interaktionstechnik darstellen.

Ein abschließendes Pilotprojekt zeigte, dass das Konzept von IIS auf andere Orte mit WA übertragen werden kann. Unsere Arbeit dient als Ausgangspunkt für die Entwicklung immersiver Anwendungen für diesen neuen Kontext. Unsere Ergebnisse liefern Learnings, Erkenntnisse und Empfehlungen für die Gestaltung und Umsetzung räumlicher Interaktionstechniken und den Einsatz von Umlenkung zur Unterstützung räumlicher Interaktion. Wir geben einen Ausblick auf zukünftige Arbeiten zur Anwendung immersiver Technologien und deren Potenzial, das Interesse und Verständnis für WA zu fördern, sowohl in einem akademischen Kontext als auch darüber hinaus.

CONTENTS

APPENDIX

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

2D Two-dimensional
3D Three-dimensional
AR Augmented Reality
AV Augmented Virtuality
CAVE Cave Automatic Virtual Environment
CE Common Era
CSMC Centre for the Study of Manuscript Cultures
DSLR Digital single-lens reflex camera
FRE Flesch Reading Ease
GHA Gaze-Hand Alignment
GUI Graphical User Interface
HCI Human-Computer Interaction
HMD Head-Mounted Display
IIS Immersive Inscribed Spaces
IPQ Igroup Presence Questionnaire
IS Inscribed Spaces
LIDAR Light Detection And Ranging
MR Mixed Reality
NASA-TLX NASA Task Load Index
RVC Reality-Virtuality Continuum
SSQ Simulator Sickness Questionnaire
SUS System Usability Scale
UCD User-Centered Design
UEQ User Experience Questionnaire
UI User Interface
UWA Understanding Written Artefacts
UX User Experience
VE Virtual Environment
VR Virtual Reality
WA Written Artefacts
XR Extended Reality

CHAPTER 1

INTRODUCTION

1.1 Motivation

The first ideas and prototypes for immersive experiences and head-mounted displays (HMD) already emerged in the 1960s [Ste16; Sut68]. In 1968, Ivan Sutherland [Sut68] created one of the first HMDs, simply called "head-mounted three dimensional display", which used head-tracking and displayed simple 3D wireframes. Up until 1990, HMDs and immersive applications were mainly developed and used at universities and research institutes [Bro+90; Fis+88; Fis+87; KGH85]. In the 1990s, there was a first hype among the general public and academia, and the entertainment industry also began to take an interest in immersive technologies [Jer15b; Ste16]. However, due to the complexity of interactive immersive systems, HMDs were still expensive, bulky, and cumbersome to setup throughout the early 2000s [Del14; Ste16] and, therefore, received little mainstream attention during the first decade of the 21st century [Jer15b]. However, 2016 marked a key year for HMDs with the launch of the Oculus Rift CV1, a high-end but comparatively affordable consumer device at the time [Ste16]. Since then, technical advances in hardware, 3D computer graphics, tracking, and display technologies have made consumer systems widely available. Today, consumer devices such as the Meta Quest HMDs come with high-resolution displays, high frame rate, and built-in multi-modal input tracking, such as hand-tracking, eye-tracking, and controller-tracking. Moreover, these HMDs are standalone devices that do not need to be tethered to a high-performance PC, which further increases their ease of use. These modern devices use different immersive technologies that integrate the physical world and virtual content at varying degrees on a so-called "Reality-Virtuality Continuum" [MK94]. While we explore multiple immersive technologies in this dissertation, our main implementation is based on *virtual reality* (VR), which is defined as follows for our work:

> **Definition 1.1.1 Virtual Reality (VR)** consists of five key elements: (i) the **users** with their individual capabilities, interpretations and experiences, (ii) the creators, who design and implement the application or system, (iii) a simulated, computer-based **virtual world**, (iv) physical and mental **immersion**, which create the sensation of being present, and (v) **interactivity** through user actions such as changing locations and interacting with objects [SC18].

Interactive VR applications can lead to more efficient understanding of a topic or concept than what can be achieved in the real world [Jer15b]. Immersive technologies can benefit a wide range of use cases in various fields and sectors. VR applications have already been successfully used in various contexts such as scientific visualizations, architecture, simulations, training, therapy, entertainment, and education [Jer15b].

This potential of immersive technologies has also attracted the attention of humanities scholars, who study so-called *Written Artefacts* (WA)[1] at the Cluster of Excellence "Understanding Written Artefacts" (UWA) [Cenb] at the Centre for the Study of Manuscript Cultures (CSMC)[1] at the University of Hamburg. The cluster follows a cross-cultural and cross-disciplinary approach for the comprehensive study of WA. The term WA has been defined by the scholars at the CSMC as follows:

> **Definition 1.1.2 Written Artefacts (WA)** is a cover term for all artificial or natural objects with visual signs applied by humans [Bau+23].

WA were produced as part of diverse and complex manuscript cultures and are an important part of our cultural heritage. The term includes written manuscripts, ranging from simple shopping lists to calligraphic boards and copies of literary works, as well as various forms of inscriptions, ranging from official kingly dedications, inscribed vases, advertisements, tombstones, and monumental rock inscriptions to anonymous ancient graffiti [Bau+23; Škr+23]. A wide range of inorganic and organic materials were used for creating WA, including metals, stones, ceramics, ivory, wood, palm leaves, and parchment. They carry purposeful messages and have been used, among other purposes, for adorning structures and buildings, for advertising, for publishing laws, and as statements in space [Cenc]. The research cluster's core objective is to investigate this rich diversity of WA and their associated manuscript cultures outside of traditionally held boundaries of academic disciplines [Cenb]. An important aspect of their work is the study of WA on-site within their spatial context and in their respective distribution of content and function [Cenc]. Based on this research focus, we use the following term throughout this dissertation to refer to WA within their spatial context:

> **Definition 1.1.3 Inscribed Spaces (IS)** are spaces that contain and are structured by WA. They can be landscapes, cityscapes, and interior spaces, including ancient ruins, historical structures, and cultural heritage sites [Cenc].

VR can be used to digitally recreate IS as interactive 3D environments. Therefore, following their cross-disciplinary approach for conducting research outside traditional academic boundaries, the UWA cluster aimed to explore the application of VR technology for spatial visualization and virtual

[1]These terms are written in British English spelling.

access to WA. The integration of immersive technology for the study of WA motivated the work for this dissertation. Within the scope of two research projects at the cluster, we conducted our research, which involved IS that contain a large amount of WA. We designed and implemented interactive VR applications to explore the potential advantages and benefits that immersive technologies, and especially VR, can offer for the study of WA. With our work in this context, we introduce the concept of so-called *Immersive Inscribed Spaces* (IIS), which we define as follows:

> **Definition 1.1.4 Immersive Inscribed Spaces (IIS)** are immersive applications with 3D representations of their physical, real-world Inscribed Spaces that integrate and visualize the research data of the WA, and provide interactive techniques and features for accessing and exploring the WA within their spatial context.

For interacting with these IIS, a fundamental task for users is to select 3D objects [LaV+17]. These selections are one type of spatial interactions, for which we use the following definition for immersive environments in HMDs[2]:

> **Definition 1.1.5 Spatial Interaction**, also referred to as 3D interaction, describes tasks performed directly in a real or virtual 3D spatial context. It involves real 3D spatial input such as hand gestures or physical walking [LaV+17].

Both of the IS that we worked with contain hundreds of WA that are distributed and partially clustered throughout a relatively large area. For their selection, we used a spatial interaction technique called classic raycasting, which is an established and widely-used technique for distant object selection in HMDs [Jer15b; LaV+17] that is defined as follows:

> **Definition 1.1.6 Classic Raycasting** is a pointer-based spatial interaction technique that casts a raycast starting from an origin point into the user's pointing direction and is commonly visualized by a virtual line segment [Jer15b; LaV+17]. Target indication is specified by intersection with a 3D object and a selection is confirmed via a user-controlled input [Jer15b].

While classic raycasting is straightforward and simple to use, its performance considerably decreases for small and distant targets [Bow99; LaV+17]. Although various extensions and further improvements as well as a wide range of other spatial interaction techniques have been introduced, they each have certain advantages and disadvantages [AA13; Jer15b; LaV+17]. Therefore, interaction techniques should be chosen according to their respective properties that are appropriate within a given use case [Jer15b]. In our IIS, the selection performance of the WA might be negatively affected due to their distribution and distance from the users. And since we applied immersive technologies in the novel context of WA, this might pose other, additional requirements for interacting with the VE. One objective of spatial interaction techniques is to facilitate effective and comfortable user interaction while minimizing the impact of inherited human and hardware limitations [LaV+17]. Thus, we were determined to address spatial interaction challenges and develop and evaluate novel interaction techniques. In this dissertation, we aimed to create IIS and improve spatial interaction techniques for facilitating the virtual access and interaction with WA.

[2] Please note that this definition excludes 3D desktop interactions, where 2D mouse input is mapped to a 3D location on the display.

In the following, we will first introduce the UWA research cluster and provide an overview of its goals and structure for the study of WA. Additionally, we present the two research projects and their IS that we worked with. After establishing the cross-disciplinary scope for our work, we then address the research questions that guided this dissertation.

1.2 Cluster of Excellence "Understanding Written Artefacts"

The research described in this dissertation was conducted within the Cluster of Excellence "Understanding Written Artefacts: Material, Interaction and Transmission in Manuscript Cultures (UWA)". The UWA cluster is dedicated to the study of WA. Today, WA are an integral part of the cultural heritage that has shaped the societies of our world. Funded by the Deutsche Forschungsgesellschaft (DFG), the cluster was established in 2019 and is part of the Centre for the Study of Manuscript Cultures (CSMC)[3] at the University of Hamburg. In addition, the Helmut Schmidt University, Hamburg University of Technology (TUHH), Universität zu Lübeck, and Bundesanstalt für Materialforschung und -prüfung (BAM) are cooperation partners of the cluster. The research at UWA aims to develop a holistic, cross-cultural, and global framework for the study of all types of WA, from all regions and time periods until the present day. It follows a cross-disciplinary approach by promoting close collaboration of researchers from the humanities, the natural sciences, and computer science. The cluster is structured into eleven research fields and six working groups that focus on different dimensions of WA, including the materiality, spatiality, content, and format[4]. The eleven research fields follow Latin alphabetical enumerations:

- **A** Artefact Profiling
- **B** Inscribing Spaces
- **C** Creating Originals
- **D** (Re-)Shaping Written Artefacts
- **E** Archiving Artefacts
- **F** Data Linking
- **G** Keeping Note(book)s
- **H** Exploring Multilingual Artefacts
- **I** Formatting Multigraphic Artefacts
- **J** Situating Graffiti
- **K** Selecting Materials

The research for this dissertation was conducted within the scope of research field "B Inscribing Spaces" [Cenc] and involved two cross-disciplinary projects with historical sites containing large amounts of WA. "B Inscribing Spaces" focuses on how WA shape the spaces, landscapes, or man made structures that they are located on or in. These WA were produced as messages and statements in space and simultaneously they structured the space itself. This research field investigates WA as three-dimensional objects within their space and takes into account their placement, accessibility,

[3] https://www.csmc.uni-hamburg.de/
[4] https://www.csmc.uni-hamburg.de/research/cluster-projects

and readability, which is central for their interpretation [Cenc]. Our objective was to leverage the potential of immersive technologies for providing the relevant spatial context. Therefore, we developed and evaluated various interaction techniques and other features for two IIS based on the following research projects of two historical sites containing WA:

RFB05 "The Interior of the Church in Lucklum", which is a closed, indoor space at the Rittergut Lucklum in Lower-Saxony, Germany. The VR application includes the 3D visualization of the interior, which is covered in WA that consists of Latin inscriptions on emblematic oil paintings. The extensive WA filter features reveal how these WA structure the space.

RFB02 "Immersive City Scripts", a large-scale outdoor area consisting of the ruins of the ancient Greco-Roman (Greek and Roman) theater of Miletus in Türkiye. The VR application integrates 3D reconstructions of former structures into the digitized theater and visualizes the various categories of WA, including ancient Greek inscriptions, within a spatial and temporal context.

Figure 1.1: The interior of the church in Lucklum is covered in well-preserved emblematic paintings (photographs ©J. A. Steiger [SSA20]).

For this dissertation, we will use the project numbers RFB05 and RFB02 when referring to the respective IIS in each project.

1.2.1 RFB05: The Interior of the Church in Lucklum

The project RFB05: "The Interior of the Church in Lucklum: A Compendium of Early Modern European Emblematics"[5] was carried out from 2019 to 2022 and aimed to document and contextualize the interior of the church in Lucklum (Lower Saxony, Germany). The church houses 156 so-called emblematic paintings (monochromatic oil paintings rich in symbolism, using the European Grisaille painting technique) with 209 Latin inscriptions, which are located all across the

[5] https://www.csmc.uni-hamburg.de/research/cluster-projects/completed-cluster-projects/rfb05

interior of the church [SSA20], including the walls, the enclosing of seating areas (church pews), the gallery, and the ceiling, as displayed in Figures 1.1 and 1.2. In the project, researchers from the fields of theology, literary and art history, materials science, and HCI collaborated to document and historically contextualize the emblematic paintings and inscriptions of the church for the first time. The research focused on investigating how the paintings and WA structure the church's interior as a space of meditation and prayer. Another project goal was to explore non-invasive methods for uncovering earlier, hidden layers in the emblematic paintings that have been (partially) painted over, and digitally restoring and visualizing the findings.

(a) (b) (c)

Figure 1.2: Exemplary selection of different emblematic paintings with Latin inscriptions from inside the church: (a) Gallery, No. VI, 8, (b) Wall, No. X, 12, and (c) Ceiling, No. XIV, 15 (photographs ©J. A. Steiger [SSA20]).

1.2.2 RFB02: Immersive City Scripts

The project RFB02: "Immersive City Scripts: Inscriptions and the Construction of Social Spaces in Miletus"[6] was carried out from 2019 to 2024. Miletus was an ancient harbor city in Asia Minor (today in the Aydın Province, Türkiye) were WA, including inscriptions, were omnipresent. The inscriptions ranged from hastily engraved ancient graffiti to official, carefully carved letters. Their formal design, material, and location reveal the spatial constructions of the inscribed public spaces [BO24]. One of these impressive inscribed spaces in the city is the ancient Greco-Roman theater of Miletus, as seen in Figure 1.3. The ruins of the Milesian theater are well preserved. The theater was built upon its Hellenistic (Greek) predecessor. In the 2nd century CE (Common Era) the theater was rebuilt (Roman) and enlarged to three seating tiers and a tree story stage building. With a width of 130m and an original height of 30m, it is one of the largest theater buildings of the entire antiquity.

[6] https://www.csmc.uni-hamburg.de/research/cluster-projects/completed-cluster-projects/rfb02

Figure 1.3: The ruins of the ancient Greco-Roman theater of Miletus contain a large number of inscriptions, which were carved into the structure, e.g. on the seating stairs (main photograph ©F. Neupert, small photograph ©A. L. Osthof [Ost25]).

The general structure of ancient Roman theaters is illustrated in Figure 1.4. They were built in a semicircular shape and consisted of the following structures in Latin terminology: the *Cavea*, a spectator area with seating stairs arranged in multiple tiers, the *Orchestra*, an area for performances at the bottom of the *Orchestra*, and the *Scaena*, the stage building in front of the *Cavea* [Höl25].

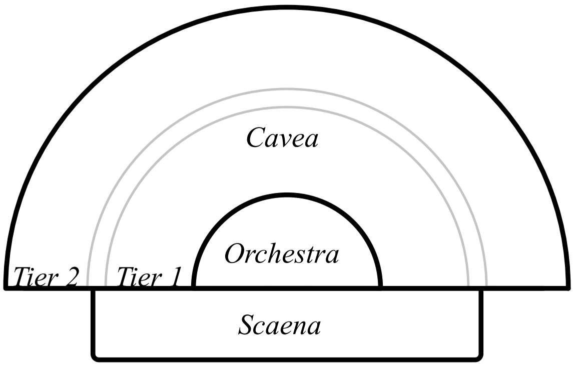


Figure 1.4: Illustrated top down view of the general structure of ancient Roman theaters.

The project combines methods from classical archaeology, ancient history, epigraphy, and HCI to investigate so-called topos-inscriptions (*topos*, "place" in Greek) that mark specific places and represent different groups of people. In this project, over 300 WA categorized into *Inscriptions*, *Game-Boards*, *Pictorial Elements* (drawings/illustrations), and *Architectural Marks* were comprehensively recorded and analyzed in their entirety for the first time. An example for each category of WA in Miletus is shown in Figure 1.5. The WA are located throughout the large, open space of the theater ruins, with some of them scattered, while others can be found in dense clusters on the seating stairs, hidden in the curved seating stair profiles, in corridors, and on different inner and outer walls. They were documented using various methods, such as sketches and drawings, photography, laser scanning, and photogrammetry.

(a) (b) (c) (d)

Figure 1.5: Exemplary selection for each of the four WA categories in the ancient theater of Miletus: (a) Ancient Greek Inscription, No. 295 (I.Milet VI 2, 936) [Her98], (b) Pictorial Element, No. 199 (Langner 2001, 2472) [Lan01], (c) Game-Board, No. 1, (Osthof 2025) [Ost25], and (d) Architectural Mark, No. 51 (I.Milet VI 2, 940 VI) [Her98] (photographs ©A. L. Osthof).

1.3 Research Questions

In this dissertation, we applied immersive technologies in the novel context of research on WA for creating interactive IIS. Our work focused on the implementation of research-oriented, interactive VR applications for the researchers who study WA, with improved spatial interaction techniques adapted to the specific context of WA. Within the scope of this dissertation, we implemented various interactive features for exploring WA, for example, by providing new perspectives, or visualizing spatial and temporal context with reconstructions and filtered views. Moreover, we investigated specific spatial interaction challenges for the IIS and the integrated WA. As this new application context comes with unique properties and requirements, this leads us to the following overarching research question for this dissertation:

Main Research Question

How can we design interactive, immersive VR applications and spatial interactions to improve the spatial understanding of WA?

Based on our leading research question, we derived further sub-questions, which consider different aspects for this novel application context. They form the basis for our research that guided our work in the individual chapters of this dissertation.

By developing IIS we do not intend to replace established on-site and analogue research methods. Instead, our aim is to explore how immersive tools can provide an additional, useful resource for researchers at the UWA cluster – especially when we consider the potential of immersive technologies for creating added value with interactive features and visualizations that are not possible in the real world. Although the main focus for the implementation of the IIS was on their usage in academia, these immersive applications can also be used for supporting science communication and for increasing awareness and interest for WA in the broader public. In order to create appropriate features and visualizations and useful interactive content for the IIS, we had to understand the researchers' goals and requirements for spatial access to the WA. Therefore, the following research question RQ-I led to our user-centered design (UCD) approach that guided our feature implementation for creating the IIS:

Research Question I (RQ-I)

How can we use immersive technologies to create interactive IIS that provide new perspectives for exploring WA?

As the core of the IIS is the spatial visualization of WA, the corresponding research data has to be integrated into the VEs. This data includes photographs, illustrations, and large amounts of textual information that researchers need to view and compare for WA at different locations. With the 156 and over 300 WA for our two IIS, this results in a substantial amount of data that needs to be displayed on spatial user-interfaces (UIs) in the form of WA information panels. Since spatial UIs are represented by 3D meshes with world-space dimensions inside the VE, they occupy a certain amount of 3D space. Taking into account the large number of WA and their dense distribution in some areas, this limits the available 3D space that one spatial information panel can reasonably occupy for displaying the research data on one specific WA. For instance, various types of textual information need to be displayed on the information panels, including transcriptions and translations of the WA, categorizations, references, and descriptions. Additionally, the research data in RFB05 includes particularly long texts. These requirements combined with the constrains for the size of the spatial UI made visualizing the WA research data a challenging task. So far, displaying large amounts of text is an uncommon use case for immersive applications and, thus, there are no established UI patterns or best practices. This leads us to our research question RQ-II:

Research Question II (RQ-II)

Which spatial UI layouts and controls are suitable for reading and navigating through long texts in immersive VR applications?

For selecting the integrated research data in their spatial context, the WA have to be marked by a 3D representation that users can interact with. These 3D markers are placed at the respective WA locations inside the IIS. The huge amount of WA are distributed over a relatively large space, with some of them loosely scattered and others densely clustered in one area. This means that researchers have to perform precise distant selections of various WA markers in different directions and at varying distances for accessing and comparing the research data. As with the spatial UI panels for the WA, we had to restrict the size of the 3D markers to minimize overlapping and occlusions in the VE.

For this distant selection task, classic raycasting with an XR controller, which is based on the pointer metaphor, is an established and commonly used interaction technique (see Section 2.3). However, selection performance significantly decreases for very small or distant targets [Bow99], since the required angular accuracy for raycasting increases with greater target distances [FHZ96; LaV+17; LG94; Pou+98]. Therefore, in situations that require the selection of distant targets, the WA markers in our IIS would be difficult to select. Furthermore, some of the researchers in our two projects have little to no prior experience with 3D or immersive applications, which might additionally contribute to difficulties with effective and precise selection of WA markers. To address this challenge, we took inspiration from *Movement Redirection*, an established concept in other domains such as VR locomotion and haptic retargeting (see Section 2.6), and applied it to raycasting interaction techniques. Guided by the following research question, we developed and evaluated our assistive raycasting-based interaction technique:

Research Question III (RQ-III)

How can we implement an assistive raycasting interaction technique that redirects the raycast towards an intended target to facilitate distant selection of spatial content?

With the technical advancements in modern HMDs, built-in hand- and eye-tracking capabilities have become more widespread and the use of novel bi-modal spatial interaction techniques such as *Gaze&Pinch* [Pfe+17] has become more prevalent. Following this development for spatial interaction design, we explored hand- and eye-tracking as input modalities in the remaining parts of this dissertation.

During the implementation phase of our VR applications for the two IIS we tested our prototypes with numerous external users at full day public events and presentations. These experiences revealed that using VR controllers might not always be straightforward for inexperienced users. In addition, it can be impractical and cause safety issues in public settings with multiple HMDs (see Section 4.6.2). This suggest that a hand-tracking based alternative for interacting with the IIS could be useful for these public settings. After investigating the effectiveness and performance of assistive raycast redirection for controller-based spatial interaction for RQ-III, our findings indicated that redirection leads to better selection performance compared to classic raycasting without assistance. Thus, we were driven to investigate whether hand-tracking can be successfully used as another modality for assistive raycasting-based selection. Moreover, we were interested in developing an additional bi-modal hand-eye-based interaction technique to further expand and improve raycasting-based spatial interaction techniques. We formulated and followed RQ-IV for extending our work on raycast redirection in this dissertation:

Research Question IV (RQ-IV)

Can we integrate hand-tracking and eye-tracking with raycast redirection to improve spatial selection?

Finally, we delved further into eye-tracking-based spatial interaction techniques. Given the growing relevance of eye-tracking, we believe that new approaches for eye-tracking based input could increase the general accessibility and user experience (UX) for spatial interactions in HMDs.

While our work so far was guided by the requirements for our IIS and the spatial interactions with WA, our proposed interaction techniques and findings can also be applied to spatial interactions in other application contexts. Therefore, we wanted to follow a more general approach for creating an interaction technique focused around eye-tracking input, that would be applicable to a wide range of immersive applications.

Existing eyes-only interaction techniques including blinks and winks were historically mainly studied as mouse emulation for desktop interfaces [KS18; Lan00; MB10a]. They rely solely on eye-tracking and do not require additional input devices or user movements for interaction. These techniques either use the gaze position, such as *Gaze+Dwell* [Han+03; Han+18; Mot+17], or gaze combined with blinking, such as *Gaze+Hold* [Ram+21]. Eyes-only interactions are a type of hands-free interaction, which also includes interactions based on other input modalities such as head movement, facial expressions, voice input, and brain-computer interfaces (BCI) [Mon+21; Mon+23]. In our research, we were particularly interested in developing and evaluating a hands-free spatial interaction technique for VEs with a focus on eye-tracking. However, we were also open to integrating head-tracking, as this input modality is already provided by modern HMDs. Within the scope of this dissertation, we will refer to our proposed eye-tracking focused interaction technique as a hands-free interaction technique.

While existing eye-tracking-based interaction techniques cover discrete interaction tasks, such as object selection or keyboard input, practical solutions for handling continuous interaction tasks, such as drag-and-drop, have not yet been established. Given this research gap for the design of hands-free eye-tracking-based spatial interaction, we aimed to create a spatial interaction technique that supports discrete as well as continuous spatial interaction tasks.

We dedicated the last part of this dissertation to creating and evaluating our novel hands-free interaction technique. Moreover, since *Gaze&Pinch* is an established hand- and eye-tracking-based interaction technique, we were interested in a performance comparison with our proposed technique. Since HMDs that implement *Gaze&Pinch* would also support our hands-free solution without requiring additional tracking hardware, our interaction technique could be implemented as an alternative alongside *Gaze&Pinch* on the same system. Motivated by these research goals, we defined the following research question RQ-V:

Research Question V (RQ-V)

Can we develop a gaze-based hands-free spatial interaction technique for discrete and continuous interaction tasks?

1.4 Outline

This chapter introduced the concept of Written Artefacts and the cross-disciplinary context of the underlying research projects as the basis for this dissertation. It explains the motivation and objectives that lead to our research question. Based on this overarching research question, five sub-questions concerning different problems and challenges were derived that guided our research.

The remainder of this dissertation is structured as follows. Chapter 2 reviews related work and the most relevant literature, and provides information on existing research and concepts for the different aspects that are investigated based on the research sub-questions. The 3D digitization process and the technical setup required for developing the IIS are described in Chapter 3. Chapter 4 addresses RQ-I and explains the development process, the implementation, and the benefits of the core features of the interactive VR applications. It reports on the identified interaction challenges and insights gained from a remote user study, field testings, and various public presentations, which guided us in our further research. This chapter also includes further ideas and prototypes for additional platforms that we explored. Chapter 5 describes the design and evaluation of different spatial UI panel layouts and controls for reading long texts based on RQ-II. Chapter 6 addresses RQ-III and reports on the design, implementation, and empirical evaluation of our proposed novel assistive raycast redirection technique for facilitating spatial selection. Building on these results, Chapter 7 further extends the concept of raycast redirection by combining it with hand- and eye-tracking, based on RQ-IV. This chapter elaborates on the design, implementation, and conducted user study for evaluating Hand-Ray redirection techniques. For investigating RQ-V, Chapter 8 describes our work on creating a hands-free interaction technique for eye-tracking-based spatial interactions, including the interaction design, implementation, and conducted user studies for our proposed solutions. Finally, Chapter 9 sums up our research results and key findings and places them in the context of our overarching research question. Moreover, this chapter presents our technical contributions, learnings and recommendations, and provides an outlook for future work.

1.5 Publications

Major parts including the main contributions of this dissertation have been published in international peer-reviewed scientific journals and conference proceedings. The publications that a chapter is based on are referenced at the beginning of each chapter. In addition, the chapters contain partially rephrased, revised, and extended content, including additional information, tables, and figures. In the following overview of publications, 🏅 indicates a Best Paper Award and 🎗 an Honorable mention for Best Paper Award.

1.5.1 Main Authorship

In the following peer-reviewed publications with me (ORCID) as the main author, my responsibilities included the ideation, conceptualization, implementation, used algorithms, experiment design, data collection, analysis, and authoring of the papers. My co-authors contributed to various parts of the papers, such as the implementation, data analysis, paper writing, or feedback.

Journal Article

- Jenny Gabel, Christof Berns, Sebastian Bosch, Jost Eickmeyer, Kaja Harter-Uibopuu, Nathalie Martin, Ann Lauren Osthof, Johann Anselm Steiger, and Frank Steinicke. "Immersive Inscribed Spaces – Bringing Virtuality to Written Artefacts for Humanities." In: *i-com Journal of Interactive Media* 21.1 (2022), pages 7–18. DOI: 10.1515/icom-2022-0012

Conference and Workshop Papers

- Jenny Gabel, Christof Berns, Sebastian Bosch, Jost Eickmeyer, Kaja Harter-Uibopuu, Nathalie Martin, Ann Lauren Osthof, Johann Anselm Steiger, and Frank Steinicke. "User-Centered Design of Immersive Research Applications for Understanding Written Artefacts." In: *2021 International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments*. (ICAT-EGVE). Sankt Augustin, Germany: The Eurographics Association, 2021, pages 31–35. DOI: 10.2312/egve.20211325
- Jenny Gabel, Melanie Ludwig, and Frank Steinicke. "Immersive Reading: Comparison of Performance and User Experience for Reading Long Texts in Virtual Reality." In: *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. (CHI EA). Hamburg, Germany: Association for Computing Machinery, 2023, pages 1–8. DOI: 10.1145/3544549.3585895
- Jenny Gabel, Susanne Schmidt, Oscar Ariza, and Frank Steinicke. "Redirecting Rays: Evaluation of Assistive Raycasting Techniques in Virtual Reality." In: *Proceedings of the 29th ACM Symposium on Virtual Reality Software and Technology*. (VRST). Christchurch, New Zealand: Association for Computing Machinery, 2023, pages 1–11. DOI: 10.1145/3611659.3615716
- Jenny Gabel, Susanne Schmidt, Ken Pfeuffer, and Frank Steinicke. "Guiding Handrays in Virtual Reality: Comparison of Gaze- and Object-Based Assistive Raycast Redirection." In: *Proceedings of the 2024 ACM Symposium on Spatial User Interaction*. (SUI). (Honorable Mention for Best Paper Award). Trier, Germany: Association for Computing Machinery, 2024, pages 1–12. DOI: 10.1145/3677386.3682080
- Jenny Gabel, Ann Lauren Osthof, Christof Berns, Jost Eickmeyer, Kaja Harter-Uibopuu, Johann Anselm Steiger, and Frank Steinicke. "Exploring Written Artefacts in Virtual Reality – Potential and Challenges for Creating Immersive Tools and Applications." In: *2025 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops*. (VRW). Saint-Malo, France, 2025, pages 25–28. DOI: 10.1109/VRW66409.2025.00013

Technical Demo

- Jenny Gabel, Sebastian Rings, and Frank Steinicke. "Ancient Theatre of Miletus: A Virtual Reality Research Tool for Studying Written Artefacts." In: *Proceedings of the 2023 ACM Symposium on Spatial User Interaction*. (SUI). Sydney, NSW, Australia: Association for Computing Machinery, 2023, pages 1–3. DOI: 10.1145/3607822.3618021

1.5.2 Co-Authorship

The following publications were mainly created by other authors and their content is not part of this dissertation. However, I contributed parts to the technical implementation, the used algorithm, the study design, the data analysis, and the paper writing.

- Judith Hartfill, Jenny Gabel, Daniel Neves-Coelho, Daniel Vogel, Fabian Räthel, Simon Tiede, Oscar Ariza, and Frank Steinicke. "Word saber: an effective and fun VR vocabulary learning game." In: *Proceedings of Mensch Und Computer 2020*. (MuC). (Best Paper Award). Magdeburg, Germany: Association for Computing Machinery, 2020, pages 145–154. DOI: 10.1145/3404983.3405517

- Jenny Gabel, Sukran Karaosmanoglu, Celeste Mason, Sebastian Rings, and Frank Steinicke. "Corona Beat - Kicking The Sedentary Habit Induced By Prolonged Social Distancing." In: *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops*. (VRW). (3DUI Contest). Virtual Conference, 2021, pages 687–688. DOI: 10.1109/VRW52623.2021.00226 [7]
- Judith Hartfill, Jenny Gabel, Lucie Kruse, Susanne Schmidt, Kevin Riebandt, Simone Kühn, and Frank Steinicke. "Analysis of Detection Thresholds for Hand Redirection during Mid-Air Interactions in Virtual Reality." In: *Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology*. (VRST). (Honorable Mention for Best Paper Award). Osaka, Japan: Association for Computing Machinery, 2021, pages 1–10. DOI: 10.1145/3489849.3489866
- Julia Hertel, Jenny Gabel, Lucie Kruse, Marcel Wollborn, and Frank Steinicke. "Co-Design of an Augmented Reality Maintenance Tool for Gas Pressure Regulation Stations." In: *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. (ISMARW). Singapore: Institute of Electrical and Electronics Engineers, 2022, pages 720–724. DOI: 10.1109/ISMAR-Adjunct57072.2022.00152
- Fabian Räthel, Susanne Schmidt, Jenny Gabel, Lukas Posniak, and Frank Steinicke. "An Evaluation of Targeting Methods in Spatial Computing Interfaces with Visual Distractions." In: *Proceedings of the 30th ACM Symposium on Virtual Reality Software and Technology*. (VRST). Trier, Germany: Association for Computing Machinery, 2024, pages 1–11. DOI: 10.1145/3641825.3687712

1.6 Preprint

Tim Rolff (ORCID) and I are both shared first authors and contributed equally to the research. While Tim Rolff's focus lied on implementing and training the deep learning model for voluntary and involuntary blink classification, my work focused on the design of the spatial interaction techniques. We were both equally involved with the implementation, user studies, data analysis, and paper writing. Our co-authors contributed to various parts of the paper, such as the implementation, data collection, data analysis, writing, and feedback. The content of Chapter 8 is largely based on this preprint article and includes partially rephrased, revised, and extended text segments as well as additional information, tables, and figures.

- Tim Rolff and Jenny Gabel, Lauren Zerbin, Niklas Hypki, Susanne Schmidt, Markus Lappe, and Frank Steinicke. *A Hands-free Spatial Selection and Interaction Technique using Gaze and Blink Input with Blink Prediction for Extended Reality*. (Author's Original Manuscript (AOM) preprint). 2025. DOI: 10.48550/arXiv.2501.11540

[7] All authors contributed equally to the paper.

CHAPTER 2

RELATED WORK

2.1 Extended Reality for Virtual Heritage

This section provides background information and context for our IIS, which will be introduced in Chapter 4. We give definitions for the relevant technologies used in this dissertation and introduce the term virtual heritage, which describes the use of immersive technologies in the field of cultural heritage. Furthermore, we provide an overview of current projects in this area.

2.1.1 Definitions

Reality-Virtuality Continuum

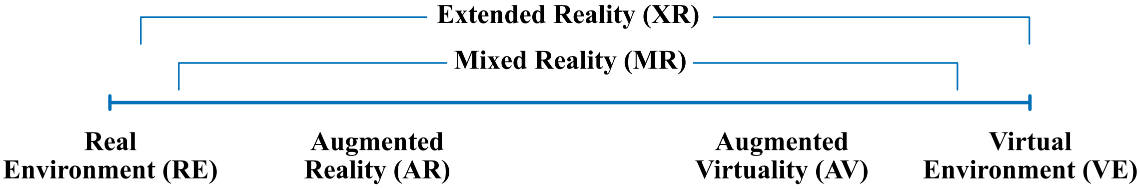


Figure 2.1: Reality-Virtuality Continuum (adapted from [MK94] and extended to include XR).

Today, the umbrella term extended reality (XR) is used to collectively refer to various immersive technologies, including augmented reality (AR), virtual reality (VR), and mixed reality (MR) technologies [Int; Kar+24]. XR encompasses the spectrum of immersive experiences, which can be described with the *Reality-Virtuality Continuum* (RVC) by Milgram and Kishino [MK94]. As illustrated in Figure 2.1 the continuum defines a spectrum ranging from real environments (RE) on the left to fully immersive virtual environments (VEs) on the right. In between, the RVC contains the space of MR, which combines real with virtual content at varying degrees. Following the continuum from left to right, the amount of virtual content increases. AR augments the real world with overlayed virtual content [Mil+95]. Augmented virtuality (AV) represents the opposite concept to AR, where a VE is enhanced with elements from the real, physical world [MK94]. VR immerses the user within a VE with virtual 3D content that they can interact with [Gig93]. Different MR technologies blend both, the physical and digital world, where users can view and spatially interact with 3D virtual content inside the real world. In this dissertation, our two IIS are implemented as VR applications. Nevertheless, we also explore further ideas and concepts for the application of AR and MR in the context of WA with simple prototypes. The majority of the development and empirical evaluation of spatial interaction techniques for this dissertation is done based on VR. In one case for our work in Chapter 8, however, we used a MR HMD.

Virtual Heritage

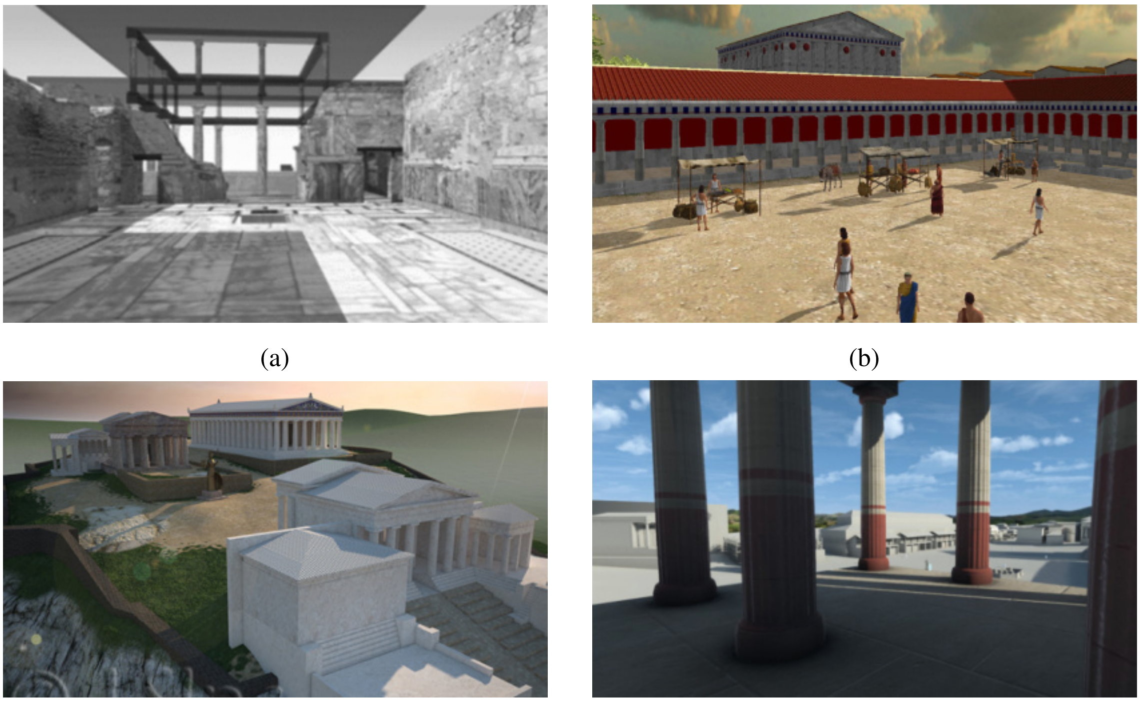
The term virtual heritage describes the cross-disciplinary application of immersive technologies such as VR, AR, and MR in the context of cultural heritage [Rou07]. Cultural heritage includes tangible physical artifacts, books, works of art, monuments, and buildings as well as intangible attributes of a group or a society that is inherited from past generations [Sul16; UNE09]. Virtual heritage focuses on the technical domain involving immersive technologies and 3D digitizations of artefacts and cultural heritage sites [San23]. This field first emerged in the 1990s after advancements in hardware and interactive 3D computer graphics made respective tools more accessible in terms of cost and ease of use [San23; Sul16]. Use cases include digital documentation, preservation, reconstructions, and visualizations of historical and cultural heritage sites. As stated by Buragohain et al. [Bur+24], by integrating technology into conservation practices, new innovative, efficient and sustainable methods for preserving cultural artifacts could be developed. Applications for virtual heritage can contribute to digital preservation, educational outreach, facilitate global accessibility, improve sustainability, and promote intercultural exchange [Bur+24].

The UWA cluster and the CSMC consider WA as written heritage that are part of our cultural heritage and cultural identity [Cena]. The work in this dissertation introduces IIS and contributes to virtual heritage by investigating and applying immersive technologies, such as VR, in the context of research on WA.

2.1.2 Virtual Heritage: Past and Present Developments

With the rise of 3D computer graphics since the late 1980s and subsequent technological advancements such as texture mapping, 3D visualizations gained the attention of humanities scholars [San23]. 3D computer graphics have since been used in various areas, including digital 3D reconstructions, historical architecture visualizations, support of archaeological field work, and scientific data exploration [MG96; Rei89; Rei91; San23]. These early 3D visualizations were typically displayed on 2D screens.

With the emergence of virtual heritage in the 1990s, 3D visualizations have been increasingly utilized in the humanities [CGC13; Par05; Rei91; San23] and interactive content such as virtual guided tours with text and images were created [San23]. In the 2000s, the project "Mediterranean



(a) (b) (c) (d)

Figure 2.2: Example of VR applications and virtual experiences over the last decades: (a) 1996: Ephesus Ruins, Türkiye (reproduced from Mazuryk and Gervautz [MG96]), (b) 2006: Miletus North Agora (north market), Türkiye (reproduced from [Foua]), (c) 2016: Acropolis, Greece (reproduced from [Lea]), and (d) 2020: Forum Romanum, Italy (reproduced from Reimann [Rei21]).

harbours, ships and sea: the invisible routes" (2000-2006) [Fouc] by the Foundation of the Hellenic World [Foue] created "A walk through ancient Miletus" [Foua] (see Figure 2.2(b)), a virtual 3D representation of the ancient city of Miletus with reconstructions, a digital city tour, and two virtual experiences using stereoscopic glasses: (i) a projection screen ("Virtual Cinema") [Foud] and (ii) a Cave Automatic Virtual Environment (CAVE) [CSD93] setup ("Kivotos") [Foub]. Projects within the last 10 years focused mainly on digital preservation and the production of virtual heritage spaces (see example in Figure 2.2) [HLH21; San23; WWC19; Woo+18]. These projects aimed to digitally preserve and expand public access to cultural heritage sites and were used in exhibitions, museums, and in the context of tourism [BBM17; Hae+16; Kyr+20; Pou+20; SS17]. One example of a touristic VR application is the self-guided tour offered at the archaeological site of Olympia (Greece) that depicts how the site might have looked like in ancient times [Oly]. Another touristic VR application shows reconstructed sites of the ancient city of Pompeii (Italy) [RA]. One ongoing long-term research and teaching project is the *digital Forum Romanum* at the Humboldt University of Berlin [Hum]. Within the project, former structures of the ancient Forum Romanum in Rome (Italy) are digitally 3D reconstructed. This project is focused on the architecture throughout the centuries. Inscriptions and smaller traces of use were therefore not included. In the VR application [Rei21] (see Figure 2.2(d)) users can experience the numerous changes to the Forum, which has been repeatedly redesigned and reinvented as a space for public communication and political representation over the course of time. Within a research project at the University of Tübingen, a VR application was created, where users can interact with avatars and ancient coins to learn about temple tax and trade at the Herod's Temple in Jerusalem [Unib]. Another recent project is the Ganjali-Khan research initiative that documents large historical landmarks in the desert city of Kerman (Iran) using state of the art 3D scanning and VR technologies [BL20]. An example of VR applications and virtual experiences over the last decades is pictured in Figure 2.2) and shows the continuous improvement in 3D computer graphics.

Moreover, new commercial and public access platforms, tools, and services for 3D visualization and interactive databases in the context of virtual heritage have emerged in recent years [Goo; Rea; Ske]. The German Archaeological Institute (DAI) has established *iDAI.world* [Ger], an open access and open source digital research platform for promoting digital research and worldwide cooperation. The platform offers tools and repositories for collecting, analyzing, visualizing, publishing, as well as storing data. This development reflects the increasing interest in digital data and in immersive technologies in this field. Another related emerging research field is the area of *Digital Epigraphy* [SR19] that explores the digitization of inscriptions and the development of novel digital tools, practices, and methodologies. The "Digital Epigraphy and Archaeology Project" of the University of Florida aims to create an open-access epigraphy toolbox with 3D visualizations. They also performed initial explorations for using AR with 3D visualizations of inscribed objects [Bar13; BB17; BBW14].

The introduced existing use cases and projects involving 3D visualizations and immersive technologies mainly focused on the 3D digitization, preservation, and visualization of historical sites and ancient structures. The interactivity is mostly limited to moving through the 3D environment and accessing further information and images from a database. In recent years, the technological advancements in HMDs and VR technology have opened up new possibilities for immersive and interactive applications in the humanities. Since the introduction of standalone VR HMDs, the use of immersive technologies has become more accessible to a wider audience, as they are less complex to setup, less error-prone, and more portable.

This creates new possibilities for the use of HMDs in immersive applications in the humanities and in the field of virtual heritage. The concept of IIS, which provide interactive VEs for displaying WA within their spatial and temporal context, has not yet been implemented and studied. Creating these immersive, interactive environments for investigating WA was one of the goals within our two research projects RFB05 and RFB02 (see Section 1.2). In this context, the application of immersive technologies for developing IIS is one research question (RQ-I) that guided the work in this dissertation. For our work, we had to identify suitable, interactive features and benefits that immersive technologies can offer in this context. These could, for example, include features that add 3D visualizations and information to provide spatial and temporal context. Consequently, it was important for us to learn about and understand the current workflows, tasks, and tools of humanities researchers such as historians and archaeologists. Thus, following an iterative UCD approach was crucial for identifying user requirements and the specific interaction challenges that the IIS and the integration of WA research data pose. These consists of reading long texts on spatial UI panels and precise selection of small and distant targets for accessing the WA data. The following sections review key literature and existing spatial interaction techniques that are relevant in the context of these interaction challenges.

2.2 Reading in Virtual Reality

For integrating the research data on the WA into the applications and visualizing it at the actual, corresponding location in the IIS, a large amount of text needs be displayed on spatial UI panels. Researchers have to read these texts in the HMDs and interact with the UI panels to explore the content and compare the information.

Since a low screen resolution leads to aliasing (visual artifacts and pixelation) of text, which negatively affects the legibility, older HMDs were not-well suited for reading [Koj+22; SR09]. However, with the technical advancements in modern HMDs and their high-quality displays, viewing and reading longer texts in VR has become feasible. Yet, there is still limited research in this area and there are no established guidelines [DKO18; Koj+20]. Given that the spatial UI panels have a 3D mesh with certain world-space dimensions, the maximum reasonable 3D world-space scale for a text panel is limited. This in turn limits the available space for displaying text and integrating UI controls for text navigation. For understanding the motivation of our work in Chapter 5, we introduce existing research for reading in VR in this section.

Previous work found that reading on printed paper resulted in significantly better reading comprehension, accuracy, and speed compared to reading on screens [Dil92; MWB13]. Wästlund et al. [Wäs+05] reported a higher cognitive load when reading and working on the computer compared to on paper within the same classroom. Rau et al. [Rau+18] evaluated reading performance for 2D desktop displays and VR HMDs. Their results revealed that participants took about 10% longer to answer multiple-choice questions when reading in VR (Oculus Rift CV1) and AR (Microsoft HoloLens) HMDs compared to reading on conventional desktop LCD screens. In their later work, Rau et al. [RZG21] found that in VR and AR (same devices as in their first study) participants read more slowly than on an LCD monitor. Their participants spent significantly more time when reading in HMDs to achieve similar accuracy to reading on LCD screens.

Kojić et al. [Koj+20] evaluated different values for text parameters and UX for reading in VR. In their user study, participants could adjust different parameters such as font size, distance, and contrast on two HMDs: (i) the Oculus Go and (ii) the first-generation Oculus Quest, and for three text samples of different length: (i) short, (ii) medium, and (iii) long. Their findings revealed significant differences for preferred values for the angular size, depending on the length of the text. The different device types had no significant influence on the text parameters. Kobayashi et al. [KKS21] investigated text parameters for reading long texts in the Japanese language. Based on their findings, they provide recommendations for suitable font type, size, and reading distance. Moreover, all their participants preferred text panels that were placed at a fixed position in the VR environment (world-space UI) instead of panels that were following the user's position and field of view. Rzayev et al. [Rza+21] observed the same preference for world-fixed UI panels for reading long texts in VR. According to their study results, they recommend displaying long texts in one paragraph. In regards to text versus background color, several studies on reading in VR all reported that light text on a dark background was preferred by the participants [DKO18; Eri+20; KKS21]. Finally, it has to be noted that these studies were conducted on older generation HMDs with lower display resolution than current, modern devices. Therefore, reading performance in HMDs might be improved by using modern devices.

As we can see from this review, existing research compared analogue reading to reading on computer displays and HMDs. Moreover, different text parameters for reading in VR such as colors, contrast, and font size have so far been analyzed. However, the question of suitable UI layouts for visualizing spatial text, and whether desktop UI patterns could be successfully transferred to spatial UIs, remains unanswered.

2.3 Classic Raycasting Interaction Technique

In this section, we introduce classic pointer-based raycasting, which is a fundamental and established spatial interaction technique for object selection in VEs [LaV+17]. In addition, we elaborate on the strengths and limitations of this technique as described in the literature.

The WA are located across a relatively large area at the two historical sites, with some of them scattered and others densely clustered. Inside our IIS the WA are represented by markers that are placed at the actual WA locations within the true-to-scale 3D models (see Section 4.3.2). The research data for the WA is integrated into the spatial context with information panels that are attached to the WA markers. In order to access and compare the data on multiple WA, users are required to perform precise distant selections of these markers. Given these characteristics of our IS and their WA, our work on spatial interaction techniques within the scope of this dissertation focuses on distant selections of small and distant 3D objects. These distant selections are referred to as *far interactions* by the software frameworks [Mica; Micb; Unia] that we used in our implementation, which we adopt and define for use in this dissertation:

> **Definition 2.3.1 Far Interactions** are spatial interactions outside of the direct interaction range and not within arm's reach of users. They are performed on any objects that users can interact with via controller-, hand-, and gaze-tracking.

There are a variety of far interaction techniques for distant spatial selection. One of the most widely used techniques for virtual pointing and spatial object selection is classic raycasting [AA13; BH97; Jer15b; LaV+17]. Many extensions and adaptations for classic raycasting exist, which are described in Section 2.4. Moreover, there are *Gaze-Based Interaction Techniques* (see Section 2.5) for distant spatial selection, as well as various techniques based on other interaction concepts, such as *World in Miniature (WIM)* [SCP95] and *Virtual Sphere* [CMS88; Han92], which use the indirect interaction metaphor [LaV+17], or the *Go-Go* technique [Pou+96], which is implemented with non-linear grasping.

Classic raycasting is based on the straightforward and easy to use pointer metaphor. A straight line representing the raycast is drawn in the 3D environment and defines the pointing direction [LaV+17]. Target indication is determined by the intersection of the raycast with a 3D object and the selection confirmation is performed via a user-controlled input [Jer15b]. In modern HMDs, raycasting can be performed with different input modalities: tracked XR controllers (Controller-Ray, Figure 2.3(a)), hands (Hand-Ray, Figure 2.3(b)), head (Head-Ray, Figure 2.3(c)), and eye-gaze (Gaze-ray, Figure 2.3(d)). The raycast is visualized as a line that is anchored to the Controller-Ray and Hand-Ray. For Head-Ray and Gaze-Ray the raycast itself is usually not visualized. Instead, often a small cursor or circular reticle [Jer15b][1] is displayed at the intersection point of the raycast and the surface of the target (see Figure 2.3(c) and Figure 2.3(d)). Since raycasting is easy to understand, simple to perform, and relatively simple to implement, it is a fundamental and widely-used spatial interaction for HMDs [Jer15b; LaV+17]. Nowadays, it is included as one of the default selection techniques on multiple established XR platforms, such as Meta Quest, Microsoft HoloLens, HTC VIVE, Valve Index, and Pico XR.

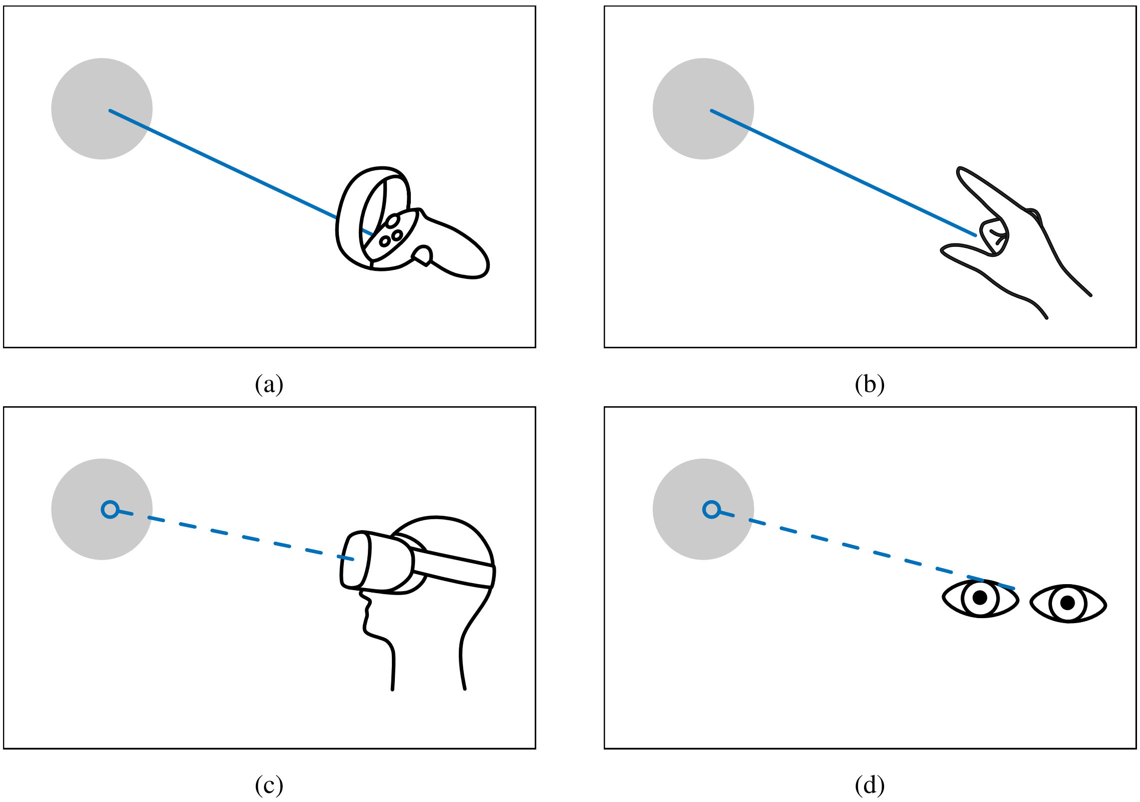


Figure 2.3: Illustrations of different raycasting modalities: (a) Controller-Ray, (b) Hand-Ray, (c) Head-Ray, and (d) Gaze-Ray.

Raycasting performs well for selections with at least medium sized objects at a close or medium range [LaV+17]. However, performance deteriorates significantly for high-precision selection tasks with small or distant objects [Bow99] due to the required high angular accuracy of pointer-based selection [FHZ96; LaV+17; LG94; Pou+98]. Selection performance is also affected by other

[1] A visual cue used for aiming or selecting objects.

parameters such as hand tremor [Bow99], tracking jitter [LG94; Pou+98], or both [Bat+21; BS19b; BS23; FHZ96; LaV+17], as well as depth perception [LG94] and visual feedback during the selection task [Pou+98]. The discrete input for selection confirmation poses another challenge due to the so-called *Heisenberg Effect of Spatial Interaction* [Bow+01]. This phenomenon describes the effect in which the required movement for a discrete input, such as pressing down a button, causes a slight shift in the position of the tracked input device, such as a VR controller [Bat+21; BS19b; Bow+01; Kop+10; Wol+20]. This, in turn, also shifts the raycast anchored to it, and the selection fails if this disturbance causes the raycast to no longer intersect with the intended selection target [BS19b; Bow+01]. Hand-tracking-based interaction can also be susceptible to this effect, due to hand movements and jitter during selection (pinch gesture) [PGG24; WBR02].

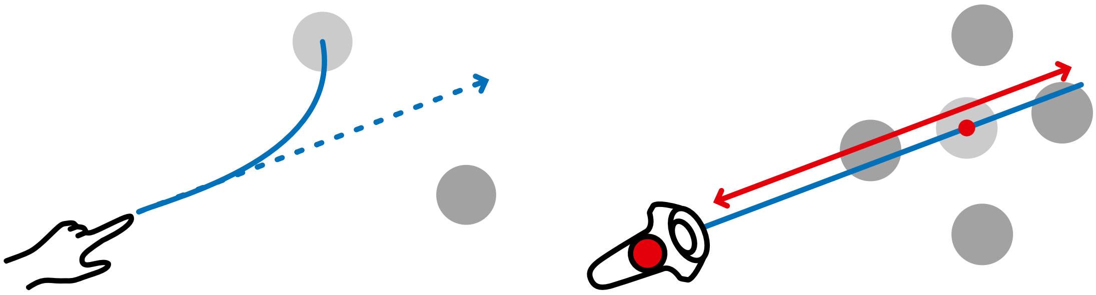
2.4 Extended Raycasting Techniques

To address the limitations of classic raycasting, various improvements and extensions have been proposed that are based on modifications of the raycast visualization [FHZ96; VGC07], the pointer behavior [Doc; SRH06], or the control mechanics [BPC19; GB06; OBF03]. Many of these techniques have been described in a survey on 3D object selection techniques by Argelaguet and Andujar [AA13]. Since then, advancements in XR and tracking technologies have opened up new possibilities for facilitating and improving spatial interaction. The following subsections provide an overview on existing techniques that extend classic raycasting.

2.4.1 Raycast Manipulation

Various approaches for facilitating raycasting-based selection by manipulating the raycast behavior have been proposed. One of these approaches is based on bending or curving the path of the raycast towards to target to support and simplify the user interaction. Snapping techniques such as *Sticky Ray* [SRH06] and Meta's *Distance Hand Grab Interactor* [Doc] both curve and instantly lock the visualized raycast to the closest target. *Sticky Ray* uses a cone-cast that is continuously increasing in size until an intersections with targets are registered. The nearest target from the initial raycast direction is then calculated and the visualized raycast is curved and snaps to the target. As illustrated in Figure 2.4(a), *Sticky Ray* visualizes the user's pointing direction with a straight raycast, while the curved raycast is always snapped to the closest target in the VE. The *Distance Hand Grab Interactor* implements a predefined conical selection frustum with a fixed size. The raycast curves and snaps to the closest target within the frustum. When there is no suitable target, the raycast visualization is hidden and continuous pointing with a visible raycast is not available. With both of these techniques, the curved raycast is instantly and automatically snapped to the center of the target. Thus, users are unable to control or adjust the raycast during snapping and cannot freely move the selection raycast between different targets to use it as a pointer.

In contrast to automatic snapping, with the interaction technique *Flexible Pointer*, the length and curvature of the visualized raycast are controlled manually based on the orientation and position of both tracked hands [OF03]. The curvature of the raycast is manually adjusted by performing hand rotations. The direction and length of the raycast are determined based on a vector formed by the position and distance of the two hands to each other. This enables users to move and curve the raycast towards the desired selection target while avoiding intersection with other objects in the VE. However, this technique requires bi-manual coordination of rotational and translational inputs, which makes performing the selections more complex.



(a) (b)

Figure 2.4: Illustrations of two Extended Raycasting Techniques: (a) Sticky-Ray (adapted from [SRH06]) and (b) RayCursor (adapted from [BPC19]).

2.4.2 Intersection Disambiguation

Another challenge for interaction techniques based on raycasting is the problem of intersection disambiguation. In cluttered or dense 3D environments the raycast can intersect with multiple targets at different depths when pointing. Different techniques that extend classic raycasting have been proposed for handling intersection disambiguation and specifying the selection target. *Bubble Cursor* [VGC07] dynamically resizes a spherical pointer around the closest target so that only one target is enclosed by the cursor. *Depth Ray* [GB06; VGC07] places a depth marker (visualized as a small sphere) on the raycast that can be controlled by moving the hand forwards and backwards. The marker acts as a cursor for target indication at the current z-depth on the raycast. *RayCursor* [BPC19] extends this idea by using the touchpad on the VR controller for controlling the depth marker, as pictured in Figure 2.4(b). The marker that can be freely moved along the raycast for accurate selection at different depths. *GazeRayCursor* [Che+23] outperforms *RayCursor* and resolves selection ambiguities by using the intersection point between Controller-Ray and Gaze-Ray for target indication, as illustrated in Figure 2.7. More details are given in Section 2.5.4.

While these techniques address the ambiguities for raycasting-based interaction techniques, they add more complexity and further abstraction to the pointer metaphor. Yu et al. [Yu+18] compared *Sticky Ray*, *Bubble Cursor*, and classic raycasting for object selection in VR. Their results showed that classic raycasting was the easiest to learn. Moreover, their findings suggest that pointing facilitators can improve performance in high-density environments, but they were only preferred when the selection tasks required efficiency and speed. When users have limited time to familiarize themselves with new selection techniques, classic raycasting with visual feedback might work best, as it already aligns with the users' mental model for pointing.

To summarize, while classic raycasting is a straightforward and widely-used interaction technique, it faces some limitations for spatial target selection. Various proposed extensions facilitate raycasting-based selection, but they come with their own limitations and disadvantages, such as increased interaction complexity and deviation from the classic pointer metaphor. Related work suggests that incorporating pointing facilitators while remaining close to the simple pointer metaphor would be beneficial to improve raycasting-based spatial interaction. Generally, many open questions and challenges remain and further research on new approaches for improving and facilitating raycast interaction techniques is required.

2.5 Gaze-Based Interaction Techniques

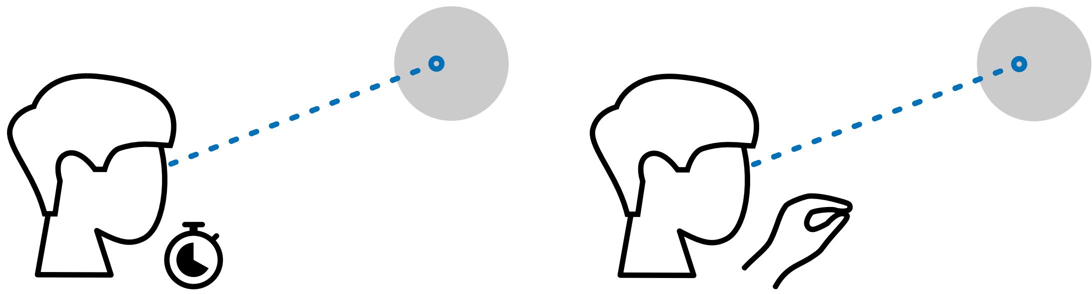
Since HMDs with built-in eye-tracking have become increasingly available, interest in gaze-based techniques has also increased. In recent years, various gaze-based interaction techniques have been developed and evaluated [Lys+22; Mon+21; Pfe+17; Plo+22; Wag+23]. While some of them only rely on eye-tracking, others use bi-modal input that combines eye- with hand- or head-tracking. Depending on the specific implementation, these techniques can either use a visible raycast or only display a small cursor at the target intersection point, as illustrated in Figure 2.3(d). This section reviews the most relevant literature and introduces existing gaze-based interaction techniques.

2.5.1 Gaze&Dwell

The *Gaze&Dwell* technique [Han+03; Han+18; Mot+17], which is depicted in Figure 2.5(a), uses gaze for targeting and gaze focus with a specified dwell time for confirming the target selection. This hands-free technique only requires gaze as input modality, with no additional input devices. Although the technique is easy to understand, the selection is slower compared to using a controller button click or a pinch gesture [MBS21]. There is a trade-off between lowering the dwell threshold (corresponding to faster dwell activation times) and an increase in error rate due to accidental selections [Mot+17; MBS21]. This is not ideal for fast interactions, such as text entry with keyboard input [Mot+17]. The text entry is either slow or, if a fairly low threshold of 300ms is used, it is prone to unintentional selections, even for experienced users [Mut+25]. Mutasim et al. [Mut+25] proposed two extensions for *Gaze&Dwell* for reducing unintentional selection: (i) *Dual-Threshold Dwell* (DTD), which increases the dwell threshold for consecutive selections of the same letter (500ms), and (ii) *Multi-Threshold Dwell* (MTD), which employs a word prediction algorithm to highlight up to three of the most likely letters for the current selection. These suggestions can be selected with a reduced threshold of 200ms, while consecutive selections of the same letter still require 500ms as with DTD. The results of their study revealed that DTD and MTD were both significantly faster and had lower error rates compared to *Gaze&Dwell* (with a constant threshold of 300ms), with MTD outperforming DTD [Mut+25]. Moreover, even novice users could already reach this performance for MTD after typing 30 phrases.

Eye&Head Dwell [SG19] modulates the dwell timer with head pointing. The dwell timer is started by focusing both, head direction and gaze on a target. The dwell timer is paused when the user looks away without moving their head and it resumes when the gaze returns. This bi-modal technique can provide more flexibility, improve input control, and allows users to freely look around in the VE, while maintaining pointing efficiency [SG19].

One major drawback of *Gaze&Dwell*-based techniques is that they do not inherently support complex interactions and continuous input such as dragging or scrolling.



(a) (b)

Figure 2.5: Illustrations of two Gaze-Based Interaction Techniques: (a) Gaze&Dwell uses a dwell timer and (b) Gaze&Pinch uses a pinch gesture for selection confirmation (adapted from [Pfe+17]).

2.5.2 Gaze&Pinch

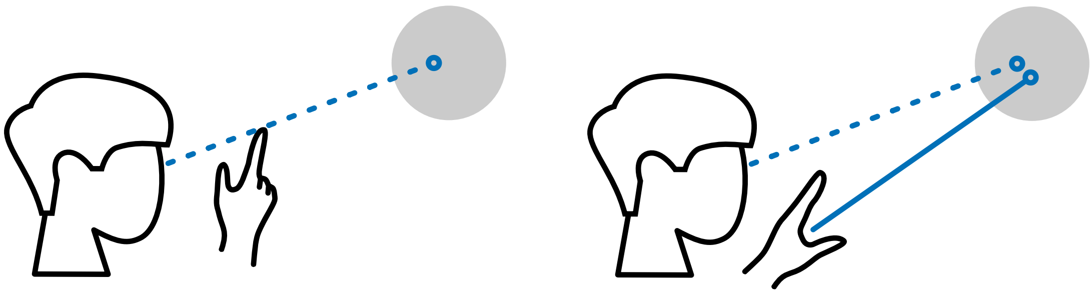
Gaze&Pinch uses gaze for targeting and hand-tracking with a pinch gesture for selection confirmation [Micc; Pfe+17], as illustrated in Figure 2.5(b). Microsoft's HoloLens 2 provides *Gaze&Pinch*, which they named *gaze and commit*, and the Apple Vision Pro, which has been released in February 2024, uses *Gaze&Pinch* as its default main interaction technique. *Gaze&Pinch* interaction can also be implemented on other modern HMDs, for example, the Meta Quest Pro and Varjo XR-4.

This interaction technique facilitates fast targeting and selection of distant objects and requires less physical effort compared to techniques that involve arm movement for targeting (e.g. Hand-Ray) [Lys+22; Wag+23]. Moreover, Lystbæk et al. [Lys+24] found that the indirect input also requires less physical effort than performing direct gestures for object manipulation. Mutasim et al. [MBS21] conducted a study on pinch, click, and dwell for selection interaction with the HTC VIVE Pro Eye and a Leap Motion device for hand tracking. Dwell caused the fewest selection errors in their study, but it was significantly slower than both, click and pinch selection. They could not find any significant performance differences between button clicks and pinch for selection confirmation.

However, *Gaze&Pinch* also faces certain interaction challenges based on five behavioral issues as described by Pfeuffer et al. [PGG24], with the following two being particularly relevant in the context of this dissertation: (i) *(Un)Learning* describes that *Gaze&Pinch* initially requires rethinking of familiar, habitual behaviors. Since users tend to intuitively move their hand towards a target that they want to acquire, they might have to unlearn this behavior when using *Gaze&Pinch* and instead, only perform the pinch gesture without changing the position of their hand. (ii) *Early and Late Triggers* can be a problem of any multi-modal interaction. For *Gaze&Pinch*, correctly matching the timing between the gaze targeting and gesture confirmation is necessary for successful interaction. The selection fails, if the pinch gesture is either performed too early or too late during a selection task [Kum+08].

2.5.3 Gaze-Hand Alignment

An interaction concept based on gaze-assisted hand-based interaction is *Gaze-Hand Alignment (GHA)* [Lys+22; Wag+23]. With GHA, a selection is triggered by aligning the fingertip or a Hand-Ray with the gaze. One interaction technique in GHA is *Gaze&Finger* [Lys+22; Wag+23], as shown in Figure 2.6(a). It is an image plane technique [Pie+97] that considers the 3D environment as a 2D image plane from the user's perspective. Target indication is performed with the user's gaze and the selection is then confirmed by aligning the index finger with their line of sight. This projects a raycast starting from in-between the user's two eyes ("cyclopean") through their raised index finger. To the user, it appears as if they are directly touching the target through the image plane [Wag+23]. While this technique eliminates the need to perform an additional confirmation gesture, it suffers from visual parallax issues at higher target distances, since users see the finger and targets at different depths [Lys+22; Wag+23]. Normally, users line up their fingertip with the line of sight of one eye, which is usually the dominant eye, but switches between the eyes can occur dynamically [BGH04]. This causes a visual shift of the finger relative to the target, which decreases pointing accuracy [Lys+22; Wag+23]. Figure 2.6(b) pictures *Gaze&Handray* [Wag+23], another GHA technique, where selection is confirmed by aligning the Gaze-Ray and Hand-Ray on the target. GHA outperforms classic Hand-Ray selection techniques for mid-sized and large targets, but the required alignment of gaze and hand input for successful selection adds more physical movement to the interaction [Azi+24; Wag+23]. These techniques might also be less familiar to users and more complex to perform due to the required gaze-hand coordination [Pfe+17].



(a) (b)

Figure 2.6: Illustrations of two Gaze-Hand Alignment Techniques: (a) Gaze&Finger and (b) Gaze&Handray both use the alignment of gaze and hand input for selection confirmation (both adapted from [Wag+23]).

2.5.4 Gaze-Assisted Pointing Techniques

Various gaze-assisted pointing techniques have been proposed for facilitating selection of small or distant targets. *MAGIC pointing* is a gaze-assisted interaction technique that instantly warps the cursor to the gaze location, where the selection can then be adjusted by manual input [FFK13; ZMI99]. This technique significantly increases pointing speed. *EyePointing* extends *MAGIC pointing* by using gaze for targeting and mid-air pointing for confirming the selections [SSM19]. *Outline Pursuits* [Sid+20] visualizes partially occluded targets with different visual stimuli that each move along a colored object outline. Users confirm their selection by matching the movement of one stimulus with their own smooth pursuit eye movement. This approach requires less physical movement compared to classic raycasting, and selection time and accuracy are less affected by increased object occlusion. *GazeRayCursor* [Che+23] resolves selection ambiguities by using the tracked gaze for estimating the depth of the desired selection target. As pictured in Figure 2.7, the intersection point between the Controller-Ray and the Gaze-Ray is used to highlight the closest selection candidate. Selection is then confirmed by pressing the trigger button on the controller. Chen et al. [Che+23] found that *GazeRayCursor* has better performance than the intersection disambiguation technique *RayCursor*. *Weighted Pointer* proposed by Sidenmark et al. [Sid+22] is a gaze-pointing technique that addresses the issue of eye-tracking inaccuracies by leveraging fallback modalities consisting of head and controller direction, as well as a relative rotational offset based on the controller or head direction. A modality weight is used to determine whether gaze input, a fallback modality, or a combination of both is used for the pointer interaction. Study results indicate that *Weighted Pointer* had higher performance and user preference compared to manually switching the input modality.

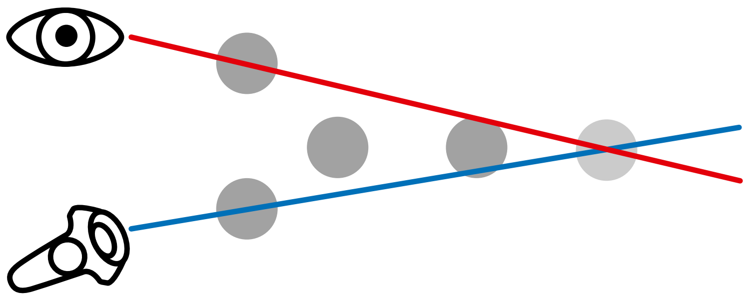


Figure 2.7: Illustration of the GazeRayCursor Technique (adapted from [Che+23]).

2.5.5 Blink-Controlled Interaction Techniques

Voluntary blinking (closing both eyes) and winking (closing one eye) can be used for hands-free interactions. Ramirez Gomez et al. [Ram+21] proposed *Gaze&Hold*, an eyes-only interaction technique with different interaction modes. Figure 2.8 illustrates one example for drag-and-drop interaction with this technique: (a) Gaze is used for target indication, (b) the closing of one eye indicates discrete input and locks-in the selected target, (c) and the movement of the open eye provides continuous input for moving the selected target. The findings of their desktop-based user study suggests that *Gaze&Hold* is effective for controlling interaction modes. Moreover, users seem to perform equally well with their dominant and non-dominant eye, since Ramirez Gomez et al. [Ram+21] could not find any significant differences in performance, usability, or workload between the eyes.

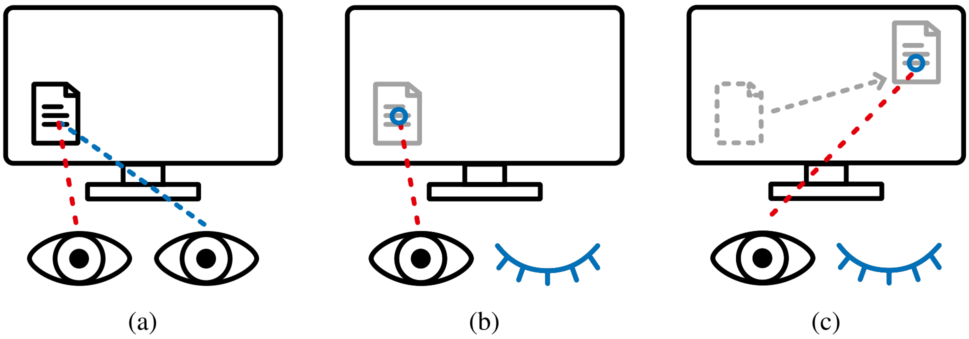


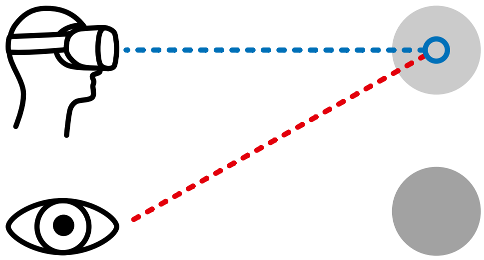
Figure 2.8: Illustrations of the Gaze&Hold Technique: (a) Looking at UI and target without input, (b) discrete selection and hold by closing one eye, and (c) continuous dragging based on gaze of open eye (all adapted from [Ram+21]).

Furthermore, blinking and winking have been proposed as an alternative to mouse input for various use cases, such as gaming [Col+24; VC16], as an accessibility feature [MB10b], and for input on virtual keyboards [AK21; Lu+20; PZ23]. Lu et al. [Lu+20] as well as Porter and Zammit [PZ23] found that blinking can improve interaction performance compared to *Gaze&Dwell*. Recent research has also explored *Blink Interaction Techniques* (BITs) in a VR gaming context [Reb+23]. In their user study on BITs, Rebsdorf et al. [Reb+23] evaluated blinking and winking for teleporting, shooting, as well as flying. Their findings indicated that the games can be successfully controlled via BITs. However, users' preferences were influenced by their affinity for specific games and by how suitable they perceived the corresponding BIT for that game. Xiao et al. [XQL19] created a VR music-on-demand system for patients with limb paralysis, where users could use blinking for searching and selecting songs. Jota and Wigdor [JW15] explored the design space of different eyelid gestures (blinking, winking, and squinting) with prototypes for activating and controlling different functionalities in mobile, desktop, and table top UIs. They argue that eyelid gestures might be most useful when combined with other input modalities and could be used for substituting or changing the behavior (mode switching) of the main modality.

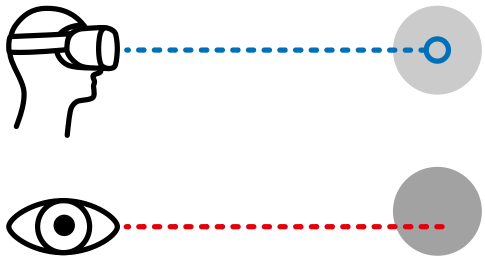
With the improved eye-tracking accuracy and increased availability of built-in eye-trackers in modern HMDs, and the numerous existing works on blink interactions, blinking and eyelid gestures appear to be a promising modality for future spatial interaction techniques.

2.5.6 Gaze-Head Interaction Techniques

Sidenmark and Gellersen [SG19] proposed *Eye&Head*, an interaction concept that uses head movements for subsequently refining gaze-based pointing in VR. This concept defines three novel interaction techniques: (i) *Eye&Head Pointing*, (ii) *Eye&Head Dwell*, and (iii) *Eye&Head Convergence*, that use head movement for modulating gaze pointing. These techniques support pointing, hovering, visual exploration around pre-selections, as well as iterative and fast confirmation for target selection. Figure 2.9 illustrates the basic interaction concept, which is implemented as *Eye&Head Pointing*. If the head moves together with the gaze, the pointer cursor is set to the gaze position, as shown in Figure 2.9(a). If the gaze shifts to another location, while the head is not moved, the cursor remains at its last position, as pictured in 2.9(b). Sidenmark and Gellersen [SG19] found that *Eye&Head* interactions might provide more control and flexibility for spatial interactions. Nukarinen et al. [Nuk+16] proposed the concept of *HeadTurn*, an interaction technique that relies on the joint usage of head rotation and gaze for system control. The gaze direction is used for target indication (selecting the control interface), and right/left head rotation is used to change the value of the associated control parameter. Nukarinen et al. [Nuk+16] tested this interaction technique for increasing and decreasing the value of numbers displayed on a desktop UI and showed that participants generally had a positive experience with using this technique.



(a)



(b)

Figure 2.9: Illustration of the Eye&Head Pointing Technique: (a) Gaze and head are on the same target and the cursor moves to the gaze position and (b) gaze shifts without head movement and cursor remains at its last position (both adapted from [SG19]).

2.6 Redirection Techniques

Parts of our work on improving spatial interaction techniques are based on using the concept of redirection to raycasting-based interaction. Hence, this section provides a brief review of existing redirection techniques that have been successfully used and studied in various other contexts.

Redirection applies gradual, subtle changes to the tracked input to manipulate the users' movements and their perception of the movement in VR. This can be implemented based on so-called "gains", a uniform or non-uniform scaling factor, which defines the mapping between the tracked movements and the corresponding visualizations inside the VE [Nil+18a; Nil+18b; Ste+08]. Figure 2.10 shows four common types of gains that are used for movement redirection when walking (locomotion) in VEs [Nil+18a].

The concept of redirection is used in different contexts, such as locomotion (*Redirected Walking*) [Lan+17; Lan+19; Ste16], tracked hand movements [EBR20; Har+21; KWB12; ZK19], and head rotation [Lan+19]. There are various techniques that manipulate the visualization of the tracked hands, such as *Redirected Touching* [KWB12], *Haptic Retargeting* [Azm+16; Che+17] as well as redirection of hand translation and rotation, which applies a gain factor to the movement [EBR20; ZKK21; ZK19]. For example, applying a gain factor of 1.5 would increase the virtual movement by 50% based on the tracked physical movement. These techniques are mainly based on mid-air hand redirection and focus on exploiting *Change Blindness* and the visual dominance

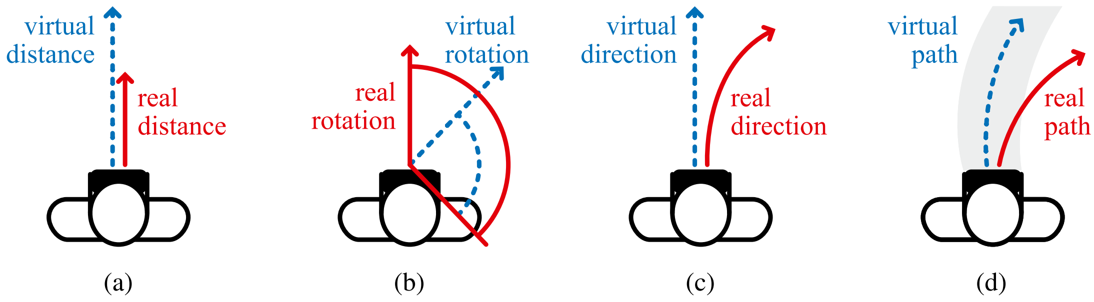


Figure 2.10: Illustrations of movement redirection gain types for walking: (a) Translation gain, (b) rotation gain, (c) curvature gain, and (d) bending gain (all adapted from [Nil+18b]).

of human perception [Gib33; GL17] to alter the perceived position in the environment (world warping) [Azm+16; Che+17; HSR18; ZK19], the virtual hand location (body warping) [Azm+16; Che+17; HSR18; ZK19], or both [Azm+16; KWB12; ZKK21]. Hand redirection is used for various purposes, including undetected manipulation or extension of the user's reach [ZK19; ZRK21] and redirecting the user's hand movements towards physical objects for creating haptic feedback that aligns with the VE [Azm+16; Che+17; KWB12; Zen+24]. Another technique is the *Go-Go* technique [Pou+96] that predates the concept of redirection but uses a similar approach for extending the user's virtual hand interaction range. *Go-Go* uses a non-linear mapping between the user's tracked hand and the visualized 3D hand model in the VE. When the user is interacting within two-thirds of their arm length, there is a 1:1 mapping between virtual and tracked position, but as soon as the user reaches further, the virtual hand position is increasingly and noticeably extended in the direction of the movement.

To our best knowledge so far, the concept of redirection has not been applied to and formally studied for redirecting raycasts for spatial selection in VR. Drawing inspiration from the existing related work, we aim to use proximity- and gain-based redirection for implementing a subtle, assistive raycast redirection technique. The redirection would help guide the raycast towards the target while the interaction remains close to the simple and straightforward classic pointer metaphor. Existing research on redirection and detection thresholds mostly focuses on keeping the redirection below a noticeable threshold. However, in this dissertation, we propose the use of raycast redirection as an assistive technique for spatial object selection. Although the implemented redirection should be subtle, it is not necessary to hide this pointing facilitation from the user. The idea of using raycast redirection is also inspired by existing interaction techniques that implement a magnetic behavior or snapping of the cursor. One example is Apple's *adaptive precision* for the iPadOS pointer [Inc] when using a mouse or trackpad. The default iPadOS pointer is visualized as a circle. When it is moved into a group of buttons, it snaps itself to the closest one while morphing its shape to fit the shape of this button and thus highlighting the currently focused UI element. The concept of *aim assist* is another example used in video games, which helps players with aiming by snapping the cursor or crosshair to nearby targets [Bat+11; Vic+14].

2.7 Fitts' Law

Fitts' Law (also Fitts's Law) [Fit54] is a predictive model for measuring user performance of selection tasks that is used in the field of HCI and ergonomics. The user's performance is determined by how fast and how accurately the movement to a target position can be executed, which is defined as *Index of Difficulty* (ID) of the given task. It is dependent on the the target width W and the movement distance A, also referred to as the movement amplitude, that specifies the distance to the center of the target. The ID is calculated according to Equation (2.1):

$$ID = \log_2\left(\frac{2A}{W}\right) \tag{2.1}$$

Fitts' Law describes this relationship as a standardized formula, which predicts the movement time MT for completing a movement to a target, dependent on A and W. Consequently, the movement time increases with greater distance and smaller target size. The original formula (see Equation (2.2)) specifies a and b as constants that depend on the specific input device and are empirically determined by linear regression analysis. The constant a represents the latency or delay of the movement and b the acceleration [GM96]. Fitts' Law has become a standard for evaluating selection techniques for traditional screens and UIs as well as for HMDs.

$$MT = a + b \cdot ID = a + b \cdot \log_2(\frac{2A}{W}) \tag{2.2}$$

2.7.1 Shannon Formulation

The original Fitts' Law formula was refined several times, one of which is the Shannon formulation by MacKenzie [Mac92a], as shown in Equation (2.3). It is the most widely used 2D extension of Fitts' Law [TL21] in the field of HCI.

$$ID = \log_2\left(\frac{A}{W} + 1\right) \tag{2.3}$$

The throughput TP is a predictive index of performance for selection tasks and conveys the "rate of information transfer" for the task [Fit54]. With the Shannon formulation, it is defined as a function of the task's ID and the movement time MT, according to Equation (2.4):

$$TP = \frac{ID}{MT} = \frac{\log_2\left(\frac{A}{W} + 1\right)}{MT} \tag{2.4}$$

2.7.2 Effective Fitts' Law Measurements

The *effective Index of Difficulty* (ID_e) incorporates both, the MT and the actual, measured selection accuracy [FP64; Wel68]. This formula uses the effective target width W_e and effective movement distance A_e, which are calculated post-study based on the actual measurements over all trials of the selection tasks with the same ID. These calculations for the effective parameters are defined as follows by Mackenzie [Mac92b] and Welford [Wel68]: W_e is obtained by multiplying S_x with the constant 4.133, following Equation (2.5). S_x is the SD of the selection coordinates x in the movement direction of the selection over all trials and indicates the distribution of the selection coordinates around the target center, as illustrated in Figure 2.11. Thus, a larger S_x value indicates higher variability and less precise targeting for the selection. The constant 4.133 was chosen based on empirical evaluation and represents the percentage of selections that miss the target [Wel68]. This is based on the fact that for normal distributed movement measurements, approximately 96% of all values fall within ± 2.066 *SD* from the mean ($2 \cdot 2.066 = 4.133$) [Cro57; Wel68].

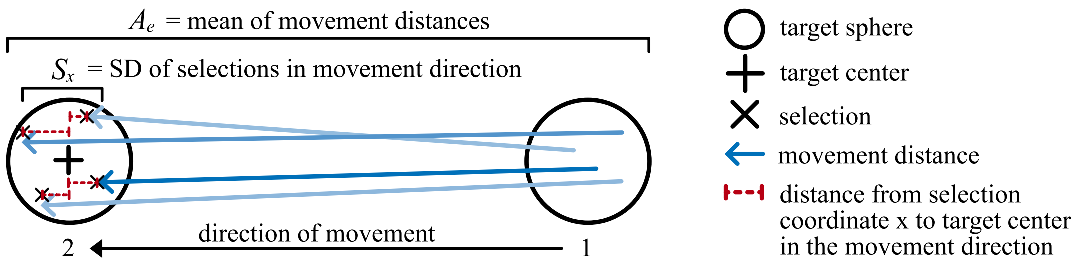


Figure 2.11: Illustrated example for calculating the effective target width W_e and effective movement distance A_e for measurements over a sequence of trials.

$$W_e = 4.133 \cdot SD_x \tag{2.5}$$

The effective movement distance A_e is defined as the mean of all movement distances for the selection task, measured across all trials, as illustrated in Figure 2.11 and calculated according to Equation (2.6):

$$A_e = \frac{1}{N}\sum_{i=1}^{N} A_i \tag{2.6}$$

where:

N is the total number of trials

A_i is the measured movement distance in the i-th trial

The effective Index of Difficulty (ID_e) is calculated using the obtained values for effective width W_e and effective movement distance A_e based on Equation (2.7):

$$ID_e = \log_2\left(\frac{A_e}{W_e} + 1\right) \tag{2.7}$$

Using the ID_e, we can then calculate the effective throughput (TP_e) according to the Shannon formulation as shown in Equation (2.8):

$$TP_e = \frac{ID_e}{MT} = \frac{\log_2\left(\frac{A_e}{W_e} + 1\right)}{MT} \tag{2.8}$$

2.7.3 Multi-Directional Tapping Task

The use of the TP_e is defined in the international standard ISO 9241-411 [Int12] (formerly ISO 9241-9 [Int00]). This standard includes the specification for a serial response multi-directional 2D reciprocal tapping task. This task uses a circular target configuration with circles or spheres that have to be selected in a predetermined sequence, where opposite targets have to be selected in each case, as pictured in Figure 2.12.

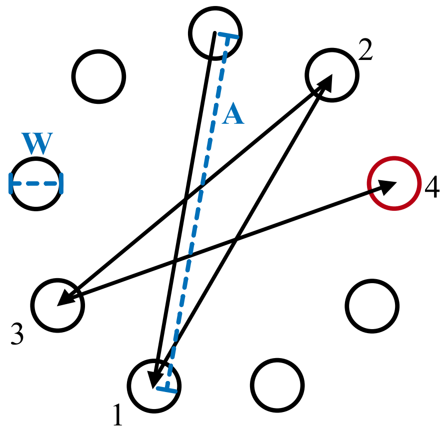


Figure 2.12: Fitts' Law ISO 9241-411 multi-directional tapping task, with movement distance A, target width W, and the first four selection targets in the sequence marked by numbers 1–4.

2.7.4 Fitts' Law for Spatial Interaction

In previous research, Fitts' Law has been frequently used to evaluate selection techniques for HMDs [Ami+25; Ber+21a]. In a review of 20 years of studies on VR object selection and manipulation by Bergström et al. [Ber+21a], 12 out of the 30 reviewed studies used a Fitts' Law selection task based on the international ISO standard. One of the earliest usage of this task in VEs was in 2011 by Teather and Stuerzlinger [TS11] for their work on spatial pointing. Amini et al. [Ami+25] conducted a systematic review of 122 XR user studies that used various adaptions and extensions of Fitts' Law. However, they found that in more than half of these studies, the Fitts' law measurements were not thoroughly examined or lacked precise reporting. Some relevant examples with raycasting-based interaction include the empirical study by Wagner et al. [Wag+23], who used this standardized task with the HoloLens 2, and Monteiro et al. [Mon+23], who used it for evaluating controller selection versus other, hands-free interaction methods in VR on the HTC VIVE Pro Eye. Batmaz et al. used the ISO task in numerous VR studies where they evaluated the effective throughput for different task execution strategies [Bat+22], investigated the influence of stereo deficiencies [Bat+22] and vergence-accommodation conflicts (VAC) on controller-raycasting-based selection performance [Bat+23a], studied the effect of raycast length [BS20] and controller grip style [Bat+23b] for spatial pointing, and tested the effects of jitter on raycasting-based spatial interaction [Bat+21; BS19b].

One criticism of Fitts' Law in the context of spatial interaction is that it cannot accurately represent 3D movements [TL21]. Extensions were developed, which try to compensate for the limitations [BS19a; CM13; MI01]. However, many previous works [Ber+21a; Wag+23] use the classical 2D formula of Fitts' Law rather than the 3D extension in their evaluations, as it can be argued that raycasting-based pointing and selection does not require actual 3D movement on the z-axis. Thus, for evaluating distant raycasting-based pointing techniques, Teather and Stuerzlinger [TS11] suggest to calculate the W_e with 2D coordinates by projecting the target width W, selection coordinates, and target center positions onto a 2D plane, as illustrated in Figure 2.13.

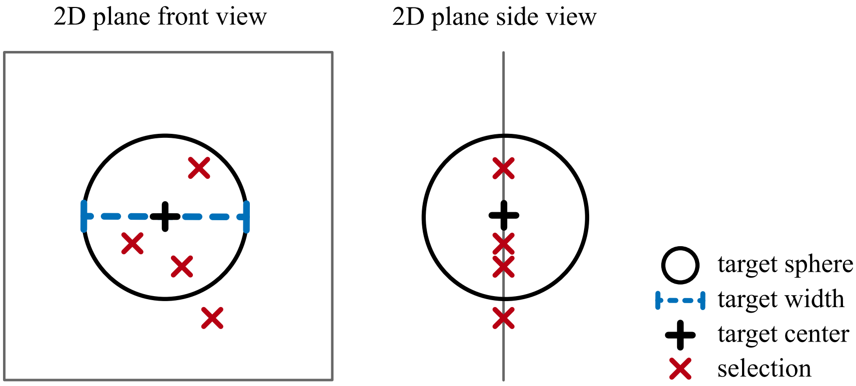


Figure 2.13: Illustrated example for 2D projection of target width W, selection coordinates, and target center position used for calculating the effective target size W_e for raycasting-based pointing techniques according to Teather and Stuerzlinger [TS11]

Fitts' Law for Distal Pointing

Kopper et al. [Kop+10] proposed another extension – the distal pointing Fitts' Law model for VEs. Their formula uses the angular width ω and angular amplitude α for calculating the ID for distal pointing (ID_{DP}). The parameter ω is the amount of angular rotation required for the movement and α is the angular size of the target from the user's perspective. The formula also incorporates a constant k as a power of ω to represent the non-linear relationship between α and ω in the ID_{DP} calculation, since the angular movement for distal pointing consists of two phases [GB05; Liu+09; Woo99]: (i) a ballistic phase with fast rotation of the wrist to move the pointer to the target region and (ii) a correction phase with slower, fine-grained movement adjustments to acquire the target, which makes up the majority of the selection time. Therefore, a value of $k > 1$ is used as a power factor to determine the weighting of ω in relation to α. The suggested value for k is 3 based on regression analysis results of Kopper et al. [Kop+10]. ID_{DP} models the increase in task difficulty as a quadratic function as defined by Equation (2.9):

$$ID_{DP} = \left[\log_2\left(\frac{\alpha}{\omega^{(k=3)}} + 1\right)\right]^2 \tag{2.9}$$

The advantage of using ω and α for the ID_{DP} calculation is that this accounts for the user's distance and position relative to the virtual display or the selection targets in the VE. Using the angular measurements allows for comparison of selection performance of targets placed at different z-distances. Kopper et al. [Kop+10] found that the increase in task difficulty is quadratic, rather than linear and that the angular width ω has a larger effect on MT than the angular amplitude α. Various existing works have used ID_{DP} in combination with the ISO 9241-411 selection task in VEs [Bat+22; Bat+23a; BS20; LWW23].

In the user studies conducted in the scope of this dissertation, we used the Shannon formulation (see Section 2.7.1) and the multi-directional Fitts' Law selection task (see Section 2.7.3) for the standardized empirical evaluation of our proposed spatial interaction techniques in Chapter 6 and Chapter 7. For analyzing the selection performance, we followed Section 2.7.2 and calculated the effective Fitts' Law measurements and effective throughput TP_e. We used the proposed 2D projection in Section 2.7.4 for calculating the effective target size W_e, because our evaluated techniques are based on distant pointing with raycasting. In Chapter 7, we evaluated targets at different depths and, therefore, chose the extension for distal pointing, as described in Section 2.7.4 for calculating the ID_{DP}.

CHAPTER 3

TECHNICAL FOUNDATION FOR IMMERSIVE INSCRIBED SPACES

3.1 3D Digitization of Inscribed Spaces

An essential basis for our further work and research in this dissertation was to first digitize both IS for RFB05 and RFB02 to obtain detailed, true-to-scale 3D models. We had specific requirements for 3D digitization that made the process more challenging. For achieving a realistic representation of both spaces, two components for 3D digitization have to be considered:

1. 3D Mesh: Defines the underlying geometry and consists of polygons, most commonly triangles in modern 3D engines. It represents the shape and structure of a 3D model. The geometry and structures of the inscribed spaces have to be precisely captured.
2. Texture: Contains visual color information stored in so-called color or diffuse maps. Other types such as normal and specular maps can be used to add further surface detail and control light reflections. The inscribed spaces and WA should have a photo-realistic appearance and coloring.

After obtaining both components, the resulting textured 3D model had to be further optimized for efficient real-time rendering in VR. In the following subsections, we will describe the technical background and the process for digitizing both inscribed spaces in detail.

3.1.1 Data Fusion for 3D Object Reconstruction

There are different active and passive systems for 3D object reconstruction of real-world objects [Luh+20]. Active systems use structured light to project illumination patterns for measuring the surface of 3D objects and reconstruct their 3D geometry. One common active system is 3D laser scanning with Light Detection And Ranging (LIDAR). This system delivers 3D point cloud data with high spatial accuracy and depth measurements with millimeter-level precision. It is not negatively affected by low light conditions and is therefore suitable for outdoor 3D data capturing. While LIDAR data provides a true-to-scale geometric structure with accurate measurements, it lacks high-resolution color information required for detailed texturing of the 3D models. Although there are scanners with built-in RGB sensors, their resolution is usually limited.

Passive 3D object reconstruction systems capture visual information based on the matching of image features. One common and established approach using image data is photogrammetry. This system requires a relatively large amount of photographs that have a certain amount of overlap. Based on the overlapping parts in different photographs, matching features across the images can be identified. With these features, the camera positions can be reconstructed and a 3D point cloud can be obtained. With a suitable, high-quality digital single-lens reflex (DSLR) camera, the captured image data provides high-resolution RGB color information that can be used for creating detailed textures with realistic coloring. Moreover, drone-based imaging can be used to capture data in areas and from view points that are not accessible otherwise. However, capturing large amounts of photographs is time consuming and especially challenging in outdoor settings with potentially bad or changing lighting conditions and hard shadows, depending on the time of day.

To leverage the strengths of both techniques while compensating for their individual limitations, an approach called multi-sensor data fusion is used. It combines the 3D LIDAR scanning data with 2D images from drone imagery and DSLR photography to create a detailed, true-to-scale 3D model and high-resolution textures with realistic color reproduction. There are different approaches for combining the data. In our projects, the LIDAR scanning data and the image data were used for creating 3D point clouds and the combined 3D mesh. The mesh was then textured based on the drone and DSLR data. Figure 3.1 illustrates the major steps of this 3D digitization process.

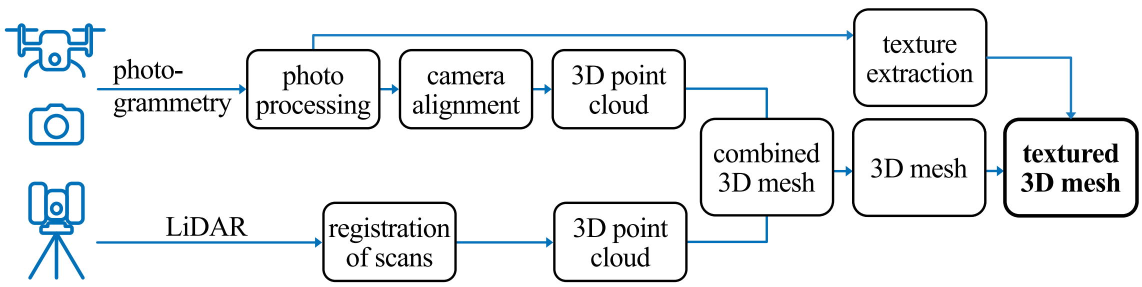


Figure 3.1: Step-by-step 3D digitization process based on data fusion of photogrammetry and LIDAR laser scanning data for obtaining a textured 3D mesh.

For both of our two IIS, we started with a limited amount of available 3D data of the inscribed spaces. After we experimented with the provided 3D data, we found that the quantity and quality of the 3D data was insufficient for creating textured 3D meshes that meet our requirements. Therefore, we used multi-sensor data fusion based on 3D LIDAR scanning and 2D image data from drones and

DSLR-photography for creating the 3D models of our IIS. The following subsection outlines the individual 3D data acquisition and 3D model processing procedure for both of our VR applications for RFB02 and RFB05.

3.1.2 3D Data Acquisition, Digitization, and Modeling

RFB05: Church in Lucklum

For RFB05 we initially used 3D LIDAR data provided by external project partners from the Institute for the History of Construction of the Technical University of Braunschweig (TU Braunschweig, Germany). In 2021, we extracted 3D point cloud data based on the scans with the 3D software Autodesk 3ds Max[1] and created 3D meshes of the church's interior with Agisoft Metashape[2]. However, the 3D scan data did not include detailed geometry of certain areas such as the pulpit and did not provide sufficient color information for creating high-resolution textures with high color fidelity. For this IIS, especially the intricate and colorful emblematic paintings had to be represented as photorealistically as possible. Since the laser scans were captured for use in technical and structural building analysis, they were not focused on capturing the small structures and fine details of the emblematic paintings. Moreover, the point cloud data did not provide enough high-density color information for properly texturing the 3D mesh.

Figure 3.2: (Left) Initial 3D model of the church's interior based on old laser scans and (right) new final 3D model based on a mixed approach combining laser scanning and photogrammetry data.

Following these initial test results, we decided to conduct a small, one-day exploratory 3D test scanning of different parts of the church's interior. In 2022, we visited the church together with the cluster's 3D digitization engineer, equipped with the university's LIDAR scanner and a DSLR camera. Our goal was to test the combination of laser scanning for obtaining accurate geometry, and

[1] https://www.autodesk.com/products/3ds-max/
[2] https://www.agisoft.com/

photos for detailed, high fidelity texturing of the 3D model. We captured overlapping photographs from different angles, which were then used to extract color information (photogrammetry) for texture creation. We created textured test meshes of small parts of the interior based on the laser scanning and image data. This test run helped us to identify and define the requirements for a more detailed 3D data acquisition of the church's interior and its emblematic paintings. We concluded that we could not achieve the required quality and fidelity for the 3D models with our own limited knowledge and the hard- and software available at our university. Thus, we commissioned ArcTron 3D[3], an experienced provider that specializes on digitization of historical and cultural sites, for the 3D scanning and processing of the data.

Figure 3.3: (Left) 3D LIDAR scanner and (right) coded target used as references in the scanning process by ArcTron 3D.

(a) (b)

Figure 3.4: Detailed comparison of initial vs. final version of 3D model and texture quality of the church in Lucklum: (a) Emblematic painting before (left) vs. after (right) and (b) church's pulpit before (left) vs. after (right).

[3] https://www.arctron.de/home/

In total, 151 laser scans and 10.711 pictures were captured with calibration for position (coded targets and ground control points) and colors. Examples of the technical scanning setup by ArcTron 3D are pictured in Figure 3.3. First, the data was post-processed, including color and lighting corrections of the photographs. Afterwards, the data was combined, meshed, and textured using RealityCapture[4]. The final 3D model and texture files were delivered in a high-resolution (15 mio. triangles, 82×8k textures) and a reduced low-resolution version (2 mio. triangles, 15×8k textures) that is suited for real-time 3D rendering in VR. To ensure a stable frame-rate on the Meta Quest 2, we further reduced the 3D model to 600k triangles. The obtained results confirmed that multi-sensor data fusion with laser scanning data and high-quality image data provides the best results. As shown in Figure 3.2, the 3D mesh as well as the colors are significantly more detailed and realistic. These major quality differences can been seen in more detail in the close-up views in Figure 3.4. The side-by-side comparison of the final 3D model in the VR environment with a photograph of the interior of the church in Figure 3.5 shows the realistic representation and coloring.

Figure 3.5: (Left) final 3D model and (right) photograph of the church (photograph ©J. A. Steiger [SSA20]).

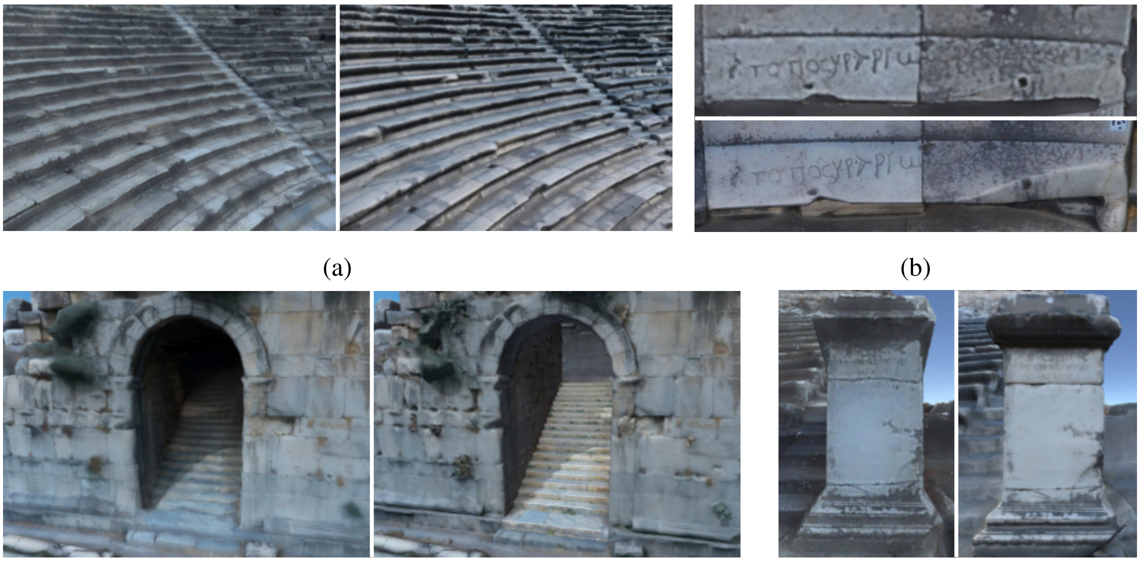
RFB02: Theater of Miletus

In RFB02 we collaborated with the laboratory of photogrammetry[5] at the Berliner Hochschule für Technik (BHT Berlin, Germany) for digitizing the ancient theater of Miletus. The initial data capturing was carried out in 2019 and an early 3D model with low mesh and texture resolution was created in Agisoft Methashape. This initial model was mainly based on drone imaging of the *Cavea* and a small amount of LIDAR scans of the remains of the stage building. Due to the limited amount of captured data, the resulting model was of insufficient quality. As visualized in Figure 3.6(a), the shape and curvature in the profile of the theater's seating stairs were distorted and blurred out. The same loss of detail and shape in the 3D mesh can be seen in Figures 3.6(c) and 3.6(d). In addition, the inscriptions in the texture were difficult to read or even illegible (Figures 3.6(b) and 3.6(d)) due to the low quality and resolution of the captured image data.

To resolve this issue, we conducted two data capturing campaigns in 2021 and 2023 for collecting large amounts of additional data. Due to the large size of the outdoor space it was not

[4] https://www.capturingreality.com/
[5] https://labor.bht-berlin.de/photogrammetrie

possible to capture the whole theater in very high detail in the given time. Moreover, the real-time rendering performance was limited by the hardware of the Quest 2. Therefore, we decided to select specific areas that needed to be captured with particularly high resolution. We choose areas and structures with relevant details in the mesh and important WA and briefed our 3D surveyor accordingly.



(a) (b) (c) (d)

Figure 3.6: Detailed comparison of initial vs. final version of 3D model and texture quality of the theater of Miletus: (a) Mesh profile and texture details of the seating stairs of the theater's first tier before (left) vs. after (right), (b) readability and details of inscription before (top) vs. after (bottom), (c) mesh profile and texture details of an entrance portal and its stairs before (left) vs. after (right), and (d) mesh and texture details of the statue base before (left) vs. after (right).

As recommended by the BHT team, we also followed a data fusion approach using LIDAR scans, drone imaging, and DSLR photography for this IIS. The data capturing included more densely placed LIDAR scans as well as scans of the inner areas, tunnels, and staircases of the ancient theater that were not captured before. In addition, a large amount of drone images from both remaining seating tiers were captured and close up DSLR camera images of the most interesting, important, or prominent WA were taken. The 3D data capturing was calibrated and the scans were geo-referenced. Given the large size of the theater and the required level of detail for specific structures and WA, the data capturing was conducted in two campaigns. The outdoor setting added additional challenges compared to the interior scan of the church in Lucklum. The changing weather conditions and the time of day influenced the shadows cast on the structures and the number of visitors in the theater. This left us with only a relatively short window in the early morning for the data capturing. The meshing and texturing of the 3D model were done in Agisoft Metashape at the photogrammetry lab of the BHT. The final version of the digitized theater was split into several parts and reduced for use in real-time rendering on the Meta Quest 2. In total, 14 textured 3D meshes

Table 3.1: Triangle and texture counts, 3D models theater of Miletus

Description	**Triangle Count**	**Texture**
1st seating tier	299k	8×8k
2nd seating tier	449k	6×8k
2nd seating tier special section	100k	4×8k
west corridor	200k	2×8k
east corridor	199k	2×8k
stairs leading to east corridor	25k	1×8k
castle building	148k	2×8k
outer walls and structures	79k	2×8k
detail, inscription 283 and 284	23.6k	1×8k
detail, inscription 282	14.3k	1×8k
detail, inscription 290	9.6k	1×8k
detail, inscription 291	14.8k	1×8k
detail, inscription 295	14.6k	1×8k
detail, inscription 332	6.2k	1×4k

were created, as listed in Table 3.1. The main part of the theater consists of 8 models representing the 1st and 2nd seating tiers, inner corridors, and structures of the outer part. In addition, 6 small, detailed models (chunks) of specific areas and structures were exported with higher relative triangle counts and a single high-resolution texture each. They include, among others, an inscribed base for a former statue (no. 295) and a stone slab of a wall with a particularly important inscription (no. 332). The before and after comparisons in Figure 3.6 show the substantial improvements in mesh and texture quality. Figure 3.7 provides a side-by-side comparison of the final 3D model of the theater and a photograph of the real inscribed space.

Figure 3.7: The theater of Miletus: (left) final 3D model and (right) drone image of the real theater (photograph ©F. Neupert)

3.2 Technical Setup and Configuration

3.2.1 Hardware

We used the Meta Quest 2[6] HMD as our main target platform for developing and testing both VR applications. Since the Quest 2 is a portable and lightweight standalone HMD, it does not require a high-performance computer, additional trackers, or a complex hardware setup. Therefore, it was easy for the humanities researchers to use this HMD by themselves. This also allowed us to send out the HMDs to each project partner during the COVID-19 pandemic.

3.2.2 Development Engine, SDKs, Packages, and Frameworks

The IIS were developed with the Unity engine[7]. Initially, we used Unity's LTS Version 2019.4, which we updated to LTS 2021.3 during the later stages of the project. The Unity project was configured with Unity's XR Plug-in Management package.

For the XR interaction implementation we decided to use Microsoft's Mixed Reality Toolkit 2 (MRTK 2)[8]. MRTK 2 offers multi- and cross-platform support, which makes it easier to port the developed applications to other devices in the future. The toolkit comes with various XR functionalities for input handling and UI setup, which we customized for our implementation, including raycasting with pointer visualizations, teleportation for locomotion, and templates for interactable 3D user interfaces elements such as buttons and toggles. For creating the user interface, we used MRTK's 3D UI elements and Unity's built-in world space canvases for 2D spatial UIs. In addition, 2D sprites were used to overlay images onto the 3D models. Since the Meta Quest 2 was our target HMD, we also used the Oculus Integration SDK (now deprecated)[9] that provided us with Quest platform specific functionality and 3D models for the Quest VR controllers. We configured the MRTK input data provider with the corresponding XR SDK Oculus Device Manager of the toolkit and the Quest 2 controller models.

As Microsoft Excel is a familiar tool that the humanities researchers use to collect and store information on the WA, we integrated an internal database asset[10] in Unity that can directly import the WA data from Excel files. We created Unity Editor scripts to extract the information from the database and automatically perform a specified setup within the virtual environment. This includes the creation and placement of 3D markers and spatial information UI panels, as well as setting up the content including texts and images for all the WA.

3.2.3 XR Controls and Interactions

During the development period of the IIS, there were no comparable research focused VR applications for exploring and analyzing WA that we could use as a reference. As there are also no established best practices or patterns for designing fundamental interactions such as object selection

[6] https://www.meta.com/de/en/quest/products/quest-2/
[7] https://unity.com/products/unity-engine
[8] https://learn.microsoft.com/en-us/windows/mixed-reality/mrtk-unity/mrtk2/?view=mrtkunity-2022-05
[9] https://developers.meta.com/horizon/downloads/package/unity-integration/
[10] https://www.bansheegz.com/BGDatabase/

and locomotion in this context, we based the implementation on commonly used patterns and interaction techniques found in VR games and applications [Jer15a].

The distinct properties of both IIS bring specific challenges for choosing and designing suitable VR interactions. The ancient theater of Miletus is a large and open outdoor space with single, scattered, as well as densely clustered WA. They are placed throughout the theater, such as on the seating stairs, hidden in their profile, in corridors, and on different inner and outer walls. In comparison, the church's interior in Lucklum consists of one main room with a high ceiling, a two-sided gallery on the first floor, and many occluding structures. The emblematic paintings with Latin inscriptions are placed throughout the church, including the pulpit, the gallery, church pews (closed seating areas), walls, and on the ceiling. Since users have to interact with WA that are placed at different distances and located in different directions, we chose to use raycasting-based pointing for selection interactions. As described in Section 2.3, it is an established and straightforward pattern based on the pointing metaphor. Since pointing is a familiar concept, we assumed that users with no prior 3D or VR interaction experience would be able to derive the corresponding mental model for this selection technique and use it without major issues. For locomotion in VR, we used two approaches: (i) The Walking Pattern based on real, actual walking with room-scale tracking [Jer15a; Ste+13] is used for moving across small to medium distances and ensures presence and ease of navigation in the VE [Uso+99] and (ii) the raycasting-based Teleportation Pattern for covering large distances via teleportation of the user to the visually indicated pointer surface intersection point [Jer15a]. We combined the teleportation with a screen fade to soften the instant relocation and prevent disorientation and motion sickness.

CHAPTER 4

USER-CENTERED DESIGN AND IMPLEMENTATION

Major parts of this chapter are based on several publications [Gab+21a; Gab+22; Gab+25; GRS23]. Large parts of the texts were rephrased, revised, and extended, and new sections and paragraphs with additional figures and tables were incorporated.

4.1 Motivation

The work in this chapter focuses on answering RQ-I:

RQ-I
How can we use immersive technologies to create interactive IIS that provide new perspectives for exploring WA?

Traditionally, field research was the only way to obtain spatial information and to investigate the WA within their spatial context. Common documentation, analysis, measuring, and visualization methods mainly involve textual notes and 2D imaging, such as photographs, sketches, and drawings. Although spatial context is essential for studying WA, these traditional methods only provide limited or relatively abstract recreations of the 3D properties of the inscribed space. Due to the cultural significance of WA, increasing efforts have been made to catalog, digitize, and provide access to WA from all over the world [Bau; CC; Ger; Liu19; SR19]. As interest in this area increases, there is also growing interest in exploring the potential of using new, innovative technologies and tools in the context of WA. Immersive interactive applications could offer truly new perspectives and enable new ways for researching WA.

This chapter presents our work on the development of the interactive features for the novel application context of IIS in our two projects RFB05 and RFB02. It also reports on our various additional prototyped ideas, our presentations and demonstrations of the applications, the user testing, and user feedback. Our aim was to identify requirements of the researchers and explore how we can built interactive immersive research environments. To achieve this objective, it was essential to ensure that we design and implement interactive tools and features that would support and benefit research on WA. Instead of only providing a basic, digital 3D replica of the historical sites, we wanted to provide researchers with true-to-scale interactive VEs of the IS that they can freely explore and interact with. The applications should give them spatial access to the WA and offer useful features that extend beyond the possibilities of on-site field research.

4.2 User-Centered Design Approach

In both RFB05 and RFB02 it was essential to follow a UCD approach [Jer15b; ND86; SPR19] to ensure the successful application of immersive technologies for the research on WA in the humanities. The humanities researchers in both projects were the target user group of the IIS. They did not have previous experience with immersive technologies and VR. Thus, they had little to no experience with 3D environments, 3D navigation, and spatial interaction techniques. The two projects were also the first introduction of immersive technologies for researching WA at the UWA cluster. Therefore, the potential for interactive applications beyond simple 3D visualization and the technical limitations of immersive technologies was not well understood by the project members. At the same time, we had no prior knowledge about the requirements, challenges, and pain points of the researchers during their work in researching WA. We were also not aware of their expectations regarding the use of immersive technologies.

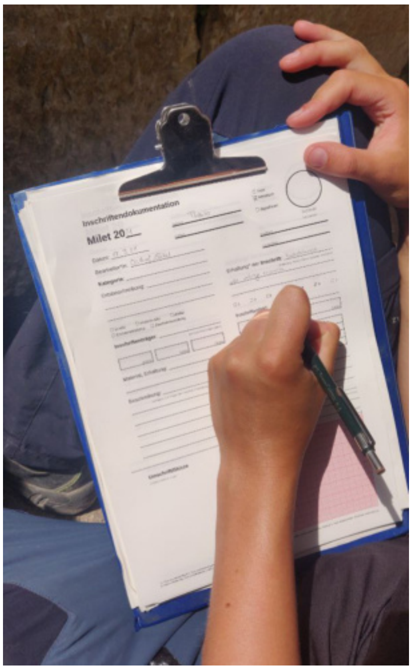
Following an iterative UCD approach allowed us to analyze and understand the researchers' needs and define appropriate requirements for the development of the IIS. Together with the researchers from the humanities, we discussed, explored, and tested various ideas and new approaches for studying WA in a virtual, immersive environment. The researchers tested the visualizations and interactive features directly on the Quest 2. Based on their feedback, we iteratively developed, extended, and refined the features. This collaborative and iterative process also enabled them to experience VR technology and gain a basic understanding of its advantages and potential. This in turn facilitated discussions and the ideation of further ideas. Using this hands-on approach combined with ad-hoc and fast testing allowed us to actively involve the researchers in the development and improvement of the implemented features. During the main design and development process we held weekly to bi-weekly meetings within a core team consisting of the doctoral and post-doctoral researchers. These meetings not only provided us with valuable user feedback and bug reports, they also served as important opportunities for brain-storming sessions and exchange, which facilitated understanding between our different research fields. For major updates and extended feedback, we arranged quarterly meetings involving all project members. During our work on the IIS, the humanities and material science researchers provided us with research data and their research results on the WA. This data includes, among others, the identified and recorded WA, images of emblematic paintings, special scans of hidden layers, photographs, translations, descriptions, and categorizations.

4.2.1 Requirement Analysis

Figure 4.1: Doctoral researcher and archaeologist in our project, conducting field research and recording an inscription in the ancient theater of Miletus (photograph ©A. L. Osthof).

We conducted interviews and made on-site observations with researchers from both projects to gain an understanding of the context and current approaches used for researching WA and IS. The researchers work in the fields of classical archaeology, epigraphy, ancient history, literary studies, theology, and material science. The researchers from the humanities reported that their current research methods are mainly analogue. This includes on-site work, as shown in Figure 4.1, with manual collection and documentation of data, as depicted in Figure 4.2(a). Other methods involve referencing existing literature, doing archival work, and recording information digitally as notes, drawings, illustrated 2D maps, and photographs. Since there are no dedicated tools for digitizing, visualizing, and analyzing the data, commonly used software such as image editing and vector illustration applications, as well as Microsoft Office applications are used. Although their work on-site happens within a spatial 3D context, when it comes to organizing and structuring the collected data, everything is still processed in 2D. The only analogue method for documenting 3D properties of WA, such as carved inscriptions, is to create analogue, reverse 3D copies with so-called squeezes. In order to create a squeeze, a special type of paper is usually wetted with water and layered on top of an inscription. The paper is pushed into the indentations of the inscription with a stiff brush, as demonstrated in Figure 4.2(b). After the paper dries out, a reversed 3D imprint of the inscription is obtained.

While squeezes capture the 3D properties of single carved inscriptions, there are currently no digital tools for spatially visualizing a collection of them with their respective placements and distribution



(a)

(b)

Figure 4.2: Recording and documenting WA in the field: (a) Handwritten documentation and (b) creating a 3D squeeze with special paper (photographs ©A. L. Osthof).

within the inscribed space. By using immersive technologies, the researchers hope to gain 3D access for analyzing their data when they are not working on-site. The focus was to develop IIS that would provide the missing spatial access and offer novel perspectives and interactive features for the analysis of WA – especially for visualizations and features that are not possible or available on-site in the real, physical environments. Based on this, we identified the most important requirements and derived potential core features for the IIS in both of our projects:

- Realistic, true-to-scale 3D models of both IS for creating the IIS
- Spatial visualization of WA data in 3D space:
  - display of extended information such as transcriptions, translations, categorization, photographs, and descriptions
- Structuring of WA data including comparing, sorting, and filtering
- Adding temporal context:
  - Reconstruction and visualization of former structures and hidden layers
  - Time of day simulation to visualize the influence of light and shadow
- Providing new perspectives and views of inaccessible areas
- Visualizing former seating capacity
- Interactive, free exploration of the IIS
- User-friendly UI menu for selecting the features

In the following section, we will describe the implementation of these interactive features for both IIS in detail.

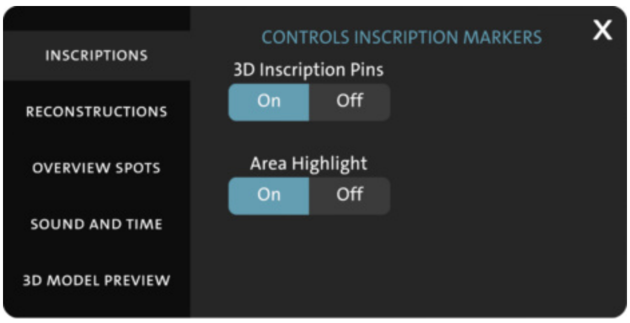
4.3 Interactive Feature Implementation

Based on the identified requirements and input from the researchers, we implemented various features to help researchers gain new perspectives in the IIS and to study the WA with spatial and temporal context. Moreover, we developed tools for visualizing, analyzing, and categorizing the WA. The menu labels and texts for RFB05 are in German as the content for this project is based on the German publication *Sinnbilder im Sakralraum. Die Kirche in Lucklum – Ein Kompendium der geistlichen Emblematik der Frühen Neuzeit.* [SSA20]. For RFB02, the ancient theater of Miletus, all content is in English. The research data for RFB02, including the data on the WA, were provided by Ann Lauren Osthof and are part of her research results for her dissertation [Ost25].

4.3.1 Main Menu User Interface

First, we will introduce the main menu setup and layout. The main menu provides access to all the implemented features. Our intention was to create an easy-to-use spatial 3D UI that displays all available features at a glance. We observed that users would lose track of the main menu while exploring the 3D environment or completely forget about it. Consequently, we tested multiple iterations for placing and anchoring the main UI to the user's position. As illustrated in Figure 4.3, in our old UI design we attached the main menu to a 3D wrist watch and anchored it to the tracked controller position. We choose to use a menu layout based on common desktop menus with a left side tabbed navigation. However, users did not like the controller wrist anchored UI. Many had difficulties with pointing at the wrist watch and selecting it with the controller in their other hand. In addition, this UI setup would always require the use of two controllers for tracking the position of both hands. Figure 4.4 depicts our updated design. For both IIS, we created a horizontal 3D toolbar that is anchored at a fixed distance in front of the user and placed in the lower part of the user's field of view. The main menu bar can be minimized to a small square button by selecting the "X" button on the right side of the tool bar. We used MRTK's 3D UI and button templates with physical, 3D mesh based spatial UI elements. The 3D UI buttons have a *button press animation* and are pushed in when they are selected. They also display individual interaction states with different coloring. Depending on the specific feature, a sub-menu is toggled when a button is pressed.

(a)



(b)

Figure 4.3: Old main menu designs and layout: (a) 3D wrist watch menu and (b) left side tabbed navigation menu.

Figure 4.4: Main menu toolbars for (top) RFB05, Lucklum and (bottom) RFB02, Miletus.

4.3.2 Written Artefacts Marker

As both IIS are true-to-scale virtual representations of their real-world counterparts, we could mark the exact 3D location of the WA inside the 3D models. The marker visualizations for both IIS were chosen to fit the users' mental model for marking locations in different types of environments.

RFB05 – Church in Lucklum

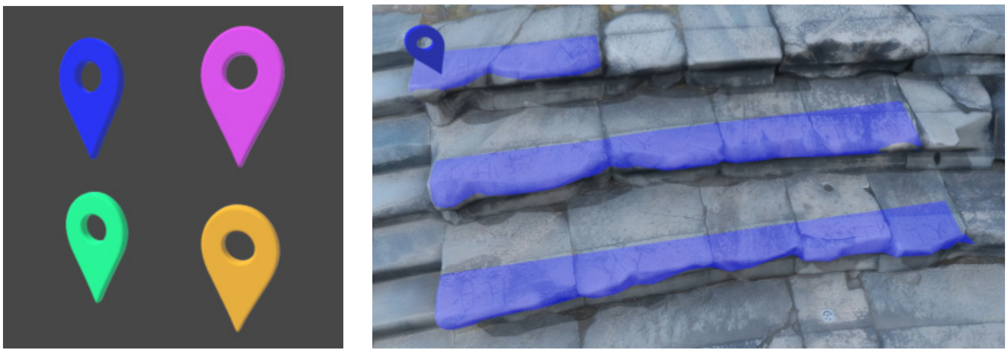
In RFB05, the emblematic paintings are located on vertical surfaces such as walls, the gallery, and the pulpit, and horizontally on the ceiling. Therefore, we choose flat spatial markers similar to commonly used labels in museums and exhibitions. We used a saturated and warm yellow color (**#EBAF0F**) for the markers to fit the color scheme of the emblematic paintings. Also, this color provides sufficient contrast against the black background and matches the yellow font color, details, and borders of the emblematic paintings. The markers were placed below or next to each WA and display a number corresponding to the existing publication ([SSA20]), as shown in Figure 4.5.

Figure 4.5: WA markers inside the church in Lucklum.

RFB02 – Theater of Miletus

The WA in RFB02 are engraved into the stones of the theater and are scattered over the large open space of its ruins. They are densely clustered in some areas and sparsely distributed in others. The WA can be found on top of the seating stairs, inside the curved profile of the stairs, or on various walls and vertical structures throughout the theater. We used spatial 3D pins (see Figure 4.6(a)), based on pins in applications with maps and location-based services, as this is a familiar metaphor for users. A billboard script is attached to each pin so that they always face towards the user's camera view. While some of the WA are large and can span across multiple seating stairs, others are small and relatively shallow or are hidden inside the curved profiles of the seating stairs, which makes them hard to spot. Thus, we implemented additional area highlights for a selection of WA. As shown in Figure 4.6(b), the area highlight applies an additive color overlay on top of the texture of the 3D model to mark the entire area that the WA occupy. Moreover, we defined four different colors for the 3D marker pins and area highlights that correspond to the four WA categories. This color-coding reveals the distribution of different types of WA throughout the theater, as shown in Figure 4.7. It visualizes the overlapping of different activities in ancient times, such as repairs, honoring individuals of the civic elite, playing games, and seat reservations. Various *Pictorial Elements* also serve as indicators of the presence of people, and their motifs convey information about the topic of conversation, important matters, or the surroundings that were visible to the theater's visitors at that time. The following four colors were specifically chosen to be accessible and barrier-free. They provide sufficient contrast to each other and to the grayish theater texture:

- Inscriptions: blue (**#0115F5**)
- Pictorial Elements: magenta (**#DC47EB**)
- Game-Boards: green (**#00FA91**)
- Architectural Marks: yellow (**#EBAF0F**)



(a) (b)

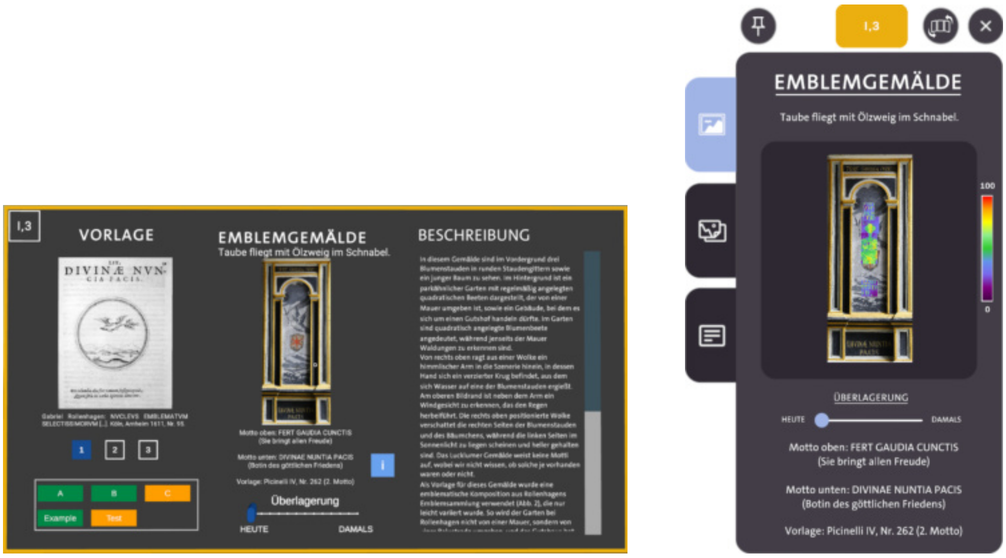
Figure 4.6: WA markers for RFB02: (a) color-coded 3D pins and (b) inscription area highlight.

Figure 4.7: Color-coded spatial 3D pins as WA markers inside the theater.

4.3.3 Spatial Information Panels

The WA markers can be selected to toggle (show/hide) the visibility of the spatial information panels. We integrated these panels to display various types information, including research data, transcriptions of the WA, translations, photographs and illustrations, as well as additional commentary for each WA. For the implementation, we used Unity's UI canvas component in *World Space* render mode with a MRTK canvas setup. In order to create a familiar UI for the users, we relied on common 2D desktop UI elements and established 2D UI patterns such as pagination and continuous scrolling [Jer15a; LaV+17; TBV20]. To ensure that information panels are always facing the user, they are controlled by a billboard script that tilts and rotates the panels in real-time according to the user's camera view direction in VR. Since the type and amount of information and research data differs for our two IIS, we designed two different information panel UI layouts with specific content for each space.

RFB05 – Church in Lucklum



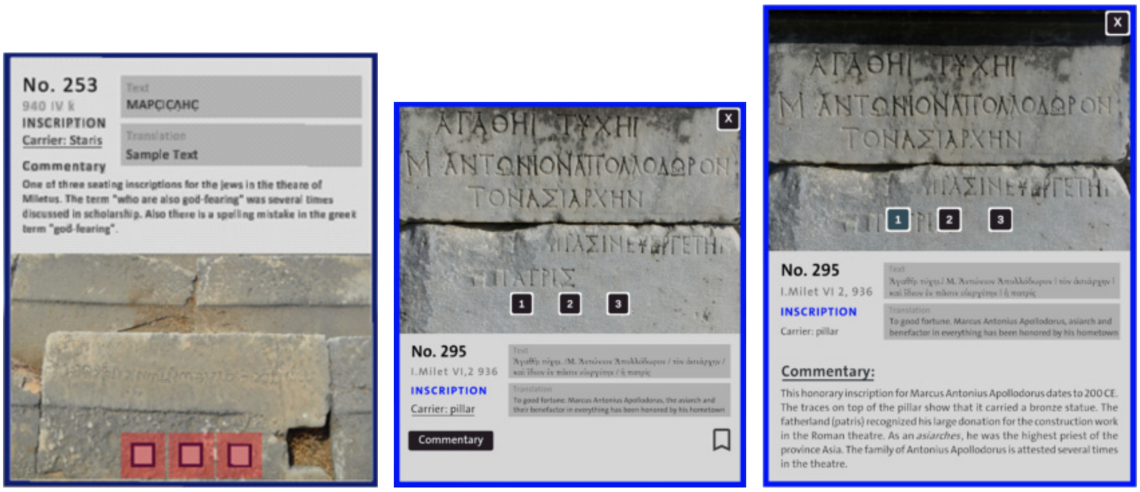
(a) (b)

Figure 4.8: Information panel design iterations for the church in Lucklum: (a) early design iteration and (b) final design.

For the 156 WA in the church, the majority of the information about the emblematic paintings and their corresponding Latin inscriptions was extracted from the existing publication by Steiger et al. [SSA20] and integrated into the spatial information panels. Since the paintings and the inscriptions are based on different reference motifs, the respective authors had to be referenced as well. Moreover, the texts for the WA are relatively lengthy due to the detailed description of the depicted motifs and the composition in the emblematic paintings. In the early development stage, we created information panels with a one page landscape layout, as illustrated in Figure 4.8(a). However, due to the amount of content and the limited space inside the church these panels where too wide, bulky, and obstructive. For the final design, we used a compact information panel in portrait format with tabbed views, as pictured in Figure 4.8(b). The displayed information includes a reference number, the title of the emblematic painting, transcriptions and translations for the Latin inscriptions, the names of the authors, an image gallery containing different templates for the motifs in the painting, and a long description text.

In the next chapter of this dissertation, we will examine and address the challenges for displaying long texts on spatial UI panels, investigate suitable UI patterns, and go into more detail about the used UI text panel layouts.

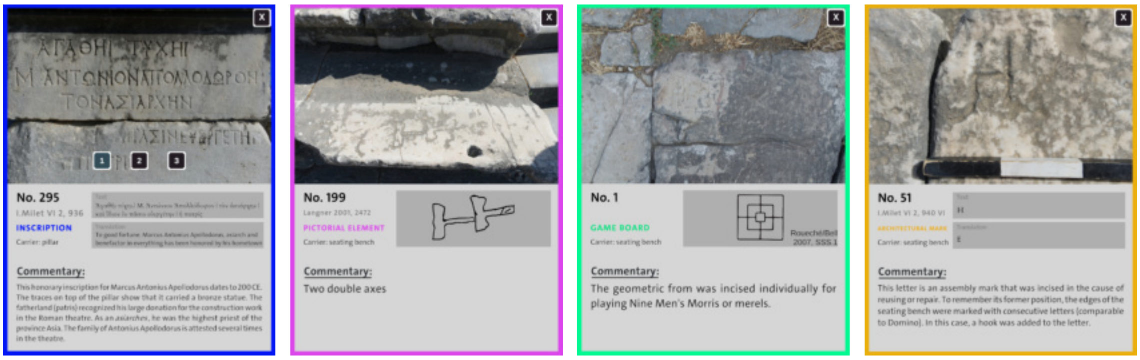
RFB02 – Theater of Miletus



(a) (b) (c)

Figure 4.9: Information panel design iterations for the theater of Miletus: (a) early design, (b) intermediate design iteration, and (c) final design.

For the theater of Miletus, it was necessary to display almost twice as many information panels due to the 300 WA that we had to integrate. Compared to RFB05, we had less textual information on the WA. Hence, we chose a simple, one page portrait layout for the spatial UI panels. The WA in the theater can be carved single letters, texts, geometric shapes, or drawings of different sizes. Therefore, we assigned half of the panel area to an image gallery. Depending on the size and placement of the WA, close-up and overview shots, as well as photographs from different perspectives, can greatly support the understanding of their visual characteristics. The recorded information on the panels include two reference numbers, WA category, WA carrier (object and surface of the engraving), transcribed text or an illustration (if it is a game-board or pictorial element), translation, and a commentary giving more background and context for the WA. The two reference numbers correspond to (i) the system used in the dissertation of Osthof [Ost25] and (ii) the numbering used in existing publications [Her98; Lan01]. Figure 4.9 depicts different design iterations of the information panel and Figure 4.10 displays the differences in color and content of the UI panels for each of the four WA categories.



(a) (b) (c) (d)

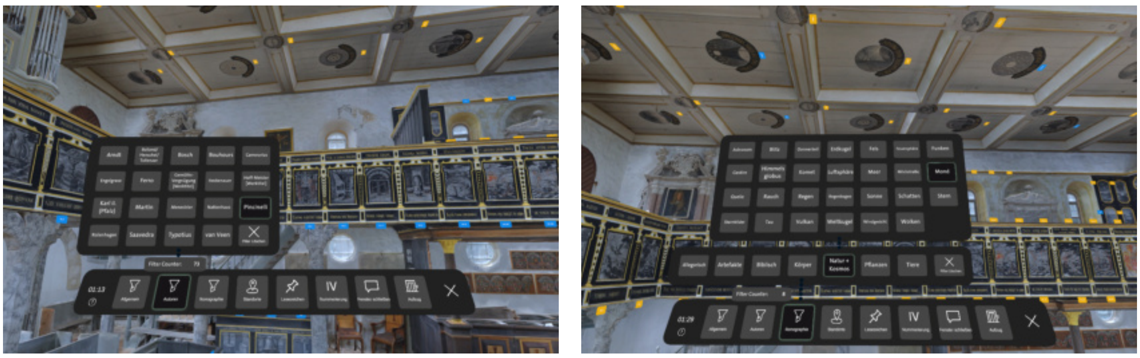
Figure 4.10: Information panel for each WA category in theater of Miletus: (a) Inscription, (b) Pictorial Element, (c) Game-Board, and (d) Architectural Mark.

4.3.4 Written Artefact Filtering

For visualizing additional clustering and detailed distributions of the WA based on specific attributes, we integrated filtered views for the WA markers into both IIS.

RFB05 – Church in Lucklum

We added extensive filtering options for RFB02. The menu pictured in Figure 4.11(a) filters the WA based on authors associated with the Latin inscriptions and the image motifs on the emblematic paintings. Figure 4.11(b) shows the large number of filter categories, with a second layer of sub-categories for further refinement based on religious iconography and different themes such as allegories, plants, and animals. In the filtered view, all matching WA markers are highlighted in a contrasting, vibrant sky blue color (#0F8AEB) to visualize the distribution of emblematic paintings with specific attributes within the church, as shown in Figure 4.11. Additionally, we added a filter counter to the filter sub-menus because not all of the paintings and their markers are visible at the same time, as they are obstructed by structures inside the building.



(a) (b)

Figure 4.11: Sub-menus for author and thematic filtering of the WA: (a) WA filtering based on authors and (b) filtering based on themes.

RFB02 – Theater of Miletus

While the color-coded 3D pins already provide an overview of the distribution and clustering of different categories of WA, we further refined the visualization by adding filters. The filtered views indicate different groups of people (for example, Jews or prominent individuals). In the future, further filtering based on time periods and functions of the WA could be added. The filter sub-menu is illustrated in Figure 4.12. The 3D pins matching the filter criteria remain in the filtered view, while the rest of the pins are hidden.

Figure 4.12: Sub-menu structure for filtering the visibility of the WA markers.

4.3.5 Reconstructions

Another important aspect was to restore visibility of former, no longer existing structures and hidden, overpainted WA. We included digital reconstructions within the IIS to visualize different temporal layers.

RFB05 – Church in Lucklum

In RFB05, the material scientist (Sebastian Bosch) in the project group conducted non-invasive X-ray Fluorescence (XRF) scanning of an emblematic painting (number I,3) at the church's pulpit. XRF-scanning outputs color mapped, square image patches for each scanned area, which visualizes the intensity of the measured X-ray fluorescence. X-ray fluorescence analysis is a non-invasive method for determining the elemental composition of inorganic materials, in this case inside the used paint of the emblematic painting. The XRF-analysis uncovered an earlier, hidden layer that was painted over in the early eighteenth century. The full XRF imaging dataset can be found in the publication of Bosch [Bos21]. The XRF-scanning process and an example output of an intensity image patch can be seen in Figure 4.13.

We visualized the uncovered hidden layer for emblematic painting number I,3 inside the IIS. After stitching (seamlessly combining) the XRF-scan patches into one image, we implemented interactive sprite overlays in Unity that are placed directly on top of the painting and, in addition, on top of the picture inside the spatial information panel. As shown in Figure 4.14, the opacity of the overlay can be freely adjusted by the user through the UI in the application.

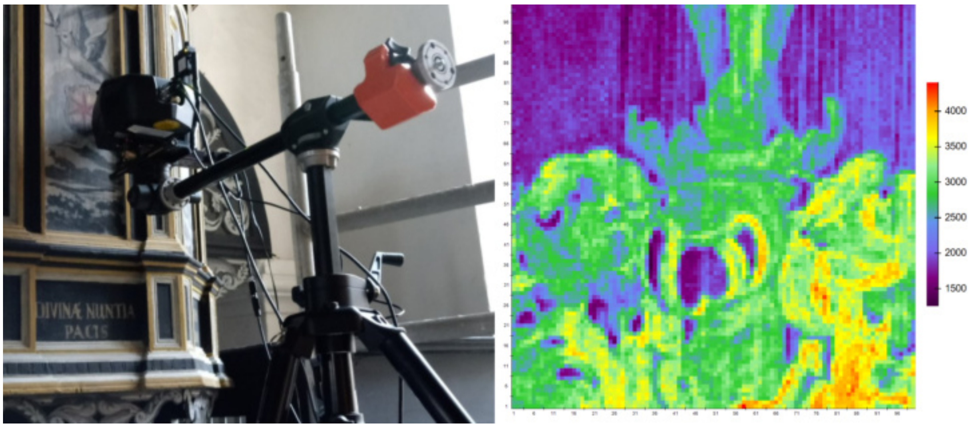


Figure 4.13: (Left) XRF-scanning of an area of the emblematic painting and (right) intensity image of one scanned area (photography ©S. Bosch).

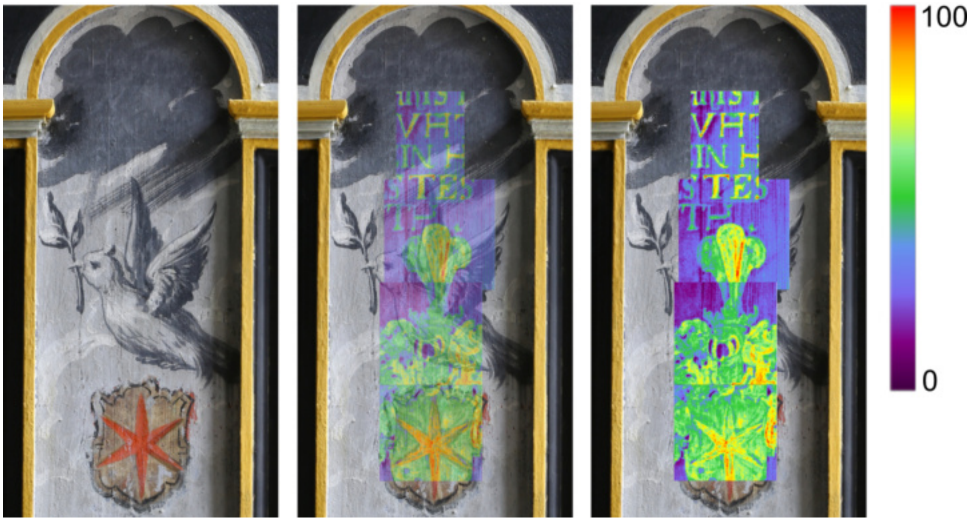
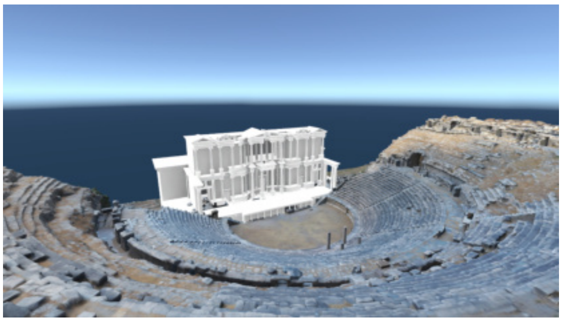


Figure 4.14: XRF intensity image of the hidden layer as interactive overlay on the emblematic painting. From left to right: 0% opacity 50% opacity, full opacity (XRF images ©S. Bosch, photography ©J. A. Steiger [SSA20]).

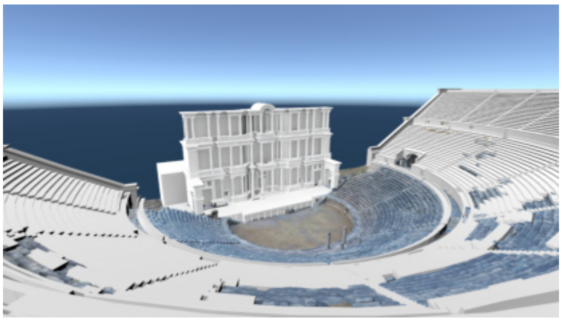
RFB02 – Theater of Miletus

The ancient theater of Miletus consisted of three seating tiers, of which only the first tier and parts of the second tier remain today. A key question for the humanities researchers in RFB02 is how people used to structure and interact with the space, and how they occupied and represented themselves in it. For example, WA could be permanent representations such as honorary inscriptions or limited, temporary ones (topos-inscriptions) that were relevant during plays and political gatherings. Moreover, Game-Boards reveal how the theater was used as an informal space for playing games. To digitally restore former structures, we created 3D reconstructions, which include the fist Roman stage building (1st century CE) and the second Roman stage building (2nd century CE), major parts of the second seating tier, the entire third seating tier, and the last seating rows of the first and second tier. We deliberately avoided using colors or textures for the reconstructions so as not to create suggestive representations. These reconstructions are pictured in Figure 4.15 and enable researchers to visualize multiple time periods of the ancient theater and experience the theater at its former full size during the 2nd century CE. Users can activate and deactivate the reconstructed structures from both time periods through the main menu.

An important advantage of our digital reconstructions is that they provide a non-invasive alternative to on-site conservation using a technique known as anastylosis, which is the re-erection of non-preserved architectural elements in the real world. Anastylosis has certain disadvantages, as only one specific time period can be restored and moreover, the process often results in damage to the ancient structures.

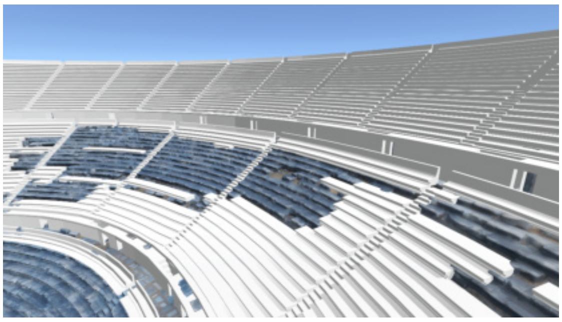


(a)



(b)

(c)



(d)

Figure 4.15: Reconstructed, former structures of the theater from two different time periods: (a) 1st Roman stage reconstruction, (b) 2nd Roman stage and seating tiers reconstruction, (c) 2nd Roman stage and analemma (supporting) walls, and (d) 2nd seating tier, partial reconstruction.

4.3.6 Overview Spots

We implemented different overview spots that enable researchers to engage with new perspectives of the inscribed space and the WA. The overview spots are placed at various points of interest (POI) and in areas or places that are inaccessible at the historical site. Each spot has a predefined starting camera orientation to ensure the best view from each specific position in the IIS. Users can select different overview spots from the main menu and are teleported to the position with a screen-fade transition. In RFB05, the lodge of the church's Komptur (Kompturlodge) is a very small and narrow room, as pictured in Figure 4.16, that can be accessed through an overview spot. Moreover, the

(a) (b)

Figure 4.16: Overview spots providing access to different, new perspectives: (a) Kompturlodge view of the church in Lucklum and (b) bird's-eye view of the theater of Miletus.

gangway for entering the lodge contains five additional WA that cannot be accessed on-site by regular visitors. In our IIS, users can enter and explore this special POI. For RFB02, we added a top down bird's-eye overview that allows researchers to view the entire theater of Miletus and its layout (see Figure 4.16). Stones from the former third seating tier were removed and used as building material for a late Byzantine castle building, which now sits directly on top of the theater. Today, it is not permitted to climb up to the top of the castle ruins. The only way to experience the view from the castle is through the overview spot inside our application.

4.3.7 Time of Day

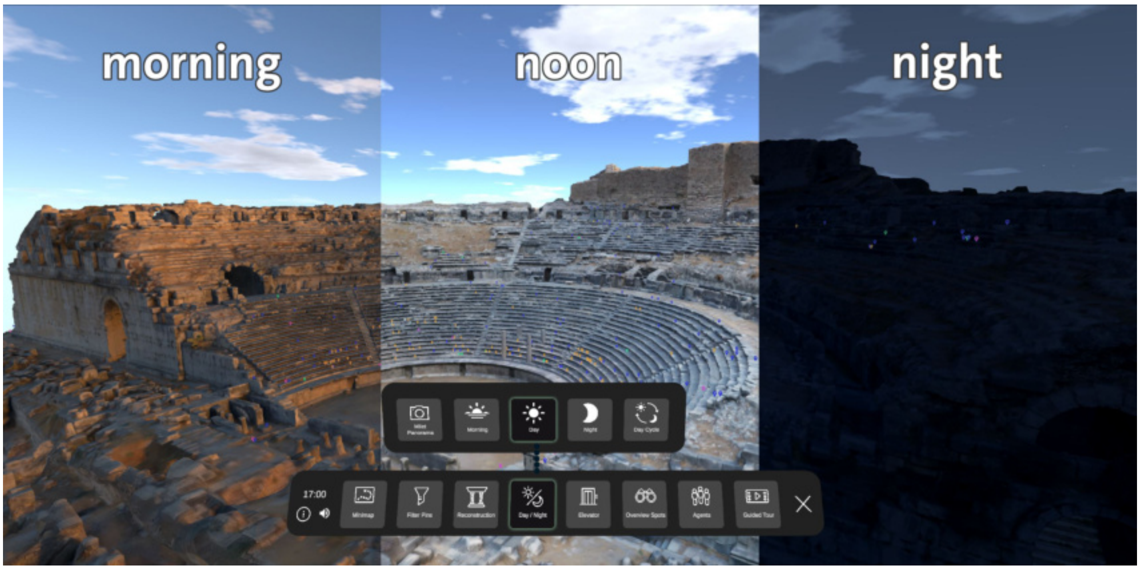


Figure 4.17: Split image showing morning, noon, and night time in the theater of Miletus.

For RFB02, we implemented an exclusive feature for simulating the time of day inside the theater. The default view displays a static 360° skybox based on stitched drone images of today's surroundings of the theater. Users can select an alternative dynamic skybox and set the visualization to three predefined times: morning (6 a.m.), noon (12 p.m.), and night (10 p.m.), as visualized in Figure 4.17. Additionally, they can select a day-night cycle that simulates 24h of a realistic summer day (mid of July) as a one minute time-lapse. Combined with the reconstructions of the second Roman stage, the sunlight and shadows of the day-night cycle revealed that the visibility of an inscription was limited and it was mostly shaded throughout the day. This allowed researchers to raise new research questions and draw new conclusions about the perception of WA.

4.3.8 Seating Capacity – Agents

Another exclusive feature for RFB02 is the display of stylized human agents for visualizing the seating capacity inside the ancient theater of Miletus. As Machidon et al. [MDC18] stated, it is also important to consider the population when creating virtual cultural heritage sites. Therefore, we included the spectators to bring the ancient theater back to life and to improve the sense of scale experienced by the users. According to calculations of archaeologists and building researchers in existing literature, it is estimated that the ancient theater had a seating capacity of around 15,000 to 18,000 people [Isl18; Kra73; Sea06].

We tested this assumption by filling the ancient theater including the reconstructed second and third seating tiers with seated agents representing the full seating capacity for spectators. In contrast to the calculations in existing literature, we found that the seating stairs of the theater were already well-occupied with 8,366 spectators. Visualizations of different seating arrangements can be selected through the main menu. These include the aforementioned full seating capacity view, a first tier only view, and special designated seating areas within the theater for particular groups of people, such as Jews, as pictured in Figure 4.18. These designated locations were derived by the analysis of the placements of the WA.

We designed the agents as neutral mannequins with uniform coloring, since we wanted to avoid suggestive representations of the appearance and clothing of the ancient visitors of the theater. We did not use any texturing on the agent models and selected a light orange color with high contrast to the color of the 3D model of the theater. A small amount of agents were also animated with basic looped motions and gestures, such as waving, cheering, and clapping.

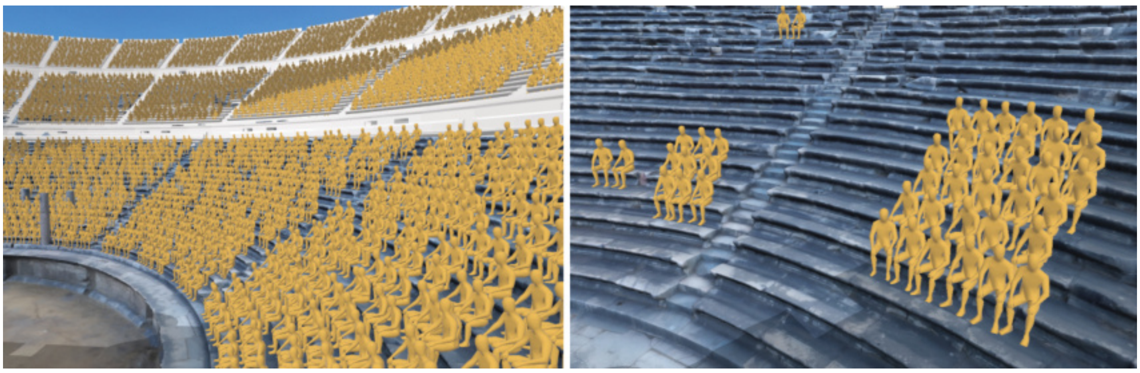


Figure 4.18: (Left) Agents at full seating capacity in the completely reconstructed theater (2nd century CE) and (right) particular groups in their seating areas.

4.4 Implementations for Additional Platforms

In addition to the main VR implementation of the two IIS, we tested and experimented with other platforms and hardware to explore their potential for future use. Furthermore, these prototypes were used for various presentations, public demonstrations, and for internal and external communication of our research projects.

4.4.1 CAVE Projection

We created a prototype for displaying the theater of Miletus, including the reconstructions and information panels inside a CAVE [CSD93]. We used a custom built 4-sided CAVE setup in the HCI laboratory at our university. It consists of four full-HD projectors with three walls and the floor as projection surfaces and has a total size of 3.15 × 4.2 × 2.36 meters (D × W × H). A detailed description of the technical setup, including an illustration of the CAVE, can be found in the work of Schmidt and Steinicke [SS17]. The system uses retroflective markers and an optical camera system by OptiTrack[1] to track position and rotation of a user's head. The markers are attached to active shutter glasses and defined as a so-called rigid body within the tracking system. With the tracked shutter glasses, the CAVE projection can be viewed in two modes – a stereoscopic view mode and a monoscopic view mode, as pictured in Figure 4.19.

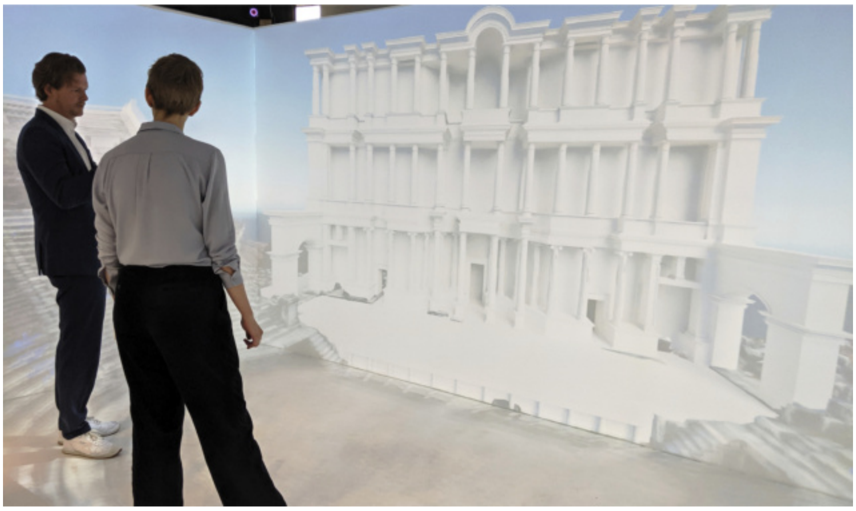


Figure 4.19: CAVE in monoscopic view mode without positional tracking of the user, allowing multiple users to join the experience (projection shows the reconstructed stage of the theater).

The CAVE prototype allows for limited, interactive exploration of the virtual theater of Miletus via movement with the tracked shutter glasses and simple controller and keyboard input. Based on specific button inputs, the reconstructions and a small selection of inscription info panels can be switched on and off.

[1] https://www.optitrack.com

4.4.2 Augmented and Mixed Reality

As a major part of the work for researching WA takes place on-site, we tested different AR and MR approaches to display spatial content in the real world.

AR Core XRF-Scan Overlay

We developed an android smartphone AR prototype for RFB05, the Interior of the Church in Lucklum. We used AR Core and Unity's AR Foundation framework with image tracking to implement a visualization for the hidden image layer that was uncovered by the XRF-scans of the emblematic painting number I,3. A photograph of the unique emblematic painting was imported into a reference image library in Unity and used as image marker for the AR tracking. The hidden, overpainted image layer of the emblematic painting is shown as an overlay on the smartphone screen as soon as the image marker is recognized. The opacity of the overlay can be manually controlled with a UI slider. Furthermore, the painting's name and reference number are displayed on the screen. We tested the application with an image of the emblematic painting on a vertical digital display and on-site at the church's pulpit, as depicted in Figure 4.20.

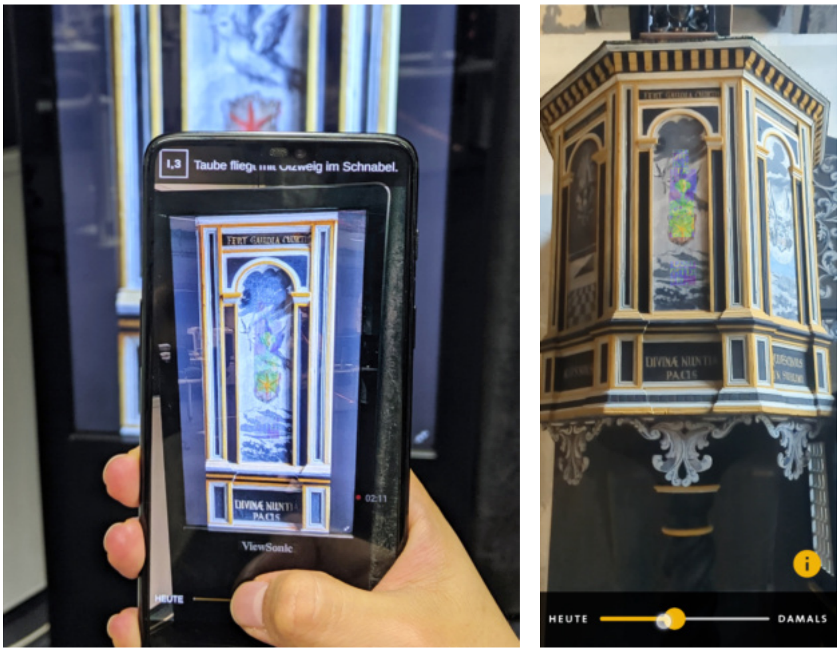


Figure 4.20: Prototype of Lucklum smartphone AR hidden layers overlay: (left) testing with an image of the emblematic painting and (right) camera view of on-site testing at the church's pulpit.

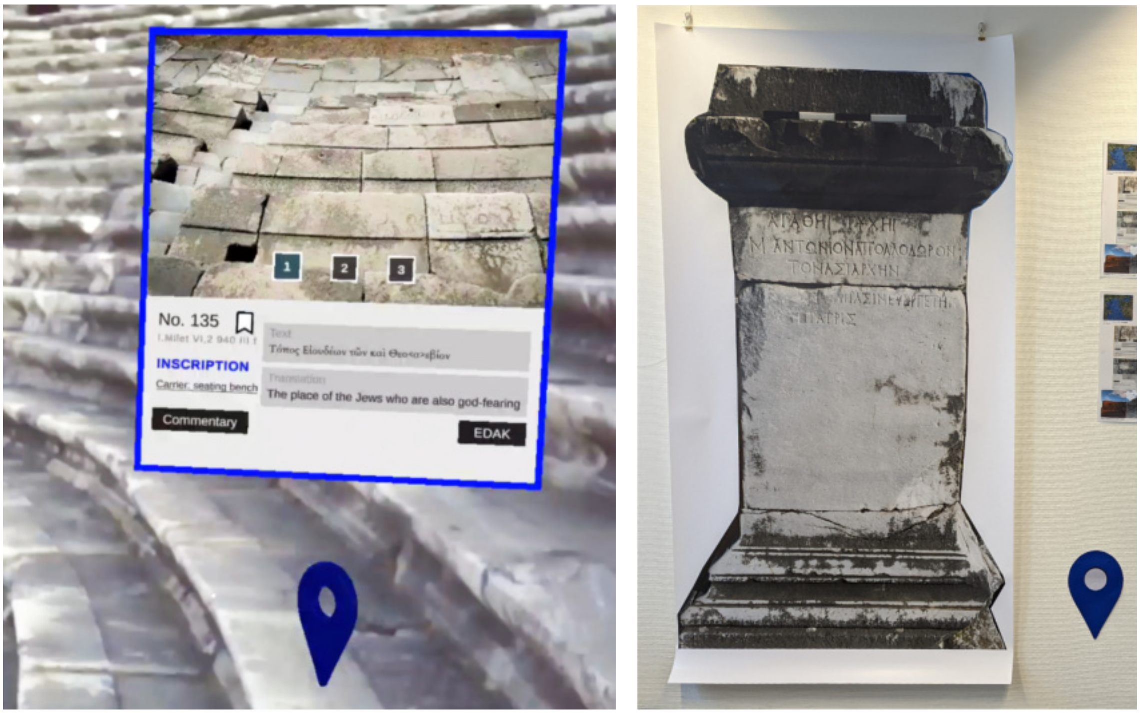
MR Inscription Pin and Infopanel

We also implemented two small MR prototypes for RFB02, the ancient theater of Miletus, for displaying spatial MR content. The first prototype application was developed for the Nreal Light glasses (the company has since been renamed to Xreal)[2]. The Nreal Light was a light-weight, consumer focused standalone HMD with spatial tracking and a 53°diagonal field of view is similar

[2] https://www.xreal.com/

to the HoloLens 2. With the prototype application, users can place 3D spatial WA markers on-site at the theater of Miletus. Users can interact with these markers by using the Neal Light controller to toggle the visibility of information panels. An in application capture (life footage) recorded on-site at the theater is shown in Figure 4.21(a).

The second prototype was designed for interacting with a WA marker in 3D space on the HoloLens 2. For this application, we placed the marker next to a large, close to true-to-scale printout of an inscription on a statue base of the theater of Miletus (inscription number 295), as pictured in Figure 4.21(b). Users could interact with real-time hand tracking and pinch gestures to select the 3D marker and toggle visibility of the information panel.



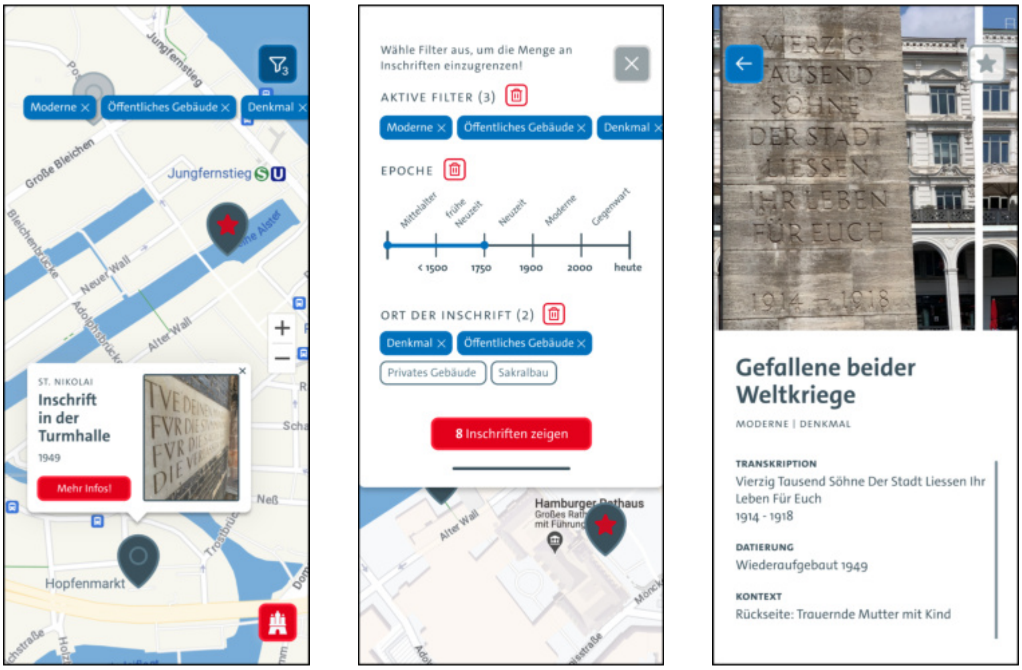
(a) (b)

Figure 4.21: Testing MR prototypes for visualizing spatial information panels in real-life settings: (a) Nreal Light application with 3D marker and info panel captured at the theater of Miletus and (b) HoloLens 2 application with 3D marker and almost true-to-scale WA printout.

Mobile Application Concept “Inschriften in Hamburg”

As part of a HCI bachelor’s project at our university, we supervised and worked with students who explored the use of AR image tracking for identifying a selection of WA in the city of Hamburg (Germany). The students created a simple Android smartphone prototype with Unity’s AR Foundation image tracking that matched reference images in its library (uploaded photographs) to WA captured in the smartphone’s camera view. If a match was found, additional information on the specific Written Artefact was displayed. All available predefined WA were displayed within a map view of the Hamburg city center.

This concept for creating a map-based application for locating and learning about WA was further extended and refined by a student in her bachelor's thesis [Kei24]. She designed and implemented progressive web application "Inschriften in Hamburg" (Inscriptions in Hamburg) for mobile platforms. The application is in German and is an initial prototype that integrates 19 inscriptions from the city of Hamburg. The inscriptions were selected in collaboration with the humanities researchers in our research project. The application uses a Google Maps integration with custom inscription location markers and provides a UI for filtering inscriptions based on their theme, time period, and placement. There is also a detail view with photographs and information for each inscription. These features are illustrated in the UI screens in Figure 4.22. The web application can serve as a basis for future addition of other WA and for further development that integrates the initial AR image tracking concept.



(a) (b) (c)

Figure 4.22: UI screens showing different features of the web application: (a) Inscription map view, (b) inscription filter menu, and (c) detailed information view (images ©N. Keilich).

4.4.3 Multi-Touch Display

We tested the visualization of digital WA data on a multi-touch display. The initial prototype was created on a custom-made multi-touch table from our HCI laboratory. The setup is built with a 55-inch ultra-high-definition TV display equipped with the infrared multi-touch frame G4S by PQ-Labs[3]. The frame uses the TUIO-protocol[4], an open source framework for multi-touch applications. The touch display is attached to an adjustable mounting stand for changing height and

[3] https://www.pqlabs.com/
[4] https://www.tuio.org/

tilt angle, which allows the display to be used in both vertical and horizontal mode. Everything is mounted on a movable platform that also carries the PC. Our implementation supports TUIO as well as Windows native touch input. The developed touch prototypes for both projects offer multi-touch gesture support for tapping, scaling, rotation, and movement of WA information panels and emblematic paintings, as seen in Figure 4.23. For RFB05 (Lucklum) we also ported the hidden XRF-scan overlay visualization for the emblematic painting to the touch display.

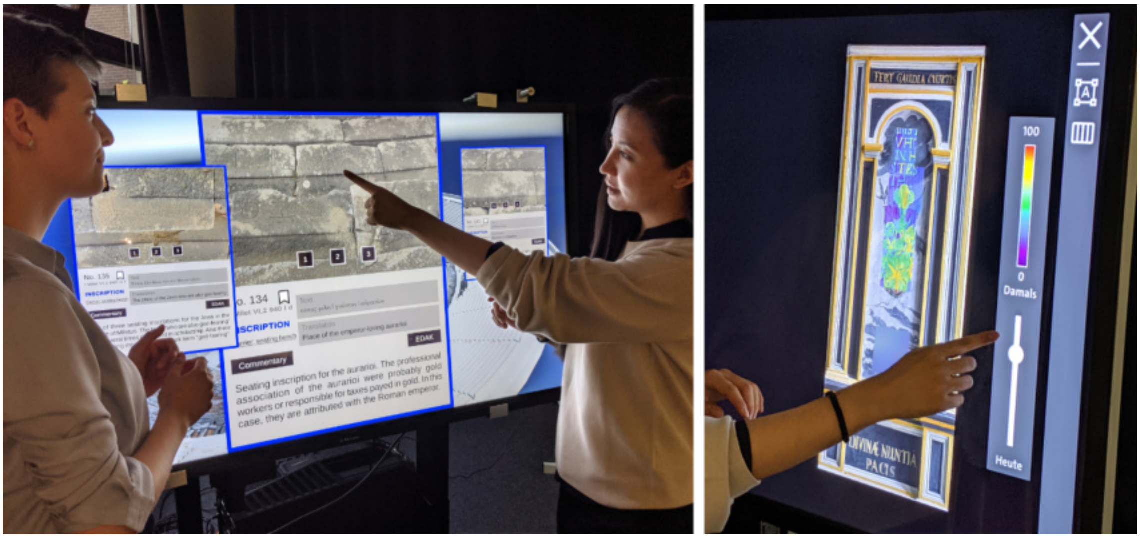


Figure 4.23: (Left) Two users discussing the design and content of the WA information panels (Miletus) on the multi-touch table and (right) adjusting the opacity of the hidden layer of the emblematic painting overlay (Lucklum).

Visualizing the WA information on a large multi-touch display could provide additional benefits. It could facilitate collaborative, multi-user scenarios in which users can freely discuss, sort, and arrange information on the WA. Moreover, touch interfaces and basic touch gestures are familiar to a wider range of users compared to using spatial interaction on HMDs. The use of touch applications could also make the research results more accessible for the general public.

Miletus Touch Display Application

The UWA cluster aims to provide the research results of project RFB02 to the general public and the visitors of the theater of Miletus. As part of these efforts, it was decided to present the research results and the developed interactive content on-site at the local museum in Miletus. Due to the limited space and technical support on-site, a touch display application would be the most suitable for use in this public space, since it would not require supervision or instruction for the visitors.

We developed a touch-application by transferring and adapting the core features of the VR application for a Windows native touch display, as shown in Figure 4.24. Additionally, we added new content including an inscriptions tour, extended commentary for the WA, new overview spots, a 3D model of an ancient Roman statue for the statue base in front of the stage building, and complete Turkish translations for the application. The inscriptions tour guides users through a selection of inscriptions in a predefined sequence and offers additional descriptions. The new content was

Figure 4.24: Touch display of the adapted touch application for the theater of Miletus: (left) touch display mounted on the frame and (right) start view of the Miletus touch application.

added based on requirements, research data, and feedback from the humanities researchers. To ensure a smooth, continuous, and long-term operation at the museum with minimal setup, the touch application was added to the computer's autostart and automatically opens up in full-screen on startup. If no touch input is detected within 2 minutes, the application resets and switches to a screensaver mode with a bird's-eye view of the theater model. As soon as a touch input is registered, the screensaver mode is ended and the interactive touch application starts with the initial theater overview, as pictured in Figure 4.24. The application was developed and tested on a commercial 49-inch touch display with 4k resolution and Windows native touch support. The display is connected to a mini-PC that is mounted to the frame of the touch display.

4.5 Remote User Study

We conducted a user study to evaluate the developed VR applications for RFB05 and RFB02. For this initial study, we focused on investigating qualitative aspects, including user feedback for interacting with the IIS. Our aim was to gain insights into requirements, UX, and challenges for the spatial interactions and the implemented features. Due to the COVID-19 pandemic, the user study was conducted remotely.

4.5.1 Study design

The application was tested with seven participants (women=2, men=5), aged between 28 to 54 ($M = 39.57$, $SD = 11.47$). Five out of the seven participants were members of the UWA research cluster, with expertise ranging from classical archaeology, ancient history, epigraphy, literary study, theology, and HCI. The remaining two participants where PhD students, one from the field of HCI and one from ancient history. Both were not part of the UWA cluster. Five of the participants were inexperienced VR users with no prior VR experience outside of testing the WA applications within the research projects. The other two participants were VR expert users from the field of HCI, who had no expertise in the research of WA. Before the study, all participants signed an informed

Figure 4.25: Humanities researchers using the Meta Quest 2 during our remote study.

consent form. During the remote user study, Zoom was used for communicating, observing, and guiding the participants. As pictured in Figure 4.25, each participant was provided with a Meta (formerly Oculus) Quest 2 HMD with the pre-installed applications. If necessary, the participants were instructed on the use of the HMD including the Meta (formerly Oculus) Touch controllers. They were guided through the basic interactions and navigation functions in the applications. As the user study was focused on qualitative evaluation and user feedback, free exploration and the thinking-aloud method were used.

Participants were tested in a room-scale setup, but they were not required to walk around. As depicted in Figure 4.25, they were asked to position themselves in front of their PC so that they are clearly visible and audible throughout the study via Zoom. Each participant was given 15 minutes to freely explore and interact with each IIS in VR. Participants tested both applications to compare them, gain insights into the requirements, and reveal challenges for both IIS. Since the two spaces have different properties and types of WA, we were interested in potential differences regarding the requested features. After approximately 30 minutes, participants were asked to take off their HMD and fill out two questionnaires on their PC, the igroup Presence Questionnaire (IPQ) [RS02; Sch03] and the System Usability Scale (SUS) [Bro+96]. Afterwards, they provided their feedback on using the IIS. In total, the study took about 60 minutes to complete.

4.5.2 Results

We analyzed the data collected from both standardized questionnaires. We also looked at the qualitative user feedback. No statistical data analysis was done, since we only collected data from a limited sample size of seven participants for this initial study.

System Usability Scale (SUS)

For the SUS answers, the mean (*M*) and standard deviation (*SD*) were calculated for the overall score and for the Usability and Learnability subscales based on Lewis and Sauro [LS09]. According

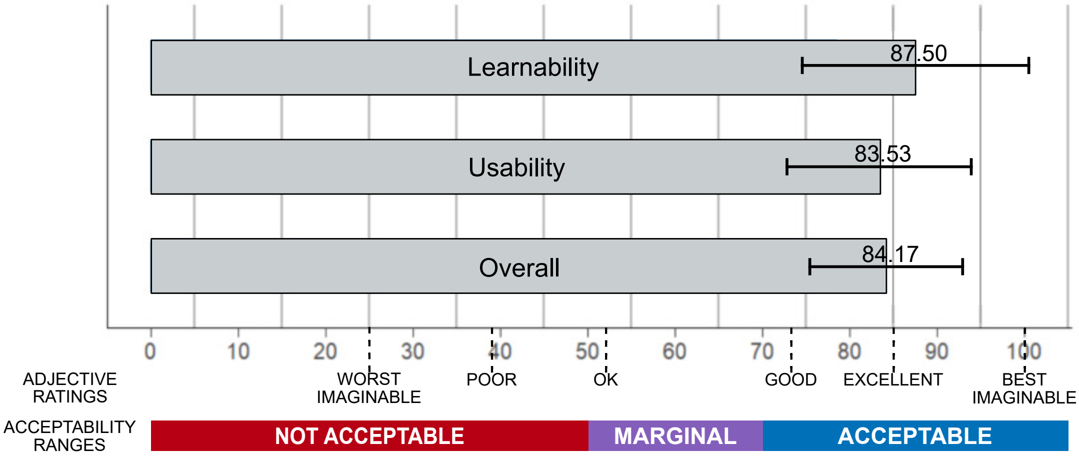


Figure 4.26: SUS mean and SD for the overall and subscale scores.

to the rating and visualization proposed by Bangor et al. [BKM09], we plotted the SUS adjective ratings and acceptability ranges for the overall and subscale scores, as visualized in Figure 4.26. The scores on all three scales are all within the acceptable SUS range. Overall, the applications were rated "excellent" on the adjective rating scale ($M = 85.36$, $SD = 8.59$). The Usability subscale ($M = 84.82$, $SD = 10.43$) is slightly lower than the overall rating, while the Learnability ($M = 87.50$, $SD = 12.50$) has an "excellent" adjective rating.

igroup Presence Questionnaire (IPQ)

The IPQ contains a General Item for measuring the general "sense of being there" and three subscales: Spatial Presence, Involvement, and Experienced Realism. It uses a seven-point scale, with scores ranging from 0 to 6, where scores below 3 indicate disagreement and scores above indicate agreement with the 14 IPQ items. The scores for the General Item and each subscale are shown in Table 4.1. The General Item ($M = 5.14$, $SD = 0.90$) and Spatial Presence ($M = 4.86$, $SD = 0.62$) were rated positively and well above the median value of 3 by the participants, indicating that the sense of being physically present in the virtual environment is perceived to be quite high. The Involvement ($M = 3.46$, $SD = 1.21$) was rated slightly above the median value, while the Experienced Realism ($M = 2.50$, $SD = 0.80$) was rated slightly below the median.

Table 4.1: IPQ mean and SD for the General Item and the three subscales.

IPQ	Mean	SD
General Item	5.14	0.90
Spatial Presence	4.86	0.62
Involvement	3.46	1.21
Experienced Realism	2.50	0.80

Qualitative User Feedback

In addition to the questionnaire results, we obtained qualitative insights about the IIS from the thinking-aloud, our observations, and from user feedback at the end of the study. All participants reported that they had a positive experience in VR. They stated that they had the feeling of being right there at the locations and that standing in the IIS, especially the large-scale theater of Miletus, felt very impressive. Everyone, including the inexperienced VR users, was able to securely navigate and interact within the VEs during the study. The inexperienced participants needed explanations for the teleport function, but were able to use it successfully after trying it out once. Two participants stated that they needed a few trials to get used to being teleported. Another important observation was the intuitive pointing and raycasting-based selection of the 3D spatial WA markers and interaction with the information panels by all participants, although we did not instruct them on the implemented selection pattern beforehand. All inexperienced participants reported that the interactions in VR were easy to understand and execute. However, reported and observed issues regarding the interactions were the limited selection range and difficulties with precise selection and interactions of distant 3D WA markers and the information panels controls. Participants had to teleport quite frequently and reposition themselves to get closer to perform interactions in the IIS. Another concern was the overlapping of spatial information panels in certain areas. We also gained the following insights into benefits and advantages of using IIS for researching WA:

RFB05: For investigating the WA inside the church in Lucklum, spatial understanding was of great importance to the researchers. In addition, the visualization of the hidden layers based on the XRF-scans added temporal context to the emblematic paintings. The researchers stated that these uncovered hidden layers in the painting have already led to earlier church interior being dated correctly in comparison to previous research. They noted that the IIS provides them with the unique opportunity to easily visualize, compare, and directly contrast the information and the overpainted layer on the emblematic painting inside the virtual interior of the church. They stated that this is particularly helpful when working remotely, especially in times of restricted travel or if access to historical sites is limited.

RFB02: The researchers working with the ancient theater of Miletus stated that being able to see and walk through the virtual theater with the reconstructions allowed them to get a good sense of the distances, scale, and height of the structures. Moreover, spatial understanding, the architectural atmosphere, as well as movement paths through the theater can be investigated by moving through the fully reconstructed virtual theater. The researchers explained that this allowed them to gain new perspectives and a better spatial understanding for the analysis of the WA. This immersive experience is not possible on-site at the real theater without the visualized reconstructions. During the study, this advantage of the IIS resulted in new research questions about the placement, visibility, and importance of a particular WA – specifically an inscription located on a remaining statue base next to the stage building. As stated in Section 4.3.7, the placement of this particular inscription raised questions after the researchers viewed it in combination with the reconstructed stage building and the time of day simulation, which revealed its limited visibility.

4.6 Field Testing and Public Presentations

4.6.1 Field Testing

We tested an preliminary version of our VR application for RFB02 on-site with researchers and archaeology students at the excavation house in Miletus. We also showed it to members of another excavation team working at the Apollon Temple of Didyma (Türkiye), which was an extra-urban sanctuary of Miletus. Different users wearing the Meta Quest 2 HMD are pictured in Figure 4.27. These field testings were intended for demonstrating the applications to relevant user groups outside of our research projects. It was important for us to inspect potential issues regarding the controls, the implemented features, and the displayed research data about the WA. The majority of the users on-site had no prior experience with using VR HMDs. We observed them during their free exploration and interaction with our IIS. The users were also encouraged to do a think-aloud and comment on their experience. Similar to our remote user study (cf. Section 4.5), there were no major problems with using or understanding the application. Users gave general feedback, including suggestions for specific overview spots and adjustments for the wording and placement of the main menu elements. Many also stated that they particularly liked the 3D reconstructions, time of day simulation, and the integrated WA data. The overall sentiment was that using the IIS of the theater of Miletus was not only interesting and informative from a research perspective, but it was also engaging and fun to interactively explore the VE and WA.

Figure 4.27: Users testing the theater of Miletus VR application at excavation houses in Türkiye.

4.6.2 Public Presentations

Although users from outside the research fields of our project members and the general public were not our target group for the development of our research focused applications, these IIS can nonetheless help to communicate the research data on WA in a visual, easy to understand, and compelling way. Therefore, we demonstrated our two VR applications at various public events and presentations (see Figure 4.28):

- Evening event *The Virtual theater of Miletus*[5]

[5] https://www.csmc.uni-hamburg.de/about/blog/2022-02-02-virtual-milet

- Public demonstration at the *Open Day of the UWA Excellence Cluster*[6]
- Demonstration at the *Museum für Kunst und Gewerbe Hamburg (MKG)*[7]
- Miletus VR demonstration at the *Rechtschutztagung "Vorsorge, Verbote und Sanktionen"*[8]
- Miletus VR demonstration at *Long Night of Science*, with our cooperation partners from the photogrammetry laboratory at Berliner Hochschule für Technik (BHT) in Berlin
- Miletus VR and CAVE demonstration for students of the *Deutsche JuniorAkademien 2023*[9]
- Miletus VR and touch-table demonstration at *Science City Day 2024* in Hamburg[10]

(a) (b) (c)

Figure 4.28: Users trying out our VR applications at various public events: (a) Evening event at the CSMC, (b) Miletus VR at the MKG, and (c) UWA cluster open day (photographs (a) ©A. Kalinichenko, (b) ©F. Hildebrandt, and (c) ©A. L. Osthof).

We did not formally collect user feedback, but these public tests and events provided us with general insights into UX and potential challenges for the future adaption to other use cases for a wider audience. They allowed us to obtain diverse, real-life feedback from users of different ages, backgrounds, and level of VR experience. We found that, in general, non-academic users would require more introduction to the historical context, the inscribed spaces, and the WA. Besides the free exploration of the IIS in VR, additional, targeted guided tours with commentary and explanations could help users learn about the historical background and the WA without overwhelming them with the large amount of integrated information and research data. Some users had difficulties with correctly using the VR controllers. Moreover, we realized that it was challenging to keep the devices charged and the controllers correctly assigned to their respective paired HMDs during all-day, frequent usage in a public setting. There were also some safety concerns regarding the use of the controllers, although they were secured to the wrists, especially with the younger user group. Overall, users were open to try out the applications and reported that they had fun exploring the interactive VEs. Especially younger users were excited about the immersive experience and

[6] https://www.csmc.uni-hamburg.de/openday-en
[7] https://www.csmc.uni-hamburg.de/news/2022-07-21-milet-mkg
[8] https://www.geschichte.uni-hamburg.de/arbeitsbereiche/alte-geschichte/tagungen/grabschutz
[9] https://www.csmc.uni-hamburg.de/news/2023-08-01-virtual-mileuts
[10] https://www.sciencecityday.de/programm/das-theater-von-milet-in-virtual-reality/

using the VR HMD. The positive feedback from these public events shows that IIS can also support activities for science-communication.

4.7 Discussion

The outlined implementation process for our IIS and the user insights validate the importance of following a UCD approach in this novel context. We benefited greatly from the user feedback and the exploratory and iterative development process, which helped us design and successively refine the interactive features and 3D visualizations for the VR applications. Following the initial remote user study, we continued to involve the humanities researchers during the development process with regular meetings.

Furthermore, our work, including the gained insights about 3D data acquisition, 3D model creation, and implementation, highlights the complex and demanding technical requirements for creating realistic, high-quality 3D representations of the IIS. After our experiments and tests with the 3D data, we realized that it is essential to have dedicated, professional support for 3D data collection and processing for projects of this scope. We could only achieve the desired results and quality for our 3D models through large-scale data collection and professional 3D data processing.

Overall, the results of our conducted remote user study were positive. As stated by the participants, the IIS allowed them to get a sense of the distances and scales in the 3D space. Together with the visualizations of former structures and layers, this enabled them to investigate WA with a better understanding of the spatial and temporal contexts. Using the IIS also facilitated the researchers' understanding for the potential advantages and benefits of using immersive technologies in their research. The applications received high ratings in Learnability in the SUS and for the sense of being physically present in the IPQ, which also matches the qualitative feedback and our observations. The IPQ Involvement score measured just slightly above the median rating. The reduced involvement experienced in VR could have been caused by the limited amount of placed inscriptions and implemented interactions in the development stage of the applications at the time of the study. *Experienced Realism* received the lowest IPQ score. This is not surprising, since the 3D models used during the study were still in a prototype stage and required additional 3D scanning data for further processing and improvement of the 3D meshes and textures (see Section 3.1.2). As we used UI elements based on familiar, established desktop UI patterns and easy to understand, raycasting-based interactions, we assumed that participants will not encounter major difficulties during the study. Nevertheless, our findings did reveal specific interaction challenges.

4.8 Conclusion

In this chapter, we introduced our UCD process, as well as the technical implementations and features for creating the IIS for our two cross-disciplinary research projects. We collaborated closely with researchers from various research fields in the humanities to develop these interactive, immersive environments that open up new ways for studying WA.

Our initial remote user study provided us with valuable insights into the UX and the potential benefits of having spatial context and interactive features for researching WA. The evaluation, however, also revealed interaction design challenges for creating these interactive, immersive applications for this novel area of research. Knowing that the completed IIS will be filled with large amounts of WA and corresponding information that must be easily accessible to users, it is especially important to address and investigate these challenges.

Within the scope of this dissertation, the following chapters will address these challenges and investigate possible solutions.

CHAPTER 5

SPATIAL USER INTERFACE FOR READING LONG TEXTS

The content of this chapter is based on published work [GLS23]. Some parts of this chapter have been rephrased and revised, and additional information, tables, and figures were included.

5.1 Motivation

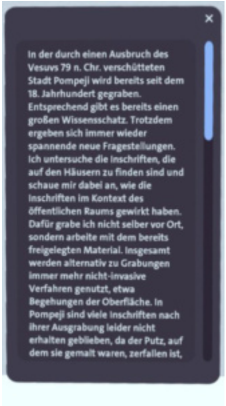
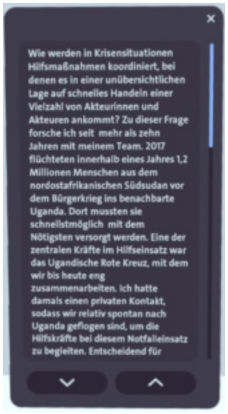
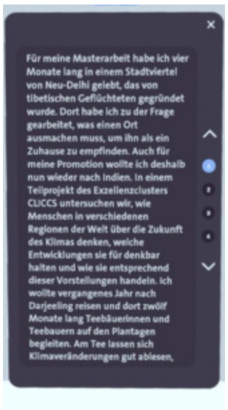
In the last chapter, we described the implementation of the IIS, whose requirements for spatial interaction techniques guided our work in this dissertation. From this chapter on, we report on our research for improving spatial interactions. First, we investigated interactions with long texts in immersive applications, based on our RQ-II:

RQ-II

Which spatial UI layouts and controls are suitable for reading and navigating through long texts in immersive VR applications?

For our specific context involving the spatial representation of hundreds of WA, a large amount of texts and descriptions have to be displayed on spatial UI panels inside the IIS. This requires researchers to access and interact with long texts in VR, especially in RFB05: The Interior of the Church in Lucklum (see Section 1.2.1), which uses long textual descriptions based on the publication *Sinnbilder im Sakralraum. Die Kirche in Lucklum – Ein Kompendium der geistlichen Emblematik der Frühen Neuzeit.* [SSA20]. Although there is some existing work on text colors and suitable text parameters, such as font size [DKO18; Eri+20; KKS21; Koj+20], research on long texts and suitable visualizations for spatial text display in VR is limited. So far, there have been no studies on suitable layouts, user interface (UI) patterns, and controls for reading and interacting with long texts in immersive applications. Findings in research on UI patterns for reading on desktop computers did not indicate any differences between scrolled versus paginated texts [Bak03]. To the best of our knowledge, this comparison has not yet been studied for text display in VR. To address this research gap, we investigated reading performance and UX for long texts with different UI panel layouts in VR. We designed four interactive spatial text panel variants for displaying long texts, as shown in Table 5.1. These panels use discrete and continuous text representations and respective UI controls for navigating through the texts.

Table 5.1: Four text panel variants with different UI controls using modified text excerpts from our university magazine "Neunzehn" [Ann21].

Continuous Scrolling		**Discrete Pagination**	
C.BAR	C.BUT	D.VER	D.HOR
			

5.2 Design and Implementation

We worked with two humanities researchers from our projects who study WA to identify requirements and to design and develop the spatial text panels. Both researchers were interviewed twice regarding their requirements and expectations for reading long texts in VR. Prior to our research projects, they had no experience with VR. As they usually work with analogue materials and 2D data on their computers, we designed the spatial text panels based on familiar 2D desktop UI patterns. We assumed that the researchers would be able to use the UI controls without prior experience due to the familiar look and interaction style of the interface [Jer15b]. Therefore, we choose to use a combination of two simple UI components: scrollbars and buttons. After discussing multiple design iterations of the text panels, we found that the researchers preferred the uninterrupted text flow, the scrolling speed, and the buttons for scrolled text. However, they also liked the overview and structure provided by paginated text with UI buttons. Based on this feedback, we adjusted, combined, and improved the UI controls, and adapted two established 2D desktop UI patterns for our use case: continuous scrolling and discrete pagination [TBV20; Tox21]. Based on these two patterns, we created four spatial text panel variants as pictured in Table 5.1:

1. Two continuous variants that display the text within a predefined view port in one continuous

flow that users can freely scroll through via a scrollbar or scroll-buttons. The variant C.BAR uses a vertical scrollbar, while C.BUT combines the scrollbar with two additional scroll-buttons (up/down) at the bottom of the panel.

2. Two discrete variants in which the text can be navigated through page by page, according to the pagination pattern. The variant D.VER uses a vertical pagination layout with individual, numbered buttons for each page, as well as up and down buttons for turning the pages. The alternative D.HOR has a horizontal pagination layout with the same setup of button controls.

The spatial text panels are implemented on *World Space Canvases* in Unity 2020.3 LTS. We used light text (white #FFFFFF) on dark background (panel #423E47, text area #3A343E) on all four designs, since previous work [DKO18; Eri+20] found that this leads to less fatigue and is preferred by readers over the reverse color scheme. For text rendering and formatting, we used the TextMeshPro package included in Unity.

For easier interaction in the continuous variants, the scrollbar's selectable bounds were setup with an increased width compared to the displayed visuals (UI sprite in Unity). The size and position of the scrollbar handle indicate the total text amount and the current position in the text. Additionally, the up/down buttons in C.BUT can be used for continuous scrolling or they can be selected briefly to perform fine adjustments when scrolling. In the discrete variants the text is divided into individual segments that can be navigated through page by page via the numbered pagination and arrow buttons. Again, the selectable bounds of the numbered buttons are bigger than the displayed visuals. Both panel variants only differ in the positioning of their UI controls and the direction of the virtual page turning animation (vertical versus horizontal movement).

5.3 User Study

Due to the differences in UI layout and interaction controls between the continuous and discrete text panel variants, we wanted to examine their influences on reading performance and UX. Since the researchers' feedback indicated that a discrete text display with pagination provides better overview and structure for the long texts, we were interested in testing potential differences compared to scrolled text with uninterrupted text flow. In our empirical study, our aim was to obtain initial insights into the text panel layouts and UI controls for reading long texts in VR. We investigated the recall (Can readers remember the information they read?) and spatial recall (Where was the information located?) of the text content, as well as qualitative aspects on UX. To test our assumptions, we formulated the following three hypotheses for our study:

H1 Discrete text display has a significantly better recall than continuous text display for reading long texts in VR.

H2 Discrete text display has a significantly better spatial recall than continuous text display for reading long texts in VR.

H3 Discrete text display has significantly better UX than continuous text display for reading long texts in VR.

5.3.1 Study Design

We used a within-participants study design and tested four different conditions corresponding to the four text panel variants introduced in Section 5.2: C.BAR, C.BUT, D.VER, and D.HOR. Participants had to read four texts, each with different content but equivalent in text length and reading difficulty. For the detailed description of the text parameter setup see Section 5.3.2. Each text was presented on a different text panel variant. For each text, three different content questions were chosen for measuring information recall after reading. In addition, spatial recall is recorded by asking for the location of the answers to the questions in the text. To account for serial positional effects, we choose questions corresponding to the content in the beginning, middle, and at the end of each text. To keep the difficulty level of the questions as similar as possible, we picked straightforward questions that could be answered with a short answer. For example, we asked about places, such as a city or geographic location, or about specific events, names, and equipment used. The questions used in the study can be found in Appendix A2.2. Unlike Rau et al. [RZG21; Rau+18], we chose to use open-ended answer text fields for the content questions instead of multiple-choice options, in order to avoid providing participants with any cues.

We administered two standardized questionnaires, the Simulator Sickness Questionnaire (SSQ) [Ken+93] and the User Experience Questionnaire (UEQ) [LHS08]. The questionnaires were setup with our university's online survey service, based on the open source tool LimeSurvey[1]. We also collected additional qualitative data with three open questions: (i) Please give a subjective ranking of the text panels, from good to bad. (ii) Please explain your ranking. What did you find particularly good or bad about each variant? (iii) Do you have any suggestions on how the user interface could be improved?

We randomized the order in which the conditions were shown to the participants, as well as the order of the four texts (Text 1–Text 4) that were assigned to the text panel variants. To counterbalance the condition order and the assignment of the text for each condition, we used the method of Zeelenberg and Pecher [ZP15] for constructing Latin squares.

5.3.2 Text Parameters

To ensure that our results in the four conditions are comparable with each other, we chose four different non-scientific texts with a similar Flesch Reading Ease (FRE) [Fle48]. FRE is a metric for measuring the readability and comprehensibility of a text, regardless of its content, with a score ranging from 0–100. We used modified excerpts from an old 2021 issue of our university's magazine *Neunzehn* [Ann21]. All four texts are in German and share the same overall topic: researchers presenting their international research projects. The FRE for German language is calculated according to Equation (5.1) [Pag20]:

$$FRE_{German} = 180 - ASL - (58.5 \cdot ASW) \tag{5.1}$$

[1] https://www.rrz.uni-hamburg.de/services/software/alphabetisch/limesurvey

where

$$ASL: \textit{Average Sentence Length}, \text{ average number of words per sentence}$$
$$ASW: \textit{Average Number of Syllables per Word}$$

The selected texts have a FRE between 47.4–50.8 ($M = 49.3, SD = 1.99$), which is rated as moderately difficult to read. All texts have 76 rows with a total word count of 351–366 ($M = 359.75, SD = 5.85$) and a *Mean* of 4.6–4.8 words per row. The detailed text parameters for each text are listed in Table 5.2. The original German texts that were used can be found in Appendix A2.1. Each text is displayed in a single paragraph on a UI panel in portrait view, with 24 rows of text within its viewport. The UI panel's position is fixed in *World Space* in VR according to the findings of Rzayev et al. [Rza+21] and Kobayashi et al. [KKS21]. We used our university's sans serif font (TheSans UHH) with *Regular* font weight and positioned the information panels at a distance of about 1.75m from the user. We chose a font size corresponding to 17.87dmm based on the findings of Kojić et al. [Koj+20] for displaying long texts in the Quest HMDs. The color contrast ratio between text to background is 11.87 : 1, which is well over the minimum contrast ratio of 4.5 : 1 defined by the World Wide Web Consortium (W3C) [W3C20]. The VR study environment was setup with a simple, neutral background to avoid any distractions, as shown in Figure 5.1.

Table 5.2: Text parameters and FRE for each of the four texts used in the study

	Rows Total	**Rows/Viewport**	**$\bar{X}$ Words/Row**	**Word Count**	**Char Count**	**FRE**
Text 1	76	24	4,6	351	2408	47,4
Text 2	76	24	4,7	358	2448	49,5
Text 3	76	24	4,8	364	2448	49,6
Text 4	76	24	4,8	366	2419	50,8

5.3.3 Study Setup and Procedure

The study was conducted in a seated, stationary VR setup, as pictured in Figure 5.1. In the beginning, participants signed a consent form and were informed about the procedure of the study. We used a Meta Quest 2 HMD and Quest Touch Controllers. Participants without VR experience received an introduction for the Quest HMD and its controls. Participants were instructed to read the texts carefully. They were informed that they had to answer multiple content-related questions and mark the spatial location of the corresponding information in the text on paper printouts. These printouts had stylized text lines and line numbers. Prior to starting the study, we handed out a sample printout to explain its usage. We also asked the participants to pay attention to the UI of the different text panels. Before entering the VR study environment, participants completed a demographic questionnaire and an initial SSQ for measuring their baseline simulator sickness. After reading each text, we recorded the recall of information as well as the spatial recall for the information's position as a measure for reading performance. In addition, participants marked the

Figure 5.1: VR study environment and stationary study setup.

corresponding spatial position of the text content for each answer on the printout. Next, they filled out the two standardized questionnaires. After finishing the study in VR, we collected additional qualitative feedback. Participants gave a subjective ranking for the text panel variants. They were asked to explain their ranking and list positive and negative aspects of each variant. Finally, they could suggest improvements for the text panels. All pre-, post-, and in-study-questionnaires were filled out on a PC. Participants were asked to re-center the view inside the HMD each time after they put the Quest 2 back on. In total, the study took approximately 60 minutes to complete.

5.4 Results

We performed both, quantitative and qualitative analyses of the collected data from the questionnaires and printouts. First, we present the quantitative results on recall and spatial recall for the text content, followed by the results of the standardized questionnaires. Finally, we report on the subjective user ranking for the text panel variants and the qualitative analysis of the user feedback. For plotting the significant differences in our quantitative results, we used the following indicators for the significance levels: $* < .05, ** < .01, *** < .001$.

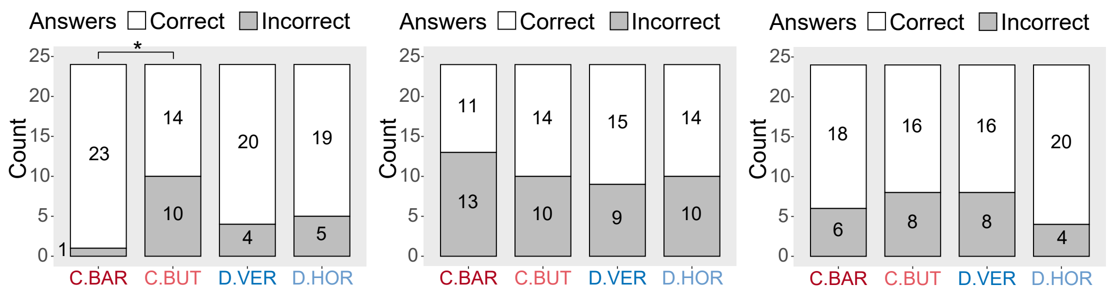
5.4.1 Participants

The participants were recruited from our university's mailing lists and were mainly students and researchers in the field of HCI. Students received study credits for their participation. We tested 24 participants (women=12, men=12) with 8 participants aged between 18–24, 12 between 25–34, and 4 between 35–44. All participants had normal or corrected to normal vision and native-speaker level proficiency in the German language. 7 participants had no prior VR experience. The other 17 participants had used a VR HMD at least once, with 2 having tried it once, 8 using VR not more than once per month, 4 multiple times a month, and 3 multiple times a week.

5.4.2 Recall

As shown in Figure 5.2, more content questions were correctly answered from the beginning and the end compared to the middle of the text for all text panel variants. We found no significant differences between the four variants with the Cochran Q-test for the middle and end of the text.

The ratios between correct and incorrect answers for each text panel variant are similar to each other, as illustrated in Figure 5.2. However, there was a statistically significant difference in information recall for the beginning of the text ($p = .021$). According to the post hoc McNemar test with Bonferroni correction, there is a significant difference between C.BAR and C.BUT ($p = .04$). As visualized in Figure 5.2(a), C.BAR has a ratio of correct to incorrect answers of 23 : 1 compared to C.BUT with 14 : 10.



(a) (b) (c)

Figure 5.2: Information recall: Correct and incorrect answers to the content questions for each panel variant: (a) beginning, (b) middle, and (c) end of the texts.

5.4.3 Spatial Recall

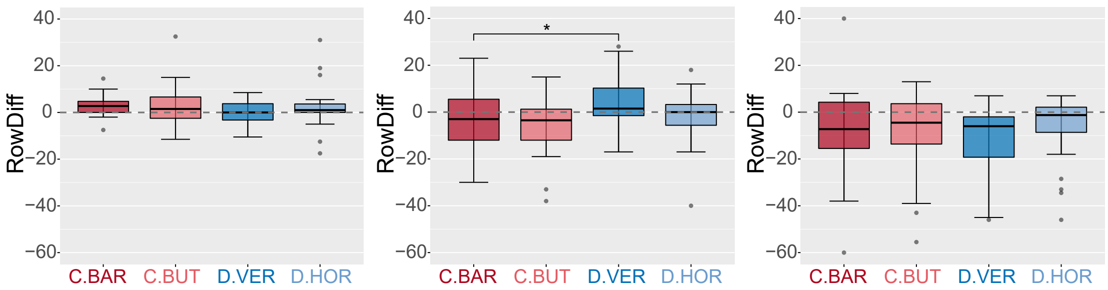
We compared the row difference *RowDiff* between the correct row number for the answer to a content question and the row numbers specified by the participants. A correct spatial recall would have a *RowDiff* $= 0$, a position too far down in the text would have a positive *RowDiff*, while a negative *RowDiff* indicates that the given position was too early in the text. Figure 5.3 shows that spatial recall within the interquartile range (IQR) generally lies between $+10$ and -10 *RowDiff*, except at the end of the text, where spatial recall is heavily skewed towards the negative numbers. This can also be seen in the *Mean RowDiffs* in Table 5.3, which lists the *Mean* and *SD* for *RowDiffs* for each text panel variant.

A normal distribution for *RowDiff* could not be assumed based on a Shapiro-Wilk test. The Friedman test revealed statistically significant differences in *RowDiff* for the middle part of the text ($p = .049$), but not for the beginning and end. A post hoc Wilcoxon signed-rank test with Bonferroni correction revealed a significant difference in the middle part of the text for *RowDiff* between C.BUT and D.VER ($p = .015$) (see Figure 5.3(b)). However, the absolute distance to the correct row number (*RowDiff* $= 0$) is similar between C.BUT (*RowDiff* $= -5.188$) and D.VER (*RowDiff* $= +4.396$).

We also looked into the absolute row difference (absolute number of rows from the correct row with the information) and performed an Aligned Rank Transform (ART) ANOVA. The result suggests that the different text panel variants do not have any significant effects on absolute row difference and that there is no significant interaction between the text panel variants and the different of the parts text. Yet, the result shows that there is a significant effect ($p < .001$) of position of the information (*beginning*, *middle*, *end*) on the absolute row difference. The ART-C post hoc pairwise comparisons with Holm's sequential Bonferroni correction indicate that C.BAR (beginning) in contrast to C.BUT, (end) are significantly different ($p = .048$). All other pairwise comparisons were not significantly different.

Table 5.3: Mean and SD for *RowDiff*, given in number of rows, between the actual row of the correct information in the text and the row that the participants specified.

	beginning		**middle**		**end**	
	M	*SD*	*M*	*SD*	*M*	*SD*
C.BAR	+2.938	4.571	-3.479	12.946	-7.167	18.516
C.BUT	+3.062	9.155	-5.188	13.414	-9.250	17.737
D.VER	-0.271	5.227	+4.396	11.909	-11.042	14.448
D.HOR	+2.354	9.455	-2.312	11.202	-6.938	14.307



(a) (b) (c)

Figure 5.3: Spatial recall: Row difference *RowDiff* between the correct row number containing the answer to each question and the number stated by the participants: (a) beginning, (b) middle, and (c) end of the texts for each panel variant.

5.4.4 Simulator Sickness Questionnaire (SSQ)

The results indicate an increase in *Mean* scores throughout the course of the study on all SSQ scales except for *Nausea*, as illustrated in Figure 5.4. The exact *Mean* and *SD* values for each subscale and the total SSQ score at each time point of measurement (pre-study baseline and after each trial) are listed in Table 5.4. Since normal distribution for the data could not be assumed (Shapiro-Wilk test), we performed Friedman tests on all three SSQ subscales, *Nausea*, *Occulomotor*, *Disorientation*, and on the *Total Simulator Sickness (TS)* score. We did not find any significant differences in SSQ scores between the trials (including baseline). While reading long texts in VR continuously increased symptoms, especially on the *Oculomotor Disturbance* scale (M = 12.95 at baseline, 19.90 after trial 4), it did not lead to a significant worsening of these symptoms.

Table 5.4: *Mean* and *SD* for each subscale and total SSQ score.

	NAU		OCU		DIS		TS	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Baseline	12.32	13.34	12.95	14.57	6.38	11.59	12.93	13.14
Trial 1	10.33	12.23	12.32	16.09	9.96	12.98	12.00	14.00
Trial 2	7.55	10.89	14.53	20.23	7.54	19.24	12.15	18.30
Trial 3	9.54	12.89	18.00	22.88	9.86	24.10	15.27	21.60
Trial 4	11.92	16.46	19.90	26.24	11.02	29.01	17.45	52.81

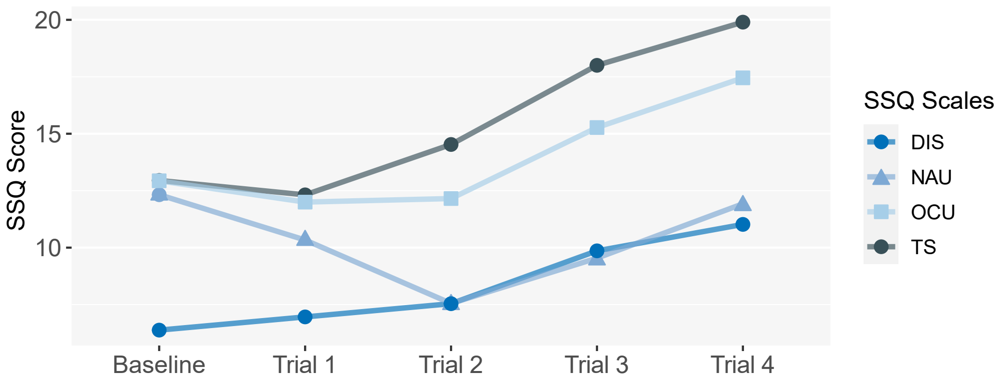


Figure 5.4: *Mean* SSQ scores after each trial for each subscale and the total score.

5.4.5 User Experience Questionnaire (UEQ)

Table 5.5: Mean UEQ ratings for each text panel variant. ▲ positive, ▶ neutral, ▼ negative.

	Attractiv.	**Perspicuity**	**Efficiency**	**Dependability**	**Stimulation**	**Novelty**
C.BAR	▶ 0.125	▲ 1.302	▲ 0.927	▶ 0.490	▶ 0.240	▼ -0.969
C.BUT	▲ 0.958	▲ 1.510	▲ 1.125	▲ 1.260	▶ 0.688	▶ -0.042
D.VER	▲ 1.083	▲ 1.938	▲ 1.313	▲ 1.583	▶ 0.573	▶ 0.063
D.HOR	▲ 1.354	▲ 2.271	▲ 1.906	▲ 1.823	▲ 0.833	▶ -0.427

Overall, the UEQ results are positive, with > 0.8 being considered as positive, -0.8 to 0.8 as neutral, and < -0.8 as negative rating for the UEQ. As listed in Table 5.5, the scales *Attractiveness*, *Perspicuity*, *Efficiency*, and *Dependability* received mostly positive ratings. The hedonic subscales *Stimulation* and *Novelty* were mainly rated neutrally, with D.HOR receiving a positive rating for *Stimulation* (0.833) and C.BAR receiving a negative rating for *Novelty* (-0.969). In comparison to the UEQ benchmark data set, only D.HOR received an excellent rating on the pragmatic scales, while C.BUT and D.VER achieved a good rating. All variants got a low rating on the *Novelty* scale compared to the UEQ benchmark. The rating is plotted in Figure 5.5, which shows that the text panel variant C.BAR received the lowest rating on all scales, whereas D.HOR scored the highest, except on the *Novelty* scale. D.VER is in second place and C.BUT places third in comparison. Although C.BUT has a slightly higher rating on the *Stimulation* scale compared to D.VER, it scored lower on all the other scales.

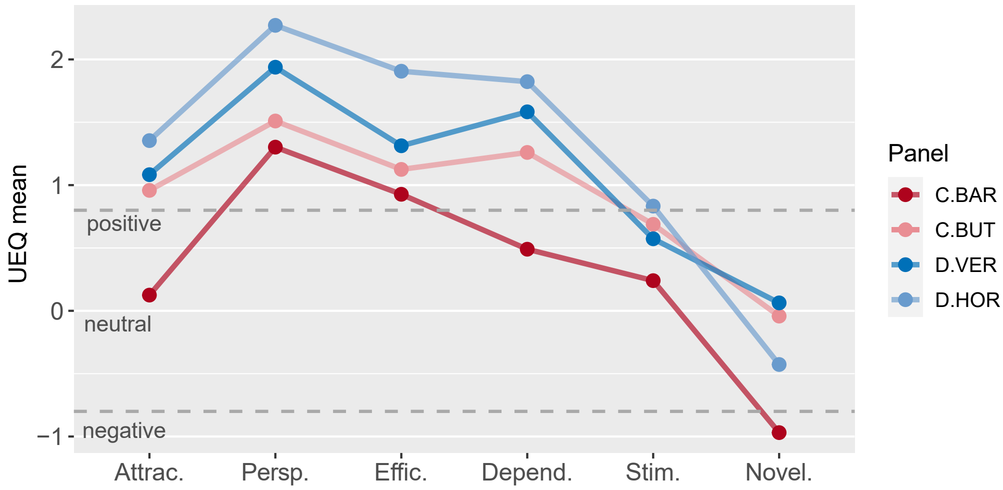


Figure 5.5: Comparison of the *Mean* ratings between the panel variants on all UEQ scales.

The *Means* for *Attractiveness* and *Efficiency* are normally distributed. We conducted a one-way repeated measurements ANOVA and found significant differences between the text panel variants for *Attractiveness* ($p = .002$) and *Efficiency* ($p = .006$). The pairwise comparison post hoc t-test with Bonferroni correction found statistically significant differences between the following text panel variants:

- D.HOR has a significantly higher *Attractiveness* rating than C.BAR ($p = .007$)
- D.HOR has a significantly higher *Efficiency* rating than C.BAR ($p = .003$)
- D.HOR has a significantly higher *Efficiency* rating than C.BUT ($p = .035$)

For the *Means* of the remaining four UEQ scales, we could not assume a normal distribution based on the Shapiro-Wilk test. We performed non-parametric Friedman tests, which indicated significant differences for the scales *Perspicuity* ($p = .019$), *Dependability* ($p = .02$), and *Novelty* ($p = .04$). The Wilcoxon post hoc tests with Bonferroni correction revealed the following statistically significant differences:

- D.HOR has a significantly higher *Perspicuity* rating than C.BAR ($p = .007$)
- D.HOR has a significantly higher *Dependability* rating than C.BAR ($p = .029$)
- C.BUT has a significantly higher *Novelty* rating than C.BAR ($p = .017$)

5.4.6 Subjective Ranking of Preference

Participants were asked to rank the text panel variants from highest (1) to lowest (4) based on their personal preference. As pictured in Figure 5.6, there was no clear primary preference for one variant. C.BUT, D.VER, and D.HOR got about a third of the votes each for rank 1, while only four participants voted for C.BAR. For the least preferred variant (rank 4), 16 participants (66.7 %) chose C.BAR and 6 (25 %) chose D.VER. The other two text panel variants only got one vote each for least preferred variant.

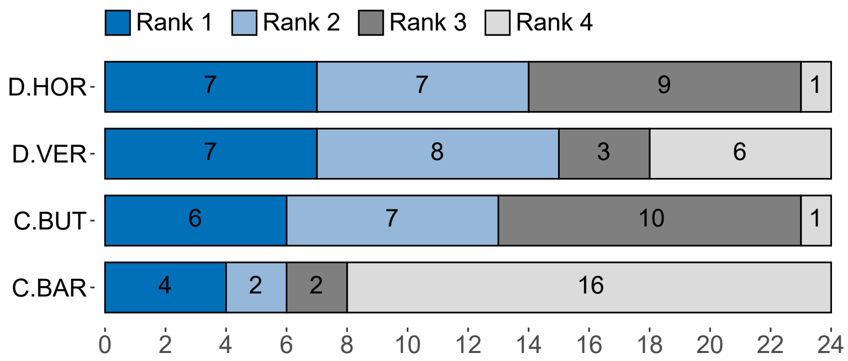


Figure 5.6: Subjective ranking with number of votes for each rank from 1 (highest) – 4 (lowest).

5.4.7 Qualitative User Feedback

We used a structured, inductive qualitative content analysis approach inspired by Mayring [May00]. We iteratively analyzed, sorted, and compared the qualitative feedback to identify reappearing themes in the statements of the participants. Finally, we consolidated these themes into the following five categories:

- Placement: size of UI controls or their placement on the text panel
- Usability: usability of individual UI controls
- View: direction of gaze or the orientation of the participant's head
- Orientation: reading flow and orientation within text
- Animation: perception of animated movements of text and UI elements

We systematically assigned the participants' statements on each text panel variant to these categories. Corresponding statements within a category were paraphrased and counted. The following four Tables 5.6 to 5.9 contain the results of the sorted and paraphrased statements in each category for each text panel variant, along with the number of counted occurrences listed under **N**. The most frequently mentioned statements are displayed in **bold** font.

Table 5.6: Sorted and paraphrased statements on text panel variant **C.BAR**.

Category	**Statements C.BAR**	**N**
Placement	Size of scrollbar better than in C.BUT	2
	Scrollbar is too imprecise/too sensitive and thus too difficult to control/exhausting to use	**14**
Usability	Easy and intuitive controls	3
View	*No statements*	-
Orientation	**Loss of orientation/impaired reading flow; jumping scrollbar handle requires searching for text passage**	**9**
Animation	*No statements*	-

Table 5.7: Sorted and paraphrased statements on variant **C.BUT**.

Category	**Statements C.BUT**	**N**
Placement	**Scrollbar size too small**	**4**
	Buttons are redundant	2
Usability	**Easy and intuitive controls**	**5**
	Better control by scrolling with buttons	3
	Scrollbar too sensitive	**4**
	Arrow buttons suggest single pages instead of scrolling	2
	Accidentally mixed up the buttons	2
View	**Text comfortably readable at eye level**	**4**
Orientation	**Loss of orientation due to fast scrolling via buttons**	**4**
Animation	**Smooth, pleasant scrolling speed via buttons**	**4**

Table 5.8: Sorted and paraphrased statements on variant D.VER.

Category	**Statements D.VER**	**N**
Placement	**UI controls placed on the right side of the text panel are easily accessible**	**4**
	UI controls on the right side not good/not well visible	3
Usability	**Easy and intuitive controls, straightforward**	**6**
	Pleasant forward and backward browsing of pages and precise control of individual pages	3
	Non-intuitive/vertical UI layout suggests scroll function	2
View	Uncomfortable head position when reading text in the lower area of the text panel	2
Orientation	**Loss of orientation due to interrupted sentences when changing pages; text between two pages not visible at the same time**	**4**
	Clear structure through division into pages	**4**
Animation	**Vertical animation is non-intuitive/irritating**	**4**
	Animation is too fast	2

Table 5.9: Sorted and paraphrased statements on variant D.HOR.

Category	**Statements D.HOR**	**N**
Placement	**Placement of UI controls at the bottom of panel impractical to reach/too far away**	**7**
	Placement of UI controls at the bottom easy to understand, clear, pleasant to use	2
Usability	**Simple, easy, straightforward to use**	**8**
	Book metaphor is pleasant	**4**
View	Uncomfortable head position when reading text in the lower area of the text panel	2
Orientation	**Clear structure of the individual pages**	**5**
	Structuring of text into individual pages is more helpful for remembering information	2
	Loss of context/orientation due to hard transition between pages	3
Animation	*No statements*	-

Additionally, Table 5.10 displays general feedback, including contains feature and interaction suggestions from the participants.

Table 5.10: Sorted and paraphrased General Feedback.

Category	Statements General Feedback	N
Placement	Font size and distance should be customizable	2
Usability	**Additional control of the texts by thumbstick**	**5**
	Speed of scrolling adjustable by pressing button, by applying different amount of pressure on trigger button or thumbstick	**5**
	Scrolling text by grabbing the text area	3
	Using pagination/flipping pages by grabbing text or by using thumbstick	2
	Combination of scrolling and pagination	2
View	Thumbstick controls in addition to UI controls, so text can be controlled without taking your eyes off the text while reading	2
Orientation	*No statements*	-
Animation	*No statements*	-

To summarize the most important findings, we created an overview of positive and negative statements with the highest occurrences ($n > 3$) for each text panel variant in Table 5.11. C.BAR received negative statements with the highest overall occurrence and no positive statements with $n > 3$. Both negative statements are related to the bad usability of the scrollbar. More than half of the participants ($n = 14$) expressed that it is too imprecise/sensitive, and 9 stated that the jumping movement of the scrollbar leads to a loss of orientation when scrolling. The other three variants received positive overall feedback in terms of usability. Nevertheless, the usability of the scrollbar was also criticized in C.BUT. Some participants stated that fast scrolling with the up and down buttons resulted in a loss of orientation, while others enjoyed the speed and smoothness of the scrolling. Both discrete text panel variants received positive feedback for the clear structure provided by the pagination. But participants reported loss of orientation as well as unintuitive and irritating animation for the vertical pagination in D.VER. While some participants stated that they appreciated the placement of the UI controls for D.VER for being easy to access, the position of the UI controls for D.HOR at the bottom of the text panel was criticized for being impractical and too far away. Four participants stated that they liked the book metaphor of D.HOR.

As general feedback, participants requested additional controls for scrolling and pagination, for example with the thumbstick of the VR controller. Moreover, they suggested adjusting the scrolling speed with the thumbstick or the trigger button, so that the reading flow is not disrupted by looking at and interacting with GUI controls in VR.

Table 5.11: Qualitative content analysis – Positive and negative statements with the highest occurrences ($n > 3$).

Panel	Overview of Statements	N
C.BAR	- scrollbar imprecise/too sensitive, difficult to use/exhausting	14
	- jumping scrollbar: loss of orientation/impaired reading flow/searching for text passages	9
C.BUT	+ simple, intuitive	5
	+ text comfortably readable at eye level	4
	+ smooth, pleasant speed of scrolling with up/down buttons	4
	- scrollbar too small and too sensitive	4
	- loss of orientation due to fast button scrolling	4
D.VER	+ easy to use, intuitive, straightforward	6
	+ control elements on right side easy to access	4
	+ clear structure through pagination	4
	- loss of orientation when changing pages; text between two pages not visible at the same time	4
	- vertical pagination animation is not-intuitive, irritating	4
D.HOR	+ simple, easy to use, straightforward	8
	- position of control elements at bottom of panel is impractical/too far away	7
	+ clear structure through pagination	5
	+ book metaphor is good	4
General	• additional control of texts with thumbstick, e.g. for scrolling and pagination	5
	• adjusting scrolling speed with thumbstick or with trigger button pressure	5

5.5 Discussion

We designed four different spatial text panels for displaying long texts in VR and investigated the reading performance and UX in an empirical user study. The results provide valuable insights about the strengths and weaknesses of each text panel variant for reading long texts in VR. In the following, we discuss the results of the tested hypotheses and their implications.

H1: Discrete text display has a significantly better recall than continuous text display for reading long texts in VR.

Our study results do not support our H1. The different text panel variants only showed statistically significant differences for recall between the the two continuous variants C.BAR and C.BAR at the beginning of the text. We did not find any significant differences in recall between continuous and discrete text panel variants. However, this finding is consistent with Baker [Bak03]'s findings for reading on desktop computers. It should therefore be further investigated if other findings in existing research for reading on desktop can also be applied to reading in VR. Further insights into this area could help with transferring and adapting already established desktop UI patterns for VR.

Figure 5.2 shows that more correct answers were given for content questions at the beginning and at the end of the texts. This observed tendency is consistent with the assumptions of serial position effects [Mur62]. This observed effect suggests that participants did read the texts carefully, which would increase the reliability of our results.

H2: Discrete text display has a significantly better spatial recall than continuous text display for reading long texts in VR.

We also have to reject H2. Although there was a significant difference in spatial recall between relative *RowDiff* for C.BUT ($= -5.188$) and D.VER ($= +4.396$) in the middle part of the text, the absolute row difference from the actual position of the information (*RowDiff* $= 0$) did not differ significantly. This indicates that no variant performed more accurately or better at this task.

While there were no statistically significant differences, it is still interesting to observe that spatial recall was most accurate for content at the beginning of the text and least accurate at the end, where participants had the tendency to select a row number that was too low. This can be seen in Table 5.3 and Figure 5.3. This result suggests that, for reading on spatial text panels in VR, some serial position effects might apply to spatial recall.

H3: Discrete text display has significantly better UX than continuous text display for reading long texts in VR.

Based on our results, we can accept H3. Both discrete text panel variants received higher ratings for *Attractiveness* and for the three pragmatic UEQ scales, with D.HOR having significantly higher ratings than C.BAR on all these scales (see Figure 5.5). All text panel variants received a below average or bad rating in the UEQ benchmark for the hedonic qualities *Stimulation* and *Novelty*. This was expected, since the designs were intentionally based on familiar 2D desktop UI patterns. According to this result, we assume that the used UI patterns can be easily understood by users without prior VR experience. Nonetheless, all variants except for C.BAR received good or even excellent ratings for the pragmatic quality compared to the UEQ benchmark data set. C.BAR received the lowest overall ratings of all panel variants, which suggest that this panel design might not provide good UX.

Our additional qualitative feedback supports these UEQ findings. The subjective ranking, as displayed in Figure 5.6), showed no clear preference for one text panel variant, but the majority of participants ranked C.BAR in the last place (66.7 %). This indicates that it is clearly the least preferred panel variant. The qualitative content analysis of the participants' feedback gave us valuable insights into the causes for C.BAR's the bad rating. The major issue seems to be the poor usability of the scrollbar in VR due to its imprecise control and high sensitivity, which makes it difficult and unpredictable to use. The control issues could have been exacerbated by the indirect, raycasting-based far interaction with the UI due to the increased Index of Difficulty (see Section 2.3). This insight suggests that scrollbar-only text navigation is not recommended for interacting with long texts in VR. The scroll-buttons in C.BUT received some positive feedback. Participants liked the scroll speed and the intuitive controls that enabled them to keep the current text at eye level while reading. On the other hand, the button scrolling speed was too fast for some participants, which caused them to scroll too far and loose their orientation within the text. Generally, participants stated that the discrete variants were easy to use, intuitive, and had a clear structure due to the pagination. However, some were irritated by the vertical layout and animation of the pagination in D.VER. This would explain the slightly lower UEQ rating of D.VER in comparison to D.HOR, which follows the established, widely used horizontal pagination layout of desktop interfaces [TBV20].

Based on our findings, we recommend using the horizontal pagination pattern instead of a scrollbar for displaying and interacting with long texts in VR, especially for text lengths that would otherwise require extensive scrolling to read. Continuous and discrete versions each seem to have their advantages and disadvantages, as participants listed various aspects they liked and disliked for each of the text panel variants. Depending on the use case or individual preference, a certain text panel variant may be favored or more suitable. Thus, users should be able to choose different options and set their individual preferences within the immersive application. In the qualitative feedback, participants also included feature suggestions, such as using the VR controller's thumbstick to control the text flow and speed. This shows that it is important to further investigate and integrate additional, alternative interaction methods, including non GUI options, for reading and interacting with long texts in VR.

5.6 Limitations and Future Work

The study was conducted with participants who were mainly from our computer science faculty and who have a background in HCI. While this gives us general insights about the feasibility, preferences, and issues of the tested text panels designs, follow-up evaluations should also include researchers from different fields in the humanities to obtain more specific results. For future studies, the virtual study environment as well as the presented long texts should be chosen based on the respective IIS and the research focus of the humanities researchers who study specific WA.

We cannot rule out that interacting with the UI controls might have distracted the participants from reading the texts and vice versa, which might have influenced our results. We also cannot rule out that the carefully chosen texts and content questions still had varying levels of difficulty for individual participants. It is generally challenging to choose texts that are as comparable as possible, as interest in specific content, reading comprehension, and attention span vary from person to person. In the future, it might be useful to follow a more practical and qualitative approach with user observation, free exploration, and thinking-aloud in various VR reading scenarios.

We observed that participants underestimated the location of information at the end of the text in the spatial recall task. Future studies could look into possible causes and generally perform a more detailed investigation into spatial recall in VR. This could be combined with built-in eye-tracking on the used HMD to provide additional insights.

As participants suggested additional alternative UI text panel controls, including non GUI solutions, future work could also focus on designing, developing, and evaluating new approaches for displaying and interacting with long texts in VR. In addition, the traditional scrollbar pattern used in 2D desktop UIs could be further improved and extended to facilitate spatial interaction in VR environments. Generally, our UI canvas based text panel designs did not fully utilize the 3D and spatial capabilities of VR. Future work could therefore study the use of 3D volumetric text panels and UI controls, as well as how information and long texts could be displayed on different depth layers.

5.7 Conclusion

In this chapter, we presented four different spatial text panel designs consisting of two variants with continuous and two variants with discrete UI controls for reading and interacting with long texts in VR. The controls were designed based on two established, familiar 2D desktop UI patterns: scrollbar and pagination. In our empirical evaluation, we did not find any significant differences in reading performance for the four variants. Moreover, none of them was clearly the most favored in terms of user preference. The classic scrollbar pattern used in C.BAR was, however, the least preferred variant due to its poor usability. D.HOR with its classic horizontal pagination pattern received the highest UX ratings regarding the pragmatic quality. Our findings indicate that desktop UI patterns can be transferred and used for creating spatial text panels for immersive applications. Nonetheless, our qualitative analysis revealed that there are still many challenges for reading and interacting with long texts in VR and that suitable UI patterns and guidelines need to be further investigated. In addition, individual user preferences as well as potential use case specific requirements should be further examined. Our work also demonstrates the importance of following a UCD approach and the collection of qualitative user feedback for research on spatial interaction and UI patterns in VR. The additional qualitative insights helped us to better understand the user preferences and pain points, including the causes for poor UX. We believe that these findings can contribute to facilitating future development and research on suitable UI patterns and improved interaction design for reading long texts in VR.

CHAPTER 6

ASSISTIVE CONTROLLER-BASED RAYCAST REDIRECTION

This chapter is based on published work [Gab+23] and contains rephrased, revised, and extended paragraphs with additional information, tables, and figures.

6.1 Motivation

After concluding our investigation on long texts in immersive applications, we address the next interaction challenge in this chapter, guided by our RQ-III:

> **RQ-III**
>
> How can we implement an assistive raycasting interaction technique that redirects the raycast towards an intended target to facilitate distant selection of spatial content?

Spatial interactions are required for accessing the research data on the hundreds of WA that are located throughout the true-to-scale 3D environment in the IIS. Raycasting is a widely used and easy to understand interaction technique for far interactions outside the direct interaction range of the user [LaV+17]. Unfortunately, the speed and accuracy for raycasting-based selection significantly degrades with decreasing target size and increasing target distance [LaV+17; Pou+98]. This affects the users of our IIS. Due to the distribution of the WA inside our IIS, they have to perform selections of various WA markers over larger distances. As described in Section 2.4, existing adaptions for classic raycasting restrict user control, diverge from the classic pointer metaphor, or increase the interaction complexity. This is not suitable for our context, as it includes first-time and novice users, as well as public presentations without additional time for long explanations and practicing spatial interactions. To address this issue, we propose three assistive raycasting techniques as an extension to classic raycasting: *RayRotation*, *RayCurve*, and *RayCurveRotation*. These three techniques differ in how they map the visible raycast based on the calculated raycast redirection (see Section 6.2). Our techniques do not add additional interaction complexity and stay close to the original pointer metaphor. Our goal is to improve spatial interaction for distant selection tasks with assistive raycasting by using subtle redirection of the tracked raycast towards the target object. Compared to existing, more restrictive techniques with automatic, instant snapping to the target center with no control over the snapping interaction [Doc; SRH06], we apply a gradual and subtle, proximity-based redirection within a pre-defined *redirection zone*. Our approach results in a magnetic pull behavior of the raycast towards the target. Users can still freely control the raycast by moving it in and out of the *redirection zone*. Our assistive raycast redirection techniques are intended to be easy to use and to provide a similar experience to interaction with classic raycasting. At the same time, they should maintain a comparable sense of agency over the interaction, while improving selection performance.

For evaluating our proposed raycast redirection techniques, we compared them to classic raycasting. To gain a more in depth-understanding of selection performance as well as subjective aspects, we combined quantitative performance measurements with qualitative approaches to provide insights into perceived workload, sense of agency, UX, pain points, preferences, and user acceptance.

Considering that this type of proximity- and gain-based rotational redirection, to our best knowledge, has not been previously studied for facilitating raycasting-based selection, our findings provide the following contributions for XR researchers as well as practitioners:

- **C1** Empirical evaluation of performance for assistive redirected raycasting-based selection techniques
- **C2** Comprehensive insights into various subjective aspects for spatial object selection, including UX and user acceptance for redirected raycasting
- **C3** Recommendations for the design and future research on redirected raycasting as an assistive interaction technique

6.2 Raycast Redirection Implementation

6.2.1 Redirection Behavior and Calculation

Our goal for the implementation was to create an easy to use raycasting-based assistive selection technique that serves as an extension of classic raycasting. Redirected raycasting should be easy to learn, maintain the same sense of agency like classic raycasting, but offer increased selection performance through the subtle, assistive redirection of the raycast towards the target.

To achieve this subtle effect, we implemented the redirection based on a simulated magnetic behavior. Our calculation uses the tracked raycast, which is hidden from the user, and gradually moves a second, visible raycast towards the target center until it finally snaps to it. As there is no existing data or evaluation on suitable raycast visualizations for redirected raycasting, we implemented three different visualizations for the visible raycast. Existing techniques that use snapping and adjusted raycast visualizations are more restrictive (cf. Section 2.4.1). *Sticky Ray*

Algorithm 1: Pseudocode for the Raycast Redirection Calculation

Input: raycast origin position $\vec{R}_o$, unit raycast direction vector $\hat{R}_d$, raycast anchor rotation $\mathbb{H}_a$, target center position $\vec{p}_t$, redirection zone radius r_{rz}, gain factor g

Output: redirected rotation $\mathbb{H}_r$

Procedure ROTATIONLERP($\vec{R}_o$, $\hat{R}_d$, $\mathbb{H}_a$, $\vec{p}_t$, r_{rz}, g)

$d_t = \text{Distance}(\overrightarrow{R_o R_d}, \vec{p}_t)$ ▷ *distance raycast to point*

$step_{adj} = \text{InverseLerp}(r_{rz}, 0, d_t) \cdot g$ ▷ *distance adjusted step*

$\hat{t}_{dir} = \vec{p}_t - \vec{R}_o$ ▷ *direction to target*

$\mathbb{H}_t = \text{RotateInDirection}(\hat{t}_{dir})$ ▷ *rotation to target*

$\mathbb{H}_r = \text{Lerp}(\mathbb{H}_a, \mathbb{H}_t, step_{adj})$ ▷ *redirected rotation*

return $\mathbb{H}_r$

and the *Distance Hand Grab Interactor* both rely on fixed, automatic, and direct snapping of the visualized raycast to the target center, so users have no control over the snap interaction. Furthermore, they apply a strong curvature to the raycast. Approaches like the *Flexible Pointer* require both hands for complete manual raycast (movement and curvature) control, which adds more complexity and places higher demands on the movement coordination for the interaction. In our approach, we use a gradual and subtle redirection within a pre-defined *redirection zone* around the target. The visualized raycast is gradually pulled towards the target position, with increasing redirection strength based on proximity and a constant gain factor. This gain-based approach is inspired by existing techniques used for locomotion or limb redirection (cf. Section 2.6). With this implementation, the redirection is eased-out when getting close to the target until the pointer finally snaps to the target center. Since this type of redirection is proximity-based, it is also gradually decreased when users move away from the target. The visible raycast returns to a 1:1 mapping of the tracked raycast as soon as it leaves the *redirection zone*. At any point during the redirection, users are able to move the pointer in and out of the *redirection zone*. This provides more control to the users and enables them to freely and continuously move the visible raycast.

For the redirection calculation, we modified a linear quaternion lerp function that uses a proximity-based gain value as the step parameter for the interpolation between the actual, tracked raycast and the target position for the redirection of the visible raycast. In addition, instead of using linear interpolation, we changed it to ease-out at the end. The redirection is applied by rotating the visible raycast into the calculated direction. Algorithm 1 provides a pseudocode of this calculation. The redirection is triggered as soon as the actual raycast (based on the tracked VR controller) intersects with the pre-defined *redirection zone* of the target, as illustrated in Figure 6.1. The $step_{adj}$ parameter determines the redirection speed towards $\vec{p}_t$ (target center position) and the strength of the snapping at the target center. The speed of the redirection towards the target corresponds to the perceived magnetic strength of the visible raycast. The calculation of the redirected raycast is updated in fixed intervals. To ensure that the redirection is frame-rate independent, the updates are

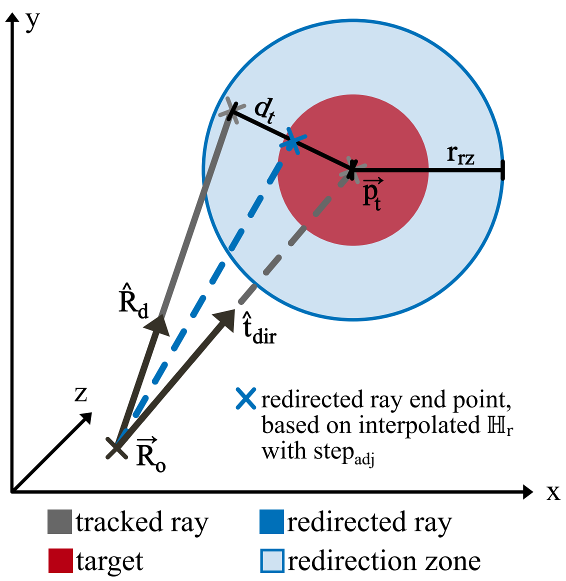


Figure 6.1: Illustrated target setup and calculation for the redirection interpolation.

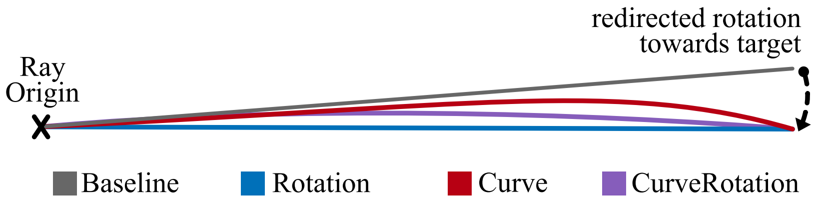
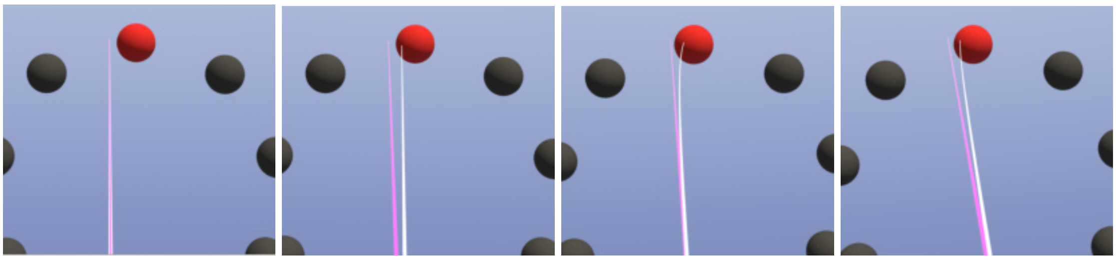


Figure 6.2: Illustration of visible raycast visualizations for each redirected *Controller-Ray Technique* compared to classic raycasting as the *Baseline*.

scaled by the *deltaTime* of the system (time elapsed since last frame was rendered). The easing effect is achieved by passing the current tracked raycast anchor rotation to the interpolation function instead of a fixed starting value.

Since there is no existing literature or recommendation for this exact type of redirection for raycasting-based interaction techniques, we conducted an informal exploratory testing to determine suitable values for the gain factor g and the size of the redirection zone rz. We asked four HCI researchers (expert users) and four members from the UWA cluster (novice users) to try out the three different raycast redirection techniques. We provided them with the same VR environment and setup as for our user study. The users could freely adjust the g and rz parameters through a special settings panel. The redirection was not required to stay below a detectable threshold, because our objective was to develop raycast redirection as an assistive technique for spatial selection in VR. Hence, we decided to base the redirection parameters for our user study on the initial user feedback that we received. We used a gain factor g of 1.5 and a redirection zone rz that is 2× the selection target size (uniformly scaled).

For the implementation of the visible selection raycast, we used the *LineRenderer* component in the Unity engine. The starting point for the raycast (raycast origin) is set to the anchor transform



(a) (b) (c) (d)

Figure 6.3: The four *Controller-Ray Techniques*: (a) Baseline, no redirection, (b) RayRotation, (c) RayCurve, and (d) RayCurveRotation. Magenta pointer: actual raycast based on tracked VR controller, hidden from the user and white pointer: visible, redirected raycast.

on the VR controller. Based on the calculated rotational redirection, we apply three different techniques to match the visualized raycast to the redirected raycasting direction. Since all three raycast redirection techniques use the same g and rz parameters, the end point position of the visible raycast is redirected by the same amount in each technique and the visible raycast ends at the same position in 3D space. The techniques only differ in how the redirected raycast is rendered from its start to its end point, as illustrated in Figure 6.2. We implemented the following three visualizations for our assisted raycast redirection:

6.2.2 RayRotation

RayRotation is the most straightforward visualization and uses a direct, 1:1 mapping of the calculated rotational redirection with a straight, non-curved visible raycast. The rotation of the raycast's anchor transform is set to the redirected rotation, which moves the end point of the visible raycast to the desired redirected pointing position. This end point is used for the raycast intersection with the target object. The start- and end-points of the visible raycast are not modified. The raycast is rendered in a straight line starting from an origin point, as displayed in Figure 6.3(b). This technique results in a slight visual offset between the actual tracked pointing direction and the rendered redirected raycast. This offset could potentially be noticed by users and might negatively affect them during the selection interaction.

6.2.3 RayCurve

The visible raycast for *RayCurve* starts out following the actual tracked pointing direction. Only the end of the pointer slightly curves towards the redirected position. This technique has no visual offset, but the curvature could potentially impact users negatively, since it deviates from the original pointer metaphor. Same as in *RayRotation*, the redirected rotation is applied to the anchor transform of the raycast, The ray's curvature is based on a cubic Bézier curve defined by four control points, as illustrated in Figure 6.4:

1. P_{start} = raycast origin
2. P_1 = control point

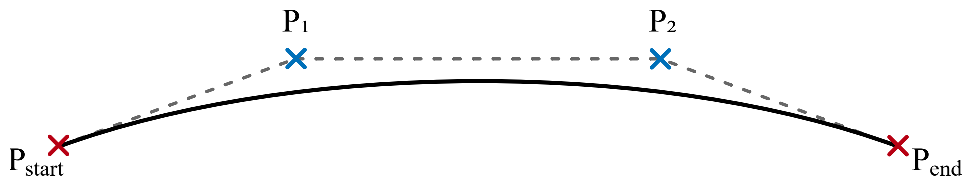


Figure 6.4: Illustrated cubic Bézier curve configuration for *RayCurve* and *RayCurveRotation*.

3. P_2 = control point
4. P_{end} = raycast endpoint

P_1 and P_2 are placed along the raycast at a distance of 30% of the raycast's length from P_{start} and P_{end}, respectively. As the *LineRenderer* of the raycast is anchored to the VR controller, this anchor is its parent transform. We implemented the curvature of the raycast by applying the inverse of the redirected rotation to the transform of P_{start}. This moves P_{start} back to the non-redirected, tracked rotation of the controller (anchor transform). Subsequently, the local rotation of P_{start} (within parent anchor transform) is mirrored to the local rotation of P_{end}. As depicted in Figure 6.3(c), this results in a visualized raycast with an increasing outwards curvature when it is moved towards the target center. Nevertheless, the overall curvature of the raycast is relatively subtle.

6.2.4 RayCurveRotation

To account for potential drawbacks resulting from the visual offset and the raycast curvature in the two other raycast visualizations, we designed *RayCurveRotation* as a technique with an intermediate visualization. It combines the two previous approaches by using a midpoint between rotation and curving of the visible raycast towards the center of the target object. First, the redirected rotation is applied to the anchor of the raycast. Next, the transform rotation of the Bézier's P_{start} is set to the half-way point of the redirected rotation. Its local rotation is mirrored to the local rotation of P_{end}, which moves the visible raycast to the midpoint between *RayRotation* and *RayCuve* and thus, reduces the curvature. To calculate the half-way point for this redirection technique, the last step in the procedure of Algorithm 1 is copied and $step_{adj}$ is multiplied by 0.5. As illustrated in Figure 6.2, the resulting redirection has a smaller rotational offset from the real raycast and a very subtle curvature. The composite Figure 6.3(d) provides a comparison of all the raycast visualizations for each of the three redirection techniques and classic raycasting.

6.3 User Study

We conducted a user study to investigate the effectiveness, UX, and the mentioned potential drawbacks of each raycast redirection technique and the used visualizations. The study also serves as an initial evaluation for the application of assistive redirection to raycasting-based interactions. Our investigation compares our redirection techniques with classic raycasting as the *Baseline* (Figure 6.3(a)) and is motivated by the following three research questions:

RQ1: Do assistive raycasting techniques improve selection performance and UX?

This type of raycast redirection has not been studied before. Although we assume that assistive raycasting would outperform classic raycasting, we do not know whether the applied redirection could also have a negative impact on the performance. For example, it could potentially be distracting or unfamiliar to users, or users could perceive it as too invasive.

RQ2: Are there significant differences in measured selection performance between the three proposed redirection techniques?

Considering that we use the same redirection parameters, the underlying redirection amount is always the same. We want to investigate whether our implemented redirection is robust, or if different raycast visualizations could have an influence on performance. Due to the visual dominance in human perception, we assume that this might be a possibility.

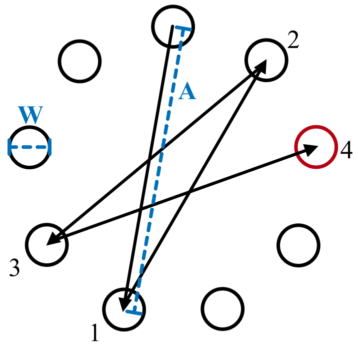
RQ3: Do UX and individual preferences differ for the three redirection techniques?

Aside from objective measurements, we believe that qualitative aspects are as important for evaluating and improving selection techniques. To gain a more in-depth understanding of users, we investigate potential individual preferences and identify subjective factors that might have an influence on the UX of assistive redirected raycasting.

6.3.1 Fitts' Law Selection Task

In order to obtain standardized and robust selection performance measurements during our study, we used a standardized Fitts' Law multi-directional selection task, as specified in Section 2.7.3. As visualized in Figure 6.5(a), our setup consisted of 9 target spheres in a circular arrangement. The target spheres are highlighted in red, one at a time, in a fixed, multi-directional selection pattern. To avoid controller selection induced pointing errors (Heisenberg Effect of Spatial Interaction [Bow+01]), we decoupled pointing and confirmation of the selection. A controller-anchored raycast extending from the dominant hand is used for pointing, while the selection confirmation is done by pressing a button on the VR controller in the non-dominant hand. If the raycast intersects with the current, highlighted target when the selection is confirmed by the button press, the selection is considered successful. Otherwise, an error is recorded. The participants did not receive any indication as to whether their selection was successful. After a participant performed a button press, the next target in the selection sequence (on the opposite side of the current target sphere) is highlighted. This selection process continues until the sequence is completed.

We chose a fixed z-distance (4m) for the targets, since our investigation focuses on selections outside the near interaction range. A distance of 4m does not necessarily require teleportation, as it lies within the recommended content zone ($1.25 - 5m$) according to guidelines [Pau+]. A viewing distance of < 10m also ensures strong stereoscopic vision in VR [Alg15]. To gain a more in-depth understanding of the performance and to validate the robustness of our redirection techniques for different object selection parameters, we defined four different *Indices of Difficulty* (IDs) based on two target widths W and four target distances A (movement distance in Fitts' Law). This resulted in four angular target sizes, as listed in Table 6.5(b). The two widths W correspond to small and large object sizes in VEs and their angular sizes match parameters that were used in previous related work [Ber+21a].



(a)

ID	W	A	Angular Size
3	0.3	2.1	4.15°
3.5	0.3	3.094	4°
4	0.1	1.5	1.4°
4.5	0.1	2.163	1.38°

(b)

Figure 6.5: Fitts' Law selection task design and parameters: (a) Illustration of the selection task and (b) *ID* listed with parameters.

6.3.2 Questionnaires

Pre-Study Questionnaire

Our pre-study questionnaire contains questions on demographics, usage frequency of interactive VR applications with controller-based interactions, and general VR experience.

Study-Block Questionnaire

We used the RAW NASA Task Load Index (NASA-TLX) [Har06b] to measure the subjective, perceived workload. We also provided four statements on perceived sense of agency that are rated on a 7-point Likert scale (*Strongly Disagree – Strongly Agree*), adapted from Argelaguet et al. [Arg+16] and Ma et al. [Ma+21]:

S1 The movement of the virtual pointer was caused by me.

S2 I felt like I was controlling the virtual pointer.

S3 The virtual pointer moved the way I wanted it to move.

S4 When I moved the VR controller, I expected the virtual pointer to move in the same way.

Post-Study Questionnaire

We used a semi-structured post-study questionnaire. Participants were asked to describe any differences in the performed selection interactions, if they noticed them. In addition, they provided a subjective ranking for the selection in each study block, including explanations for their ranking. Afterwards, we revealed the redirection techniques and showed print-outs with images similar to Figure 6.3, displaying the *actual* versus the *redirected* raycast position. Participants could then assign each technique to a study block based on their assumptions, and provide further feedback for each technique.

6.3.3 Study Setup and Procedure

The redirection techniques and study environment were implemented with the Unity engine (version 2020.3 LTS) and the Oculus Integration (version 51.0). The study was conducted on the Meta Quest Pro HMD with the self-tracking Touch Pro VR controllers in a stationary VR-setup.

We followed a within-participants design, where each participant tested all four conditions. Our independent variable *Controller-Ray Technique* corresponds to the three raycast redirection techniques, as described in Section 6.2 (*RayRotation*, *RayCurve*, *RayCurveRotation*), and a *Baseline* condition using classic raycasting without redirection. The four conditions were presented in subsequent blocks. In each block, the selection task was configured with the four IDs defined in Table 6.5(b), with each ID occurring exactly five times. This amounts to $4 \times 5 = 20$ tasks per block. The condition order was pseudo-randomized based on the repeated measures Latin-square design. The order of the selected IDs for the 20 tasks within a block was randomized with the Fisher-Yates shuffle algorithm [Ebe16]. A mandatory training block containing two tasks (with ID 3 and ID 4) had to be completed before starting the actual study blocks. The short training took about 30 seconds and was integrated to allow participants to familiarize themselves with the task and VR controls. It was restricted to only two tasks to minimize practice effects.

Before starting the study, participants gave their informed consent and were briefed about the study procedure. Next, they filled out the pre-study questionnaire on a PC and received a short introduction for the HMD and VR controls. The selection task was explained and participants were instructed to perform the selections as accurate and as fast as possible. In total, they had to complete four study blocks in total and answer questions about the selection after each block. Participants were seated inside the VR study space. We assisted them with correctly putting on the HMD and re-centering their view in the VR study environment. If participants completed the training block and had no further questions, they were instructed to start the actual study. After each block, participants answered the study-block questionnaire on a close-range UI panel in VR. Once participants completed the fourth block and the corresponding questionnaire, they filled out the post-study questionnaire on a PC. The duration of the study was approximately 60 minutes.

6.3.4 Participants

We recruited participants through our university's participant sign-up platform, which is mainly used by students in the field of HCI and psychology. Students received study credits for their participation. The study was completed by 26 participants (women=15, men=9, non-binary=2), aged from 19 to 55 years (M=26.3, SD=7.68). All participants were right-handed and had normal or corrected-to-normal vision. The majority of participants ($n = 19$) rated their general VR experience with 5 or under (*from 1 = no experience to 10 = very experienced*), with 11 of them giving a rating of 2. Only 7 participants gave a rating of 6 or higher (overall Median = 3). Participants reported their frequency of using interactive VR applications as follows: 3 never used VR before, 12 only used it 1–2 times in total, 7 use VR about 1–2 times a year, 3 use it about 1–2 times per month, and one participant uses VR 1–2 times per week.

6.4 Results

Before performing the statistical analysis, we cleaned and pre-processed the collected data. We then tested the quantitative data for normal distribution using the Shapiro-Wilk test. According to the results, we chose the appropriate parametric or non-parametric statistical tests and post hoc tests. For plotting the significant differences, we used the following indicators for the significance levels: $* < .05, ** < .01, *** < .001$.

6.4.1 Fitts' Law Measurements

Data Cleaning and Processing

The collected amount of recorded data for each participant is listed as follows:

- 1 multi-directional selection task $\times$ 9 target selections = 9 target selections per trial
- 5 multi-directional selection tasks $\times$ 4 IDs $\times$ 4 Conditions = 80 trials
- 80 trials $\times$ 9 target selections = 720 data points

In total, we collected 720 $\times$ 26 participants = 18.720 data points, which gives us 4.680 data points for each *Controller-Ray Technique*. To remove invalid measurements, we filtered the data based on the recorded movement time MT for each selection, using two steps:

(1) We removed accidental double-clicks. For identifying double-clicks, we examined all error selections where the raycast was either within 2 $\times$ the distance of the width of the previous target or within its redirection zone. Additionally, we plotted and visually inspected all error selections. With both methods, we found the same two trials with a double click in the condition *RayCurve*. We also removed the next trial in the sequence, since the target got skipped by the double click. By excluding these 4 trials, we removed 0.09% of the data points for the *RayCurve* condition.

(2) We filtered out all trials for each participant, where $MT > Mean_{MT} + 3 \times SD$ for each condition. This accounts for potential tracking errors, unintentional interruptions, and deviations from the task (e.g. participant slowed down intentionally to examine the curvature of the ray). This removed 120 trials consisting of: 26 trials in *Baseline* (0.6%), 27 in *RayRotation* (0.6%), 34 in *RayCurve* (0.7%), and 33 in *RayCurveRotation* (0.7%).

Altogether, we excluded 124 out of 18.720 trials from our statistical analysis, which corresponds to 0.7% of the recorded data points.

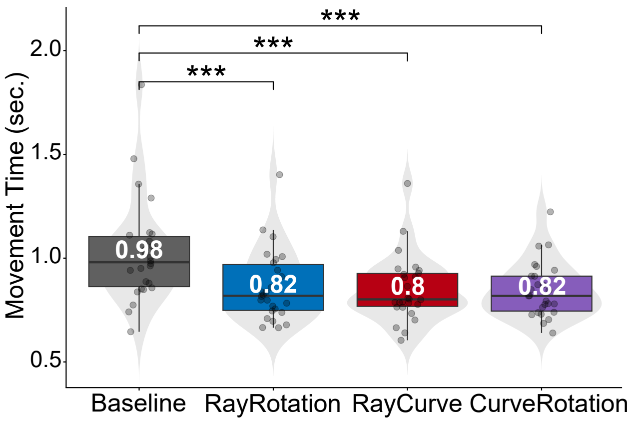
Fitts' Measurement Calculations

Using the cleaned data, we calculated the effective Fitts' Law measurements W_e, A_e, ID_e, and TP_e, as described in Section 2.7.2. For calculating A_e, we also used the selection coordinates of excluded trials (except for the two double-clicks) as starting position for the subsequent target in the sequence. Moreover, we projected the target size, selection coordinates, and target center positions on a 2D plane for the calculations as suggested for distant-pointing techniques by Teather and Stuerzlinger [TS11].

We performed initial two-factor statistical testing on the Fitts' Law measurements to confirm that there were no interaction effects between *Index of Difficulty* $\times$ *Controller-Ray Technique*. In the following, we will only focus on the factor *Controller-Ray Technique*.

Movement Time

The median MTs in seconds for each *Controller-Ray Technique* are: *Baseline* Mdn=0.98, *RayRotation* Mdn=0.82, *RayCurve* Mdn=0.8, and *RayCurveRotation* Mdn=0.82. As shown in Figure 6.6(a), the median values and interquartile ranges for each redirection technique are very similar, while the *Baseline* shows a higher MT. The descriptive statistics are listed in Table 6.6(b). Following a significant Friedman test ($\chi^2(3) = 37.25, p < .001$), we performed pairwise Wilcoxon signed rank post hoc tests with Holm-Bonferroni adjusted p-values. They revealed a significantly higher median MT for *Baseline* compared to all redirection techniques, as marked in Figure 6.6(a): *RayRotation* ($V = 333, p < .001$), *RayCurve* ($V = 347, p < .001$), and *RayCurveRotation* ($V = 344, p < .001$). We found no significant differences between the proposed redirection techniques.



(a)

Condition	Median	Mean	SD	[Q1-Q3]
Baseline	0.98	1.03	0.25	[0.86-1.10]
RayRotation	0.82	0.86	0.17	[0.75-0.97]
RayCurve	0.80	0.85	0.16	[0.77-0.92]
RayCurveRot.	0.82	0.85	0.13	[0.75-0.91]

(b)

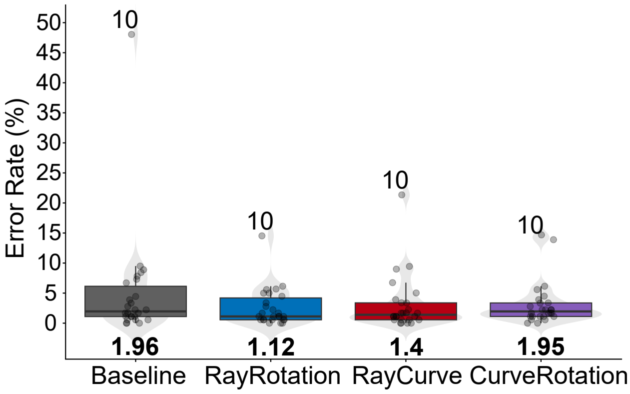
Figure 6.6: Movement time (sec.): (a) Boxplots and (b) descriptive statistics summary.

Error Rate

As shown in Figure 6.7(a), the median error rates are relatively low ($< 2\%$) for all *Controller-Ray Techniques*. However, the plotted individual data points show that especially one participant (id no. 10) made a lot more mistakes in the *Baseline* condition, with an error rate of nearly 50% for the task. This is also evident when looking at the maximum values of the descriptive statistics in Figure 6.7(b). We performed a Friedman test ($\chi^2(3) = 6.96, p = .07$) on the data, which did not indicate any statistically significant differences in the median error rates.

Effective Throughput

Mean TP_e in bits/s are: *Baseline* M=4.08 (SD=.61), *RayRotation* M=5.13 (SD=.82), *RayCurve* M=5.1 (SD=.81), and *RayCurveRotation* M=5.17 (SD= .8). Additional descriptive statistics are listed in Table 6.8(b). We performed a one-way repeated measures ANOVA ($F(3,75) = 28.82$, $p < .001$, $\eta^2_G = 0.269$), which indicated significant differences for the TP_e. The post hoc pair-wise t-tests with Holm-Bonferroni correction revealed a significant higher effective TP_e for all redirection techniques compared to the *Baseline*, as visualized in Figure 6.8(a): *RayRotation* ($t(25) = 7.59, p < .001$), *RayCurve* ($t(25) = 6.94, p < .001$), and *RayCurveRotation* ($t(25) = 8.12, p < .001$). We did not find any significant differences in TP_e between the redirection techniques.

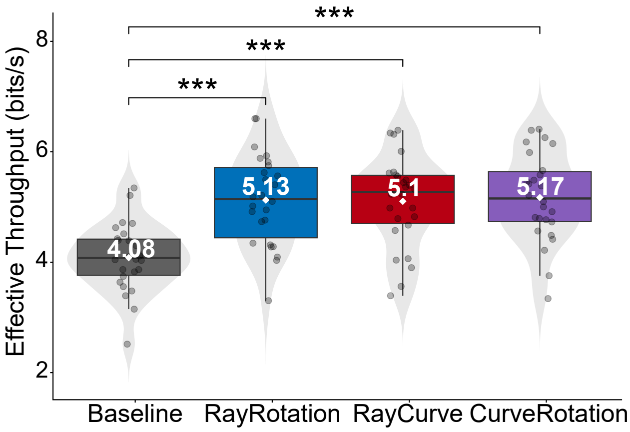


(a)

Condition	Median	Mean	SD	[Q1-Q3]
Baseline	1.96	4.88	9.31	[1.09-6.14]
RayRotation	1.12	2.58	3.14	[0.56-4.20]
RayCurve	1.40	3.04	4.55	[1.11-3.36]
RayCurveRot.	1.95	3.10	3.64	[0.56-3.38]

(b)

Figure 6.7: Error rate (%): (a) Boxplots and (b) descriptive statistics summary.



(a)

Condition	Median	Mean	SD	[Q1-Q3]
Baseline	4.08	4.08	0.61	[3.76-4.42]
RayRotation	5.14	5.13	0.82	[4.44-5.72]
RayCurve	5.27	5.10	0.81	[4.70-5.57]
RayCurveRot.	5.15	5.17	0.80	[4.74-5.64]

(b)

Figure 6.8: Effective throughput (bits/s): (a) Boxplots and (b) descriptive statistics summary.

6.4.2 NASA RAW TLX

The detailed descriptive statistics for all the TLX subscales and the Total Workload (TW) are listed in Table 6.1. Overall, the median task load for the *Baseline* is 27.5 out of a maximum score of 100. In comparison, all three redirection techniques had lower median total workload: *RayRotation* Mdn=20.4, *RayCurve* Mdn=22.5, and *RayCurveRotation* Mdn=25. The median scores for Temporal Demand were the highest of all scales ($33-35$) across the conditions, while the scores for *Physical Demand* were the lowest ($<= 15$).

We performed a Friedman test on each subscale and for the Total Workload (TW). The tests indicated significant differences for the subscales Performance ($\chi^2(3) = 12.98, p = .005$), Effort ($\chi^2(3) = 22.43, p < .001$), and Frustration ($\chi^2(3) = 8.84, p < .031$), as well as for the TW ($\chi^2(3) = 19.21, p < .001$). Figure 6.9 visualizes the median scores and error bars for the median spread for the TLX scales with significant differences. In the bar plots, the median score 42.5 for Effort in the *Baseline* condition stands out, which is the highest score across all scales. Post hoc pair-wise Wilcoxon signed-rank tests with Holm-Bonferroni correction revealed the following significant

Table 6.1: Descriptive statistics for NASA RAW TLX subscales and Total Workload.

	Baseline		**RayRotation**		**RayCurve**		**RayCurveRotation**	
Scale	M±SD	Mdn[Q1-Q3]	M±SD	Mdn[Q1-Q3]	M±SD	Mdn[Q1-Q3]	M±SD	Mdn[Q1-Q3]
Mental	33.1±22.4	25[16.2-45]	25.8±21.0	17.5[10-37.5]	28.7±24.2	20[15-41.2]	27.9±22.6	25[10-37.5]
Phys.	24.6±21.5	15[10-33.8]	16.4±14.6	15[5-20]	16.2±18.6	10[5-20]	21.2±20.9	15[5-28.8]
Temp.	35.6±24.3	27.5[20-55]	33.0±25.1	27.5[15-48.8]	35.6±26.4	30[16.2-53.8]	35.2±24.5	32.5[16.2-55]
Perf.	22.5±14.5	20[10-33.8]	15.4±12.4	10[6.25-23.8]	20.2±18.4	12.5[10-25]	18.9±15.0	15[10-20]
Effort	41.7±24.8	42.5[20-63.8]	24.2±22.8	15[10-28.8]	29.4±22.3	25[11.2-43.8]	31.9±24.3	20[11.2-58.8]
Frustr.	25.4±20.5	20[10-38.8]	16.4±18.8	7.5[1.25-23.8]	19.6±22.9	15[1.25-30]	19.0±18.1	17.5[5-25]
Total	30.5±16.2	27.5[20.8-39.8]	21.9±14.6	20.4[12.1-27.3]	24.9±17.8	22.5[15-29.2]	25.7±15.7	25.0[16-36.2]

differences: Performance was significantly better ($V = 24.5, p = .009$) (inverted scale in TLX, a higher score represents poor performance) and the Frustration significantly lower ($V = 33.5, p = .015$) for *RayRotation* compared to the *Baseline*. Effort was significantly lower for all redirection techniques compared to the *Baseline*: *RayRotation* ($V = 27, p = .003$), *RayCurve* ($V = 55, p = .009$), and *RayCurveRotation* ($V = 32.5, p = .006$). The same also applies to the TW compared to the *Baseline*: *RayRotation* ($V = 38, p = .002$), *RayCurve* ($V = 71, p = .034$), *RayCurveRotation* ($V = 64.5, p = .023$). A plot containing all TLX scales can be found in Appendix A3.1.

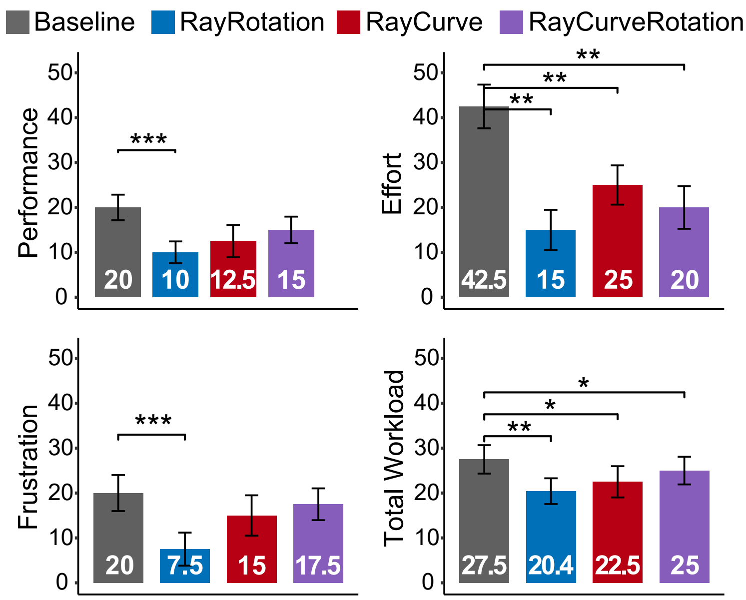


Figure 6.9: Barplots for the significant scales of the Nasa RAW TLX.

6.4.3 Sense of Agency

The redirection techniques had no negative effect on the perceived sense of agency, as their ratings were only marginally lower compared to the *Baseline*. Figure 6.10 shows that for every statement (S1–S4), no participant stated (strong) disagreement with the statements for any of the *Controller-Ray Techniques*. Overall, there is a very high agreement rate ($>= 88\%$), with the majority of participants (> 50%) (strongly) agreeing with each agency statement in all conditions. *RayCurveRotation* is the only *Controller-Ray Technique* that consistently received slight disagreement for each statement (<= 8%). We performed a non-parametric Friedman test on the ordinal Likert scale data, which did not indicate significant differences between the *Controller-Ray Techniques* for any of the four statements.

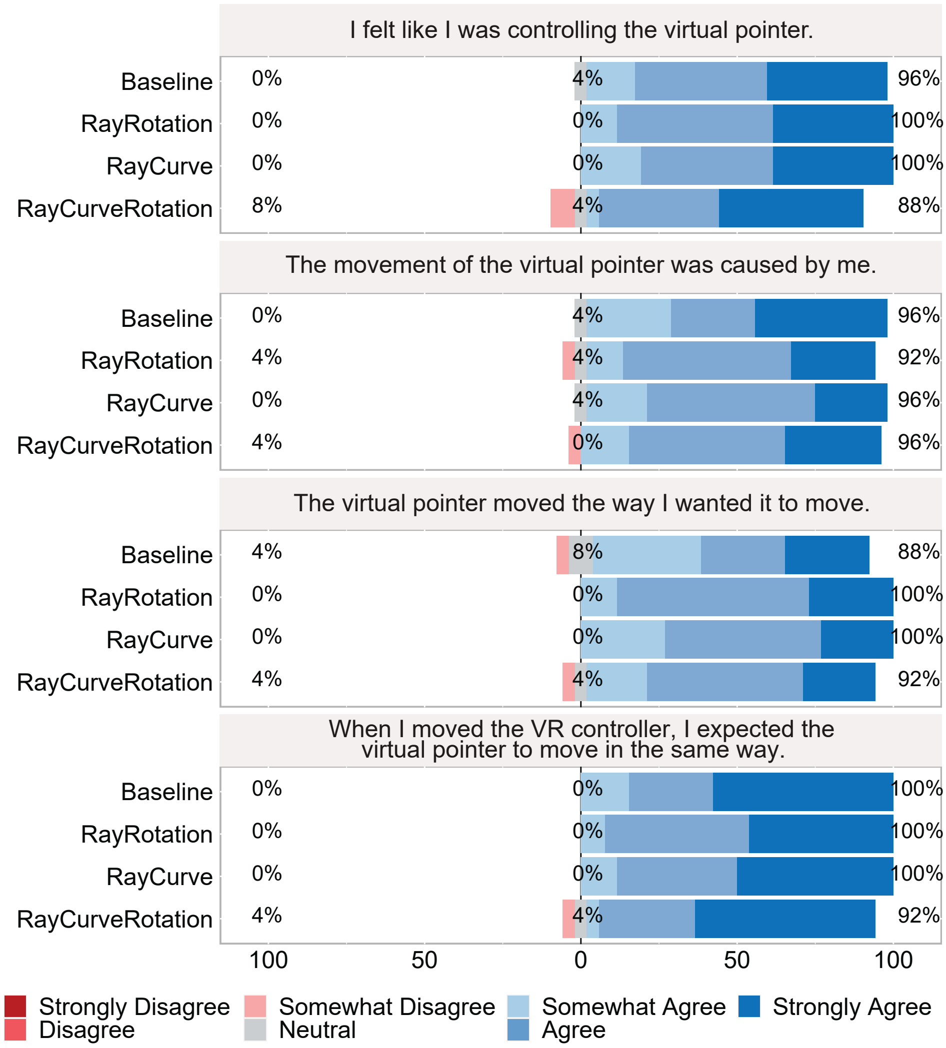


Figure 6.10: Agreement-disagreement Likert-scales for sense of agency statements.

6.4.4 Subjective Ranking of Preference

Participants ranked the selection in each block from 1 (most preferred) to 4 (least preferred). Three participants stated that they did not have a preference or did not perceive any difference between them, and thus ranked the interaction in two or more blocks the same. However, $n = 15$ (58%) participants ranked *Baseline* as their least preferred selection (see Figure 6.11). If we do not account for the tied rankings, *RayCurve* is the overall favored technique with the highest number of top rankings. *RayRotation* is placed in the middle with a comparable number of votes for ranks 1, 2, and 3. *RayCurveRotation* received lower rankings than the other two redirection techniques.

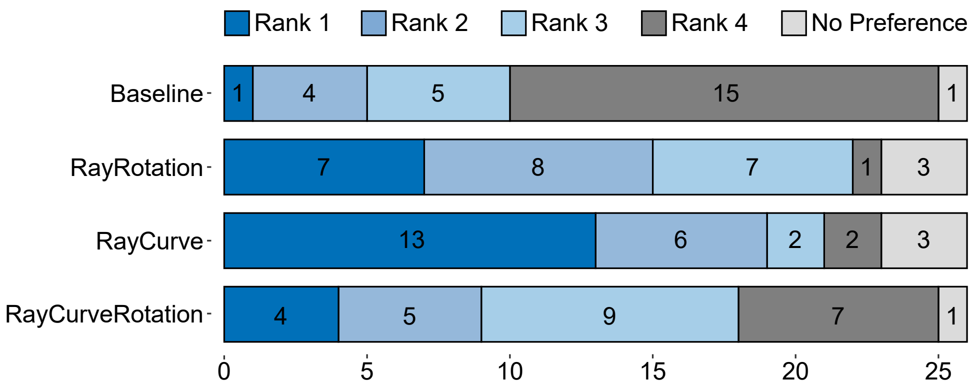


Figure 6.11: Subjective ranking for each *Controller-Ray Technique*.

6.4.5 Qualitative User Feedback

For the qualitative feedback, we followed an approach inspired by qualitative content analysis methods [Arm17; May00]. We iteratively defined three initial feedback categories based on the post-study questionnaire:

(i) Any differences that participants noticed in the selection
(ii) Feedback on subjective selection performance
(iii) Positive and negative aspects for each *Controller-Ray Technique*

Next, we familiarized ourselves with the data and generated initial codes for categorizing the qualitative data. During the coding process, we inductively defined new codes, re-coded statements, and formed new categories. Finally, we re-examined the coded data and determined the following final eight categories for our qualitative analysis:

Redirection Noticed

While 22 participants (85%) noticed differences between the selection interactions, 4 (15%) did not. From the former group, 12 (46% total) noticed the visual curving in *RayCurve*, while 7 (27% total) reported that they did not visually see a difference in *RayRotation*, and 4 (15% total) participants did not perceive a difference between the three redirection techniques. 7 (27%) out of the 26 participants stated that the perceived differences between the redirection techniques were generally very subtle. 7 (27%) participants reported that they were familiar with the concept of *aim assist* from video games.

Subjective Selection Performance

Performance-related aspects were frequently mentioned by participants when describing the selection in each of the four study blocks. Many participants reported that the selection was faster and/or less difficult with assistance. It was mentioned that the effort was higher and the accuracy was not as good when they had no assistance during the selection.

Perceived Pointer Behavior

The participants used various descriptions for the perceived redirected pointer behavior, such as sticky, flexible, curving, and slowing-down. Most frequently, the raycast was described as magnetic and jumping/snapping to the target center.

Redirection Strength

Overall, participants thought that the strength of the raycast redirection was adequate and not overly strong or intrusive. One participant reported that it was a good balance between maintaining control and getting assistance. Another mentioned that the redirection strength should depend on the target size. 3 participants felt that *RayCurve* used the strongest assistance, a few even reported that it was too strong. 6 participants stressed that the redirection should be minimal, subtle and not too strong. However, a few stated that they would not mind more assistance for the selection.

Visual Pointer Curvature

7 participants perceived the curving of the pointer in *RayCurve* negatively and described it as visually irritating, annoying, frustrating, or interfering. These participants reported that they only use interactive VR applications 1–2 times a year, or that they had used it only 1–2 times in total. In contrast, others had a positive impression of *RayCurve* and saw the curvature as a visual aid. Some participants reported that they were not expecting the pointer to curve. They were initially surprised or confused by this behavior, but quickly adapted to the curvature and did not find it irritating or distracting afterwards.

Sentiment and Preferences

The majority ($n = 23$) expressed positive sentiments for the redirection with statements such as: helpful, more pleasant, more fun, and smooth. Several participants reported that they especially appreciated the assistance for the selection of smaller targets. For the *Baseline* technique, it was often mentioned that it felt jittery/unstable. There was no clear consensus on the most preferred technique. Participants had different preferences and mixed feelings regarding each redirection technique. While multiple participants did not like *RayCurve*, it was favored by others. It was also stated that *RayRotation* felt more static and unnatural compared to the curved rays, while others preferred *RayRotation* over the curved visualizations. *RayCurveRotation* received the least attention in the user feedback. A few participants reported that the redirection felt too strong or imprecise, while others liked it. Several participants noticed when there was no redirection in a study block and remarked that they missed the previously provided assistance. A few stated that they began "to get used to it/dependent on it" (the assistance of the redirection) and thus put less effort into aiming. In contrast, others felt that the assistance in *RayCurve* and *RayCurveRotation* was not working as well as they hoped – the pointer's curvature did not make the selection easier for them, as they were unsure about the exact behavior of the curved raycast.

Agency and Self-Attribution

While a few participants felt that their control over the pointer was reduced by the redirection in *RayRotation* and *RayCurve*, others stated that they always felt in control of the interactions. Several participants attributed the perceived increase or decrease in performance between blocks entirely to their own selection performance, noting that they improved due to the extended practice with the selection task. Others were uncertain whether they should attribute the change in performance or the jitter of the pointer to themselves.

Expectations Pointing

Some participants expressed that they would like to manually adjust the redirection of the pointer. In addition, they would prefer to be informed about the assistive redirection beforehand.

6.5 Discussion

We implemented three different assistive interaction techniques based on raycast redirection for facilitating spatial selection in VR. We evaluated and compared their performance, perceived workload, and qualitative aspects with classic raycasting as a baseline in a user study with a Fitts' Law selection task. In the following, we will discuss the results of the study and their implications in the context of our research questions.

RQ1: Do assistive raycasting techniques improve selection performance and UX?

For the Fitts' Law measurements, the post hoc tests revealed significant differences between the *Baseline* and each of the three redirection techniques for median MT and mean effective throughput TP_e (all $p < .001$). These measurements were similar for all redirection techniques. Overall, MT was around 160ms faster and TP_e around 1 bit/s higher with redirected raycasting. The error rates in each condition were not significantly different. All median error rates were below 2%, which implies that the majority of participants performed the selection task properly and with attention to the task. This suggest that these redirection techniques, which were new to the participants, did not disturb the pointer interaction or lead to more selection errors than the *Baseline*. Interestingly, one participant (id no. 10) made noticeably more mistakes in each condition, but especially in the *Baseline* condition (c.f. Figure 6.7(a)), with an error rate of nearly 50%, which is around double the error rate compared to the redirection conditions. This participant stated that they are a frequent VR user (1-2x per month). Moreover, they started with the *Baseline* condition in the study. We suspect that this participant might have been too fast and confident with the selections due to their VR experience. The participant reported in the qualitative feedback that they had to correct their movements a lot in the *Baseline* condition. However, this case shows that the redirection techniques were able to reduce the selection error rate for this participant by at least 25% compared to no assistance. Since we found no interaction effects between *ID* and *Controller-Ray Technique*, this suggests that the redirection techniques are robust enough to significantly improve spatial selection for different target sizes and distances.

This is also supported by our findings for the Total Workload and the Effort subscale in the NASA TLX, which were significantly higher for the *Baseline* compared to the redirection techniques. Interestingly, we also found that Performance and Frustration were significantly higher for *Baseline* compared to *RayRotation*, both with $p < .001$. We will elaborate on possible explanations for this finding in the discussion of RQ3. Our redirection techniques do not seem to negatively affect the sense of agency. The ratings were high across all *Controller-Ray Techniques* ($>=$ 88% agreement) and participants felt in control of the pointer and the interactions.

In terms of UX, the majority of participants noticed and responded positively to the assistance of the redirected raycast selection. Many stated that the amount of redirection was adequate and that the selection was faster with the assistance. In comparison, the effort was higher, and it felt more difficult without the assistive redirection. This supports our results for the Fitts' Law and NASA TLX measurements.

RQ2: Are there significant differences in measured selection performance between the three proposed redirection techniques?

We did not find any significant quantitative differences between the three raycast redirection techniques based on our Fitts' Law selection task measurements. Our results indicate that the implemented assistive raycast redirection is robust in terms of selection speed, error rate, and effective throughput. The findings suggest that this type of proximity- and gain-based redirection might be unaffected by different visualizations of the redirected raycast, at least for our evaluation with the Fitts' Law selection task and within our tested parameters.

RQ3: Do UX and individual preferences differ for the three redirection techniques?

The results of the qualitative feedback revealed that individual perception and preferences differ for the three redirection techniques. There was no clear overall favorite, and opinions were mixed and often contrasted each other. On the one hand, *RayCuve* was the preferred technique for some participants, who liked the visible curvature. On the other hand, it was irritating for others, especially those with little VR experience. Although all techniques used the same redirection parameters, some participants felt that *RayCuve* applied the strongest redirection. This suggests that pointer visualizations which are more noticeable, such as a curved ray, might be perceived as stronger redirection by some individuals. The other two redirection techniques were not mentioned as frequently. Similar to *RayCuve*, *RayCurveRotation* received contrasting statements from participants. *RayRotation* was described as more static and unnatural by one participant. However, *RayRotation* has the lowest perceived workload in the four significant NASA TLX scales of our study (c.f. Figure 6.9). It is the only redirection technique with significantly higher Performance and lower Frustration than the *Baseline*. One possible explanation could be that the curved redirection techniques were irritating for some participants and unexpected for others, who stated that they required some time to get used to the technique. In addition, participants primarily noticed the curved redirections and may have just felt that their own performance was better in the *RayRotation* condition (self-attribution). In summary, there are many subjective factors that could potentially influence the UX for redirected raycast selection.

Recommendations

Drawing from our study results, we provide the following recommendations for using assistive raycast redirection in spatial interaction techniques. Since different raycast visualizations did not result in significant performance differences, the same underlying redirection calculation could be applied to additional raycast visualizations. Likewise, various other combinations of visualizations and redirection parameters could be used for achieving specific effects, such as perceived redirection strength. For practical use, the appropriate visualization should be chosen according to the use case and based on the VR experience of the target user group. For example, novice VR users might not be familiar with curved raycast visualizations, which might cause confusion or distraction during the interaction. Another important aspect is the transparency in the communication of the used interaction technique. Users should be clearly informed when redirected raycasting is applied, and an option to toggle the assistive selection should be provided. A subtle redirection seems to be generally accepted by users. However, if feasible, settings for changing the raycast visualization and for adjusting the redirection parameters should be offered, since the individual preferences appear to differ widely between users.

6.6 Limitations and Future Work

Since this is our initial evaluation of the proposed assistive raycast redirection techniques, we chose an abstract Fitts' Law selection task as a standardized measure of selection performance to reduce possible confounding variables. Therefore, the validity of our results and the practical application of the redirected raycast selection needs to be further investigated in follow-up studies with different, realistic use cases and interactive immersive environments.

To avoid selection induced pointing errors during our study, we decoupled the pointing and selection confirmation of the selection task. However, in real-world applications, the confirmation of selections is typically done by pressing the trigger button on the same VR controller, which might introduce pointing errors. The target snapping of our assistive redirection should theoretically compensate for this. This assumption should be investigated in the future.

As we only used a pre-defined size of $2\times$ the target size for the *redirection zone*, the transferability to other potential use cases might be limited. Future work should evaluate different *redirection zone* sizes. Additionally, other combinations and variations of redirection parameters could be investigated. Even solutions for adaptive parameter setups could be considered. Since participants felt a stronger redirection with *RayCurve*, an interesting aspect to explore in future research could be the modulation of perceived redirection properties, such as speed or strength, with different visualizations. For our initial study, we did not compare assistive raycasting to other existing raycasting-based selection techniques, such as those based on instant snapping and curving of the raycast. A comparative study on performance and qualitative aspects with other established spatial interaction techniques should be conducted in the future.

Furthermore, we only tested our redirection techniques with controller-based raycasting for this initial investigation. Since modern HMDs are increasingly equipped with multi-modal built-in tracking, it would be interesting to transfer assistive redirection to hand- and eye-tracking-based raycast interaction techniques in future studies.

Lastly, our redirection-based raycast techniques have limitations in cluttered 3D environments.

Since the redirection does not account for overlapping targets or occlusions, this would lead to conflicts if multiple objects or redirection zones overlap. There is no inherent distinction as to which object the user wants to select. Even if the raycast is slightly closer to the center position of one 3D object, the actual intended selection target could be the other, overlapping object. This limitation should be addressed in future research.

6.7 Conclusion

We developed three raycast interaction techniques based on redirection that subtly assist with spatial selections, but stay close to to the original pointer metaphor: *RayRotation*, *RayCurve*, and *RayCurveRotation*. To compare the redirected raycasting techniques with each other and with classic raycasting as a baseline, we conducted an empirical user study with a standardized Fitts' Law selection task. For investigating each technique, we formulated three research questions concerning performance, UX, and subjective aspects of the spatial interaction techniques. Our findings suggest that selection with redirected raycasting is faster, leads to higher effective throughput, and decreases perceived workload for users. The sense of agency during the interaction was not negatively affected by the applied redirections, and most participants preferred selection with assistive redirection. The results did not reveal any significant differences in measured performance and perceived workload between the three different raycast redirection techniques. Our findings suggest that different visualizations, such as a straight versus a subtly curved ray, may not influence the selection performance. However, participants' opinions for each redirection technique and the curved raycast visualization were mixed and sometimes even contrasted each other. This indicates that individual preferences and prior experience with VR interactions play a role in user acceptance of using redirection for raycasting-based interaction techniques.

To summarize, for addressing RQ-*III*, this chapter proposes the novel concept of raycast redirection for spatial interaction and provides an algorithm as well as three different types of raycast visualizations for implementing the proposed redirection. This work presents an initial evaluation with insights and recommendations for the design and implementation of redirected raycasting techniques. With our implementation and user study, we have demonstrated the potential of incorporating assistive redirection for facilitating spatial interaction, which serves as a starting point for further research on raycast redirection. Our results also highlight that individual, subjective factors and user preferences are equally as important as the selection performance when developing new spatial interactions. This emphasizes that a user-centered approach and UX considerations are essential for the design and practical application of redirected raycasting as an assistive spatial interaction technique.

CHAPTER 7

HAND-RAY AND GAZE-BASED RAYCAST REDIRECTION

This chapter is based on published work [Gab+24] and contains rephrased, revised and extended paragraphs. In addition, new information, tables, and figures have been included.

7.1 Motivation

Following our promising findings for controller-based assistive raycast redirection, we continued with the investigation of raycast redirection with hand- and eye-tracking, led by our RQ-IV:

> **RQ-IV**
>
> Can we integrate hand-tracking and eye-tracking with raycast redirection to improve spatial selection?

Our results presented in the previous chapter suggest that selection with redirected raycasting has significantly better performance and decreased perceived workload compared to classic raycasting. Moreover, the sense of agency was not negatively affected by the redirection, and the assistance was generally well received by the participants. Encouraged by these findings, we further explored the use of hand- and eye-tracking input in combination with assistive raycast redirection.

With the increasing availability of advanced built-in hand- and eye-tracking capabilities in HMDs, both input modalities have become more widespread for spatial interactions. For example, hand-based raycasting is one of the standard interaction techniques for spatial interaction on the Microsoft HoloLens 2 and Meta Quest HMD series. And *Gaze&Pinch* [Pfe+17], which uses eye-tracking for targeting and hand gestures for confirming the selection, is a standard interaction technique on the Microsoft HoloLens 2 and the Apple Vision Pro HMDs. By using these interaction techniques, there is no need for additional controller tracking and handling, which might be impractical in certain situations, such as for usage in public spaces. Moreover, using unfamiliar XR controllers might be more complicated and difficult for inexperienced users (see Section 4.6.2). While the advanced hand- and eye-tracking capabilities in modern HMDs generally enable fast and precise selections, there is still room for improvement. Since Hand-Ray interaction uses the pinch gesture (index finger and thumb) for selection confirmation, it is more prone to selection-induced errors (Heisenberg Effect of Spatial Interaction [Bow+01]). Natural hand tremor and tracking jitter could also negatively impact the selection [BS19b; Bow+01; LaV+17]. In addition, since Hand-Ray selection is a raycasting-based interaction technique, it also suffers from a decrease in performance for the selection of small and distant targets (see Section 2.3).

Based on our previous findings for Controller-Ray redirection, we assumed that redirection could potentially counteract these limitations and increase selection performance for Hand-Ray interactions. Following the same approach as for Controller-Ray redirection in Chapter 6, our goal was to facilitate Hand-Ray selection while remaining close to the classic pointer metaphor.

Additionally, we were also interested in further extending and improving Hand-Ray redirection with the added support of eye-tracking input. Therefore, our second goal was to develop a novel redirection targeting approach that uses the tracked gaze position without altering the core behavior of raycast redirection. By leveraging gaze to dynamically set the redirection center rz of the Hand-Ray, the assistive redirection becomes independent of the target object position (cf. redirection to target center in Section 6.2). This would allow for a more general and scalable practical application of raycast redirection.

For our empirical evaluation, we formulated the following two research questions for testing both techniques – Hand-Ray redirection and its gaze-assisted extension – and for comparing them with classic Hand-Ray as a baseline:

- **RQ1** Does assistive Hand-Ray redirection improve object selection performance and UX compared to classic Hand-Ray?
- **RQ2** Are there differences in selection performance and UX between the two Hand-Ray redirection techniques?

Our work in this chapter contributes comprehensive results and insights into assistive Hand-Ray redirection:

- **C1** *RayToTarget* as an adaptation of controller-based raycast redirection and *RayToGaze* as a novel redirection concept for gaze-assisted Hand-Ray interaction.
- **C2** A Fitts' Law evaluation comparing the proposed techniques to classic Hand-Ray with no redirection.
- **C3** Insights gained from the study, demonstrating differences in UX and performance advantages over no redirection.

7.2 Raycast Redirection Implementation

Building on our previous work on assistive Controller-Ray redirection, we used the same redirection algorithm and parameters as described in Section 6.2 and applied them to Hand-Rays. Instead of the controller origin, we attached the redirected Hand-Ray to the respective left and right hand-tracking anchors of the VR-Rig. We disabled the visualization of the actual tracked Hand-Ray and used the tracked hand pointer pose for our redirection calculation. The redirected pose is then applied to a second, visible, redirected Hand-Ray. The redirection algorithm is implemented with a modified quaternion linear interpolation (lerp) function that uses a proximity-based rotational gain as the interpolation step parameter, according to Equation (7.1):

$$\mathbb{H}_r = \text{Lerp}(\mathbb{H}_a, \mathbb{H}_{rc}, step_{redir}) \tag{7.1}$$

As illustrated in Figure 7.1, $\mathbb{H}_r$ is the output rotation quaternion that is applied to the redirected Hand-Ray. $\mathbb{H}_a$ is the rotation of the tracked Hand-Ray anchor in 3D space, and $\mathbb{H}_{rc}$ is the required amount of rotation from $\mathbb{H}_a$ to the redirection center (position that the Hand-Ray is pulled towards to). The proximity-based gain is calculated with a float inverse linear interpolation: $step_{redir} = \text{InverseLerp}(r_{rz}, 0, d_{rc}) \cdot g$. The equation uses the radius of the redirection zone r_{rz} (set to $2\times$ of the Fitts' target width), the current distance to the redirection center d_{rc}, and a fixed gain factor g of 1.5. As soon as the hidden, tracked Hand-Ray enters the redirection zone rz, the redirected Hand-Ray is gradually pulled towards the redirection center rc. At the end, the redirection is eased-out and the visible raycast subtly snaps to the center rc position. This results in a magnetic behavior of the Hand-Ray. Users can freely move the Hand-Ray in and out of rz at any time during the interaction. The amount of redirection increases and decreases based on d_{rc}. As the redirection is proximity-based, the redirected Hand-Ray gradually returns to a 1:1 mapping of the tracked Hand-Ray as it leaves rz. This implementation seamlessly blends the spatial interaction between redirected and tracked Hand-Ray movement.

Using this calculation, we implemented two Hand-Ray redirection techniques, *RayToTarget* and *RayToGaze*. Both techniques differ in the setup of their redirection center rc and redirection zone rz, as listed in Table 7.1, and are described in the following paragraphs.

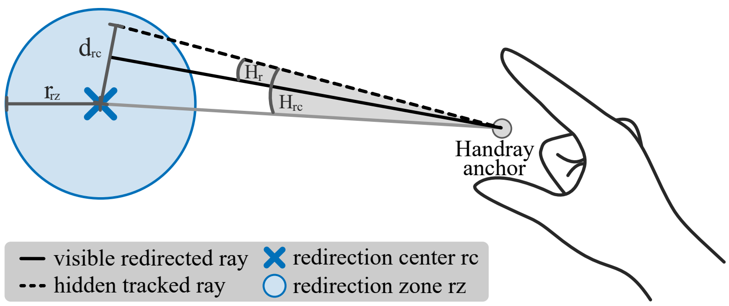


Figure 7.1: Illustration of the Hand-Ray redirection calculation.

7.2.1 RayToTarget

RayToTarget is a Hand-Ray redirection technique that corresponds to our implementation for Controller-Ray redirection (cf. Section 6.2) and follows the same rc and rz setup. As depicted in Figure 7.2, a predefined rz surrounds the target object. Depending on the distance to rc, the

Hand-Ray is pulled towards the object's center. As long as the hidden, tracked Hand-Ray stays inside *rz*, the current redirection gain is updated and applied to the visible redirected Hand-Ray. This technique requires prior knowledge of the interactable objects in the immersive environment for *rz* setup. In addition, the redirection is always applied towards the object's center position.

7.2.2 RayToGaze

RayToGaze implements a novel redirection targeting approach by setting the redirection center *rc* based on the user's gaze position. As pictured in Figure 7.2, *rc* is set to the position in 3D space where the gaze intersects with the target object. The redirection zone *rz* is then placed around *rc*. If the gaze intersects with a target and the Hand-Ray is inside the dynamically set *rz*, the Hand-Ray gets redirected towards the gaze *rc* position. Following the example of other interaction techniques, where the eyes assist the user's manual input [Che+23; Kyt+18; Wag+23], we implemented *RayToGaze* to further extend the Hand-Ray redirection technique with added gaze support. This dynamic, gaze-based redirection takes inspiration and partially derives from concepts of existing techniques such as *Magic Pointing* [ZMI99] and *Gaze&Pinch* [Pfe+17]. However, in contrast to our subtle, proximity-based redirection, these techniques instantly warp the pointer to the gaze position. While *RayToTarget* requires prior knowledge of a predefined redirection zone around potential target objects, the dynamic gaze-based redirection of *RayToGaze* is independent of target object positions. This eliminates the need for additional setup of the immersive environment and could offer more flexibility for various practical use cases.

Table 7.1: Comparison of the two Hand-Ray redirection techniques.

	RayToTarget	**RayToGaze**
Redirection Center	target object center	gaze hit position
Redirection Zone	around target object	around gaze position
Modalities	hand	hand and gaze

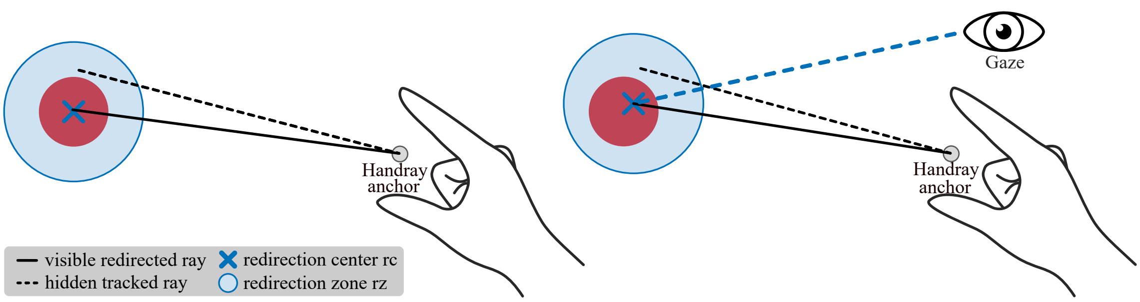


Figure 7.2: Illustrated Hand-Ray redirection setup: (left) RayToTarget and (right) RayToGaze.

7.3 User Study

To evaluate our two proposed assistive Hand-Ray redirection techniques, *RayToTarget* and *RayToGaze*, we conducted a comprehensive user study and compared both techniques to classic Hand-Ray as a baseline condition. We chose classic Hand-Ray as a baseline since we designed this assistive redirection for facilitating raycasting-based pointing with the intention of using the redirection to compensate for over- and undershooting and jitter during spatial selection tasks. We expected that users would perceive the assistance provided by Hand-Ray redirection in a similar way to *aim assist*. We were specifically interested in investigating the integration of redirection for assistive Hand-Ray interaction that would allow users to maintain control over their pointing and targeting interactions. Therefore, we did not include instant, direct raycast-to-target snapping (see Section 2.4.1) or existing gaze-based interaction techniques (see Section 2.5.4) as additional conditions in our user study.

For investigating RQ1, we formulated six hypotheses that were partially derived from our previous work on controller-based raycast redirection (cf. Chapter 6). Based on our previous findings for Controller-Rays, we hypothesized that the integration of redirection for Hand-Ray interaction would result in an overall higher measured selection performance of our two Hand-Ray redirection techniques compared to our *Baseline*, the classic Hand-ray without redirection:

H1a Hand-Ray redirection has faster movement times than *Baseline*.
H1b Hand-Ray redirection has lower error rates than *Baseline*.
H1c Hand-Ray redirection has higher effective throughput than *Baseline*.

In addition, we tested the following hypotheses on subjective aspects of Hand-Ray interaction:

H2 The perceived workload is higher in *Baseline* compared to Hand-Ray redirection: Since we expect Hand-Ray redirection to facilitate selection, this should reduce the workload for users.

H3 The sense of agency is not negatively affected by Hand-Ray redirection: Similar to Controller-Ray redirection, the subtle redirection with seamless transition between redirected and tracked movement should not affect users negatively.

H4 Users prefer Hand-Ray redirection over classic Hand-Ray: Due to improved selection performance and UX, we expect users to prefer Hand-Ray with redirection.

To our best knowledge, gaze-assisted raycast redirection has not been previously studied. Thus, for answering RQ2, we formulated four hypotheses for comparing the two Hand-Ray redirection techniques, *RayToTarget* and *RayToGaze*, in terms of selection performance:

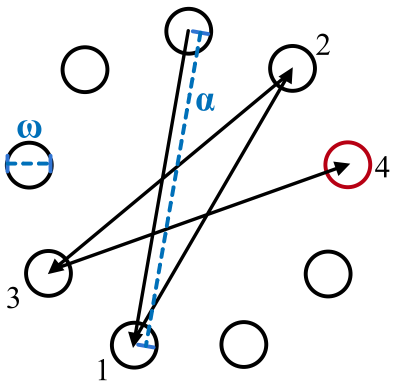
H5a Movement times differ between *RayToTarget* and *RayToGaze*.
H5b Error rates differ between *RayToTarget* and *RayToGaze*.
H5c Effective throughput differs between *RayToTarget* and *RayToGaze*.

Moreover, drawing from our previous results for Controller-Ray redirection, we hypothesized that users would have individual preferences regarding the two proposed Hand-Ray redirection techniques:

H6 Users either prefer *RayToTarget* or *RayToGaze*.

7.3.1 Fitts' Law Selection Task

We used the same multi-directional Fitts' Law selection task with 9 target spheres as in our Controller-Ray redirection study (see Section 6.3.1). The targets light up in red color in a fixed sequence, as illustrated in Figure 7.3(a). The active target (red) has to be selected with the Hand-Ray of the dominant hand by pointing at the sphere. The selection is then confirmed by performing a pinch gesture with the thumb and index finger of the same hand. Each target lights up in blue color as soon as it is intersected by the Hand-Ray. We recorded a successful selection if the Hand-Ray intersects with the active target on pinch confirmation, and an error otherwise. Participants did not receive any indication of whether their selection was successful. After performing a selection, the next target on the opposite side of the circular target setup is activated until the sequence is completed. To test the robustness of Hand-Ray redirection, especially with the addition of gaze for *RayToGaze*, we used two different depths for the targets: 1.3m and 4m. These distances represent close and distant targets in VEs, such as VR menus or 3D objects. Both values are within the recommended stereo content viewing zone (1.25 – 5m) [Pau+], but outside the direct interaction range, so that users are required to use far interactions. Since we included two different depths, unlike in our previous evaluation of Controller-Ray redirection, we used ID_{DP} for distal pointing for setting our target parameters, as described in Section 2.7.4. We chose six ID_{DP} (calculated with k=3), based on three angular target widths ω and two angular amplitudes α, as listed in Table 7.3(b). The ω values are within the range of parameters reported in related work [Ber+21a; Gab+23]. The α and ω values were also selected to create relatively equidistant ID_{DP} ranging from 1.9 to 4.4. This results in twelve unique configurations of 2 depths × 6 ID_{DP} for the selection task.



(a)

ID$_{DP}$	ω	α	**depth**
1.9	2.4°	22.5°	1.3m/4m
2.3	2.4°	26°	1.3m/4m
2.7	2.2°	22.5°	1.3m/4m
3.2	2.2°	26°	1.3m/4m
3.7	2°	22.5°	1.3m/4m
4.4	2°	26°	1.3m/4m

(b)

Figure 7.3: Fitts' Law selection task design and parameters: (a) Illustration of the selection task and (b) ID_{DP} listed with parameters.

7.3.2 Questionnaires

We used a pre-study questionnaire on demographics, VR usage frequency, and experience with hand-tracking in VR. Within-study questionnaires were presented after each study block and included the NASA RAW TLX [Har06b] and sense of agency statements. The post-study questionnaire collected subjective rankings and preferences, as well as qualitative feedback. For measuring the perceived sense of agency, we used four statements adapted from related work, including our previous study on Controller-Ray redirection [Arg+16; Gab+23; Ma+21]. The statements are rated based on a 7-point Likert scale ranging from *Strongly Disagree – Strongly Agree*:

S1 The movement of the virtual pointer was caused by me.

S2 I felt like I was controlling the virtual pointer.

S3 The virtual pointer moved the way I wanted it to move.

S4 When I moved my hand, I expected the virtual pointer to move in the same way.

7.3.3 Study Setup and Procedure

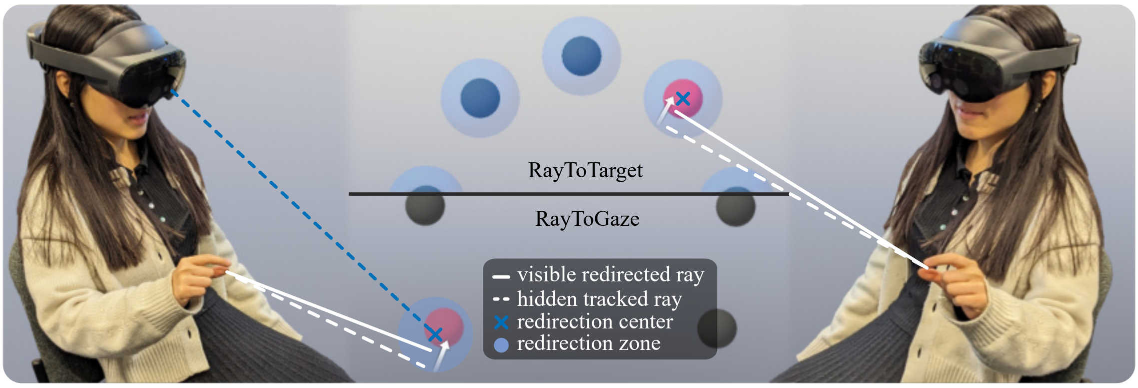


Figure 7.4: Participant during the user study, with the Fitts' Law task showing both Hand-Ray redirection techniques: (left) *RayToGaze* and (right) *RayToTarget*.

The study environment was implemented with the Unity engine (version 2021.3 LTS) and the Meta XR All-in-One SDK (version 60.0). In the study, we used a Meta Quest Pro HMD with built-in hand- and eye-tracking. The study was conducted in a stationary, seated VR-setup, as pictured in Figure 7.4. We applied an exponential moving average (EMA) filter to smooth out the rapid saccades in the eye-tracking input and used Unity's SphereCasting function to increase the accuracy of the gaze input.

Our within-participant study design had three *Hand-Ray Technique* conditions: *RayToTarget*, *RayToGaze*, and classic Hand-Ray as the *Baseline*. Before starting the study, participants were informed about the study procedure and signed a consent form. Next, they answered the pre-study questionnaire on a desktop PC and received instructions for using the HMD and the hand-tracking gestures. Inside the VR study environment, participants followed a short eye-tracking and seating height calibration sequence. They also had to complete a short, initial training block with two Fitts' selection tasks (ID 1.9 and ID 2.7). Participants were instructed to select the targets as accurately and quickly as possible. We made sure that they understood the task correctly and had no problems with performing the pinch gestures. Afterwards, participants started with the actual tasks for the study. They could take a break after each fully completed multi-directional selection task sequence to avoid fatigue of the arm and hand. In total, participants completed three subsequent study blocks. In each block, one of the three conditions was presented. All 12 target configurations (see Table 7.3(b)) were used twice, amounting to 6 ID_{DP} $\times$ 2 depths $\times$ 2 repetitions = 24 selection tasks per block. We counterbalanced the condition order with a Latin-square design, and randomized the order of the 12 task configurations within a block with the Fisher-Yates shuffle algorithm [Ebe16]. After completing each block, participants answered a within-study questionnaire with a VR controller on a UI panel. As the Meta Quest Pro does not provide an option to re-calibrate the eye-tracking while an application is running, participants could not take off the HMD during the study. However, they were allowed to take a break to close and rest their eyes after each study block. Finally, participants completed the post-study questionnaire on a PC. After subjectively ranking and explaining their preference for each condition, we revealed and explained the two different Hand-Ray redirections with images and videos. Participants could then provide additional feedback. The study took about 60 minutes in total.

7.3.4 Participants

Participants were recruited through our university's study platform, which is mostly used by HCI and psychology students. The students received study credits for their participation. An a priori power analysis indicated a minimum sample size of $n = 27$ to detect medium effects ($\eta_p^2 = .06$) with 80% power for an ANOVA (repeated measures, within factors) with an alpha level of .05. In total, we collected data from 32 participants. Two data sets were excluded due to data recording and hand-tracking issues. The remaining 30 participants (women=23 and men=7) were aged between 20-67 years (M=26.16, SD=8.44). 28 participants were right-handed, 2 were left-handed, and all had normal or corrected-to-normal vision. 22 out of 30 participants have used hand tracking in VR before, and 8 participants have never used it. The participants' VR experience is summarized as follows:

- $n = 12$ no to very little experience ($n = 7$: never, $n = 5$: less than 1x/year)
- $n = 6$ moderate experience ($n = 2$: 1x/year, $n = 2$: 1x/semester, $n = 4$: 1x/quarter)
- $n = 10$ regular VR users ($n = 7$: 1x/month, $n = 3$: >1x/week)

7.4 Results

For testing H1[a-c] and H5[a-c], we first looked at the Fitts' Law measurements in our collected quantitative data. Outliers were removed, as described in the following Section 7.4.1. We tested the data for normality of residuals with the Shapiro-Wilk test and used appropriate data transformations in case the assumption was violated. Sphericity was confirmed using Mauchly's test. If we could not assume normal distribution of the data, we used appropriate non-parametric tests and post hoc test. For each of the three dependent variables movement time, error rate, and effective throughput, we performed a two-way ANOVA with the factors *Hand-Ray Technique* and *Depth*. Following the Fitts' Law data analysis, we report our results for the NASA RAW TLX, sense of agency statements, and the subjective ranking and qualitative feedback. For plotting the significant differences, we used the following indicators for the significance levels: $* < .05, ** < .01, *** < .001$.

7.4.1 Fitts' Law Measurements

Data Cleaning

Before starting the statistical analysis, we cleaned and pre-processed the data to remove invalid measurements. The amount of collected data for each participant is listed below:

- 1 multi-directional selection task $\times$ 9 target selections = 9 target selections per trial
- 2 multi-directional selection tasks $\times$ 2 depths $\times$ 6 $\mathrm{ID}_{\mathrm{DP}}$ $\times$ 3 Conditions = 72 trials
- 72 trials $\times$ 9 target selections = 648 data points

This provides us with 648 $\times$ 30 participants = 19.440 total recorded selections, with 6.480 data points for each condition (*Hand-Ray Technique*). From the collected data, we removed invalid selections in two steps:

(1) All error selections where the distance between the Hand-Ray and the previous target in the Fitts' sequence is below the threshold of 2 $\times$ the target width are classified as double pinches. This removed 45 selections: 15 for *Baseline*, 17 for *RayToTarget*, and 13 for *RayToGaze*. If a double pinch occurred in the middle of a Fitts' sequence, the next target in the sequence would be skipped. In this case, we also had to remove the next recorded pinch, since it would not start at the correct position in the multi-directional setup. If a double pinch was performed on the last target in the sequence, we only removed the double pinch. This resulted in the removal of 37 selections: 13 for *Baseline*, 15 for *RayToTarget*, and 9 for *RayToGaze*.

(2) In the second data cleaning step, we identified and removed unusually high MTs as outliers. We did this to exclude selections that were disrupted by tracking issues and by users that intentionally slowed down or paused during the selection. This removed 271 outliers where movement time $MT > Mean_{MT} + 3 \times SD$: 92 for *Baseline*, 92 for *RayToTarget*, and 87 for *RayToGaze*.

In total, we removed 353 selections with both data cleaning steps, which corresponds to 1.8% of the total recorded selection data. For each condition, this amounts to: 120 selections (1.8%) for *Baseline*, 124 selections (1.9%) for *RayToTarget*, and 109 selections (1.6%) for *RayToGaze*.

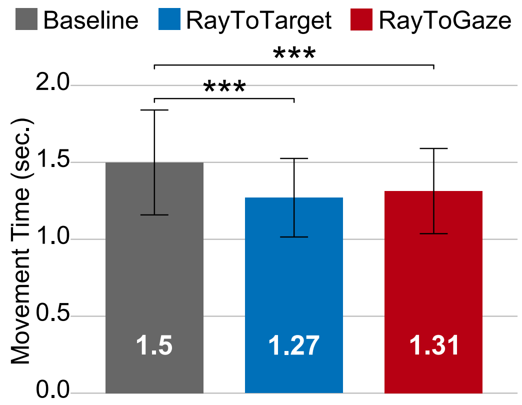
Fitts' Measurement Calculations

Using the cleaned data, we calculated the effective Fitts' Law parameters W_e, A_e, ID_e, and TP_e, as described in Section 2.7.2. For the A_e calculation, we followed the same procedure as in our previous study (see Section 6.4.1). The pinch selection coordinates of the removed outliers from step (2) were used as movement starting positions for the next target in the sequence. Following our previous study on Controller-Ray redirection, we also projected the target size, selection coordinates, and target center coordinates onto a 2D plane for the calculations, as suggested by [TS11].

Movement Time

Table 7.5(b) shows that the Hand-Ray redirection conditions have mean MTs that are around 0.2 seconds lower than the *Baseline*. We applied a Box-Cox transformation with $\lambda = -1$ [BC64] and confirmed the normality assumption on the residuals. A two-way ANOVA revealed significant main effects for both *Hand-Ray Technique* ($F(2,58) = 28.867, p < .001, \eta_p^2 = .499$) and *Depth* ($F(1,29) = 43.112, p < .001, \eta_p^2 = .598$), but no significant interaction effect. For *Hand-Ray Technique*, a post hoc test with Bonferroni-adjusted p-values revealed a significantly higher MT

for *Baseline* compared to both redirection techniques, *RayToTarget* ($p < .001$) and *RayToGaze* ($p < .001$). The significant findings are also visualized in the barplots in Figure 7.5(a). No significant differences between the two redirection techniques were found ($p = .424$).



(a)

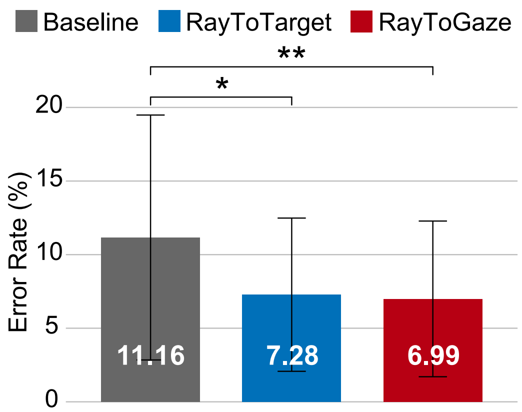
Technique	N	Mean	SD
Baseline	30	1.50	0.34
RayToTarget	30	1.27	0.26
RayToGaze	30	1.31	0.28

(b)

Figure 7.5: Movement time (sec.): (a) Barplots and (b) descriptive statistics summary.

Error Rate

The *Baseline* condition has visibly higher mean error rates with higher standard deviation, as shown in Figure 7.6(a)). This is also in evident in Table 7.6(b). The measured error rates were sqrt-transformed before the normality assumption was confirmed with a Shapiro-Wilk test on the residuals. A two-way ANOVA revealed a significant main effect of *Hand-Ray Technique* on error rate ($F(2,58) = 9.328, p < .001, \eta_p^2 = .243$). There was no significant main effect of *Depth* or interaction effect between both factors. For *Hand-Ray Technique*, a post hoc analysis with Bonferroni correction showed that the error rates for *Baseline* was significantly higher than for both redirection techniques, *RayToTarget* ($p = .012$) and *RayToGaze* ($p = .001$), as illustrated in Figure 7.6(a). However, the error rates of the two redirection techniques were not significantly different from each other ($p = 1.000$).



(a)

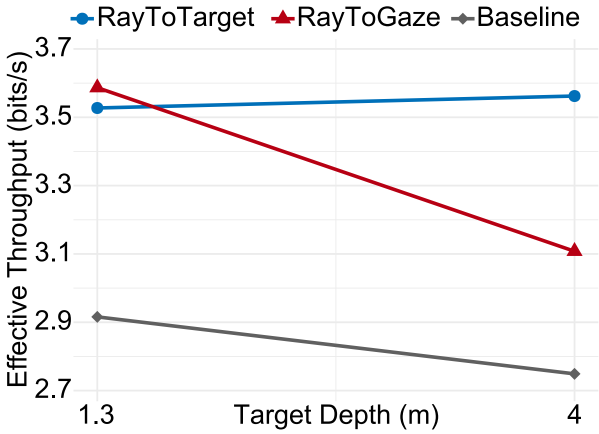
Technique	N	Mean	SD
Baseline	30	11.2	8.32
RayToTarget	30	7.28	5.21
RayToGaze	30	6.99	5.29

(b)

Figure 7.6: Error rate (%): (a) Barplots and (b) descriptive statistics summary.

Effective Throughput

The residuals of the TP_e were normally distributed, as confirmed with a Shapiro-Wilk test. As listed in Table 7.7(b), the mean TP_e in the *Baseline* condition is lower compared to both conditions with Hand-Ray redirection. However, TP_e is lower for *RayToGaze* at a depth of 4m compared to 1.3m. For a better visualization of the TP_e for each condition at both depths, we plotted Figure 7.7(a), which also clearly illustrates the decrease in TP_e with increasing depth for *RayToGaze*. This is also visible in Figure 7.8(a).

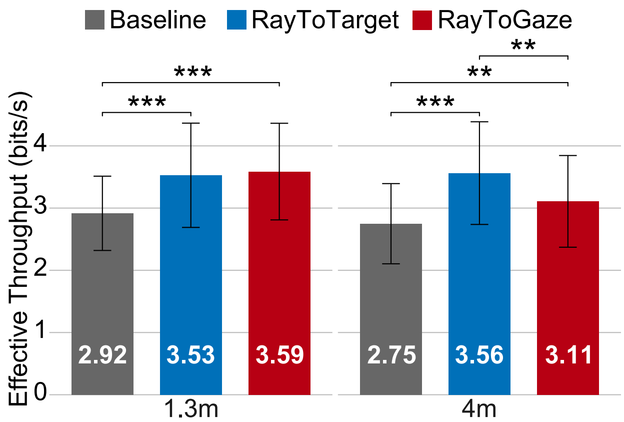


(a)

Technique	N	Depth	Mean	SD
Baseline	30	1.3	2.92	0.60
RayToTarget	30	1.3	3.53	0.84
RayToGaze	30	1.3	3.59	0.78
Baseline	30	4	2.75	0.64
RayToTarget	30	4	3.56	0.83
RayToGaze	30	4	3.11	0.74

(b)

Figure 7.7: Effective throughput and depth: (a) Line graph TP_e for depth 1.3m and 4m and (b) descriptive statistics summary.



(a)

Depth	Simple Comparison	p-value
1.3 m	RayToTarget–Baseline	p<.001
1.3 m	RayToGaze–Baseline	p<.001
4 m	RayToTarget–Baseline	p<.001
4 m	RayToGaze–Baseline	p=.003
4 m	RayToTarget–RayToGaze	p=.005

(b)

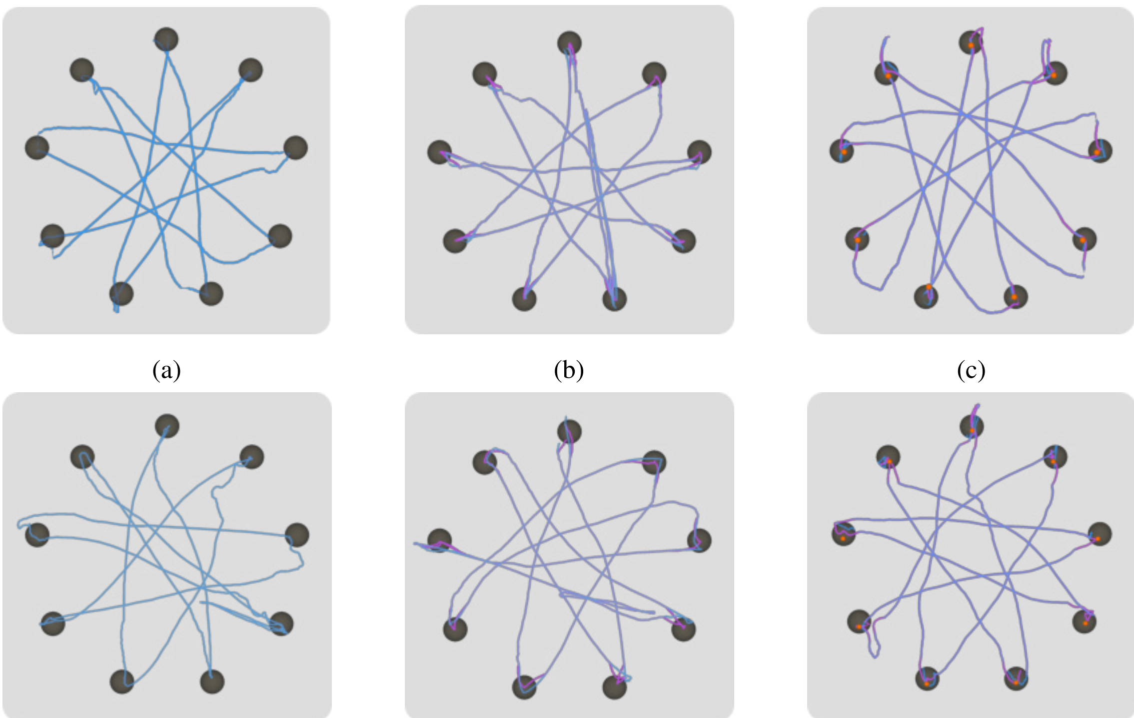
Figure 7.8: Effective throughput (bits/s): (a) Barplots split by depth and (b) significant simple comparisons.

We performed a two-way ANOVA and found significant main effects for both, *Hand-Ray Technique* ($F(2,58) = 23.446, p < .001, \eta_p^2 = .447$) and *Depth* ($F(1,29) = 9.938, p = .004, \eta_p^2 = .255$), on TP_e. The interaction effect between *Hand-Ray Technique* and *Depth* was also significant ($F(2,58) = 16.702, p < .001, \eta_p^2 = .365$), which limits the conclusiveness of the two main effects.

Therefore, we conducted a follow-up simple main effects analysis. The tests revealed a significant simple main effect of *Depth* on TP_e for *RayToGaze* ($F(1,29) = 28.9, p < .001$). Moreover, they indicated a significant simple main effect of *Hand-Ray Technique* on TP_e at a target depth of $1.3m$ ($F(2,58) = 20.6, p < .001$) and $4m$ ($F(2,58) = 24.3, p < .001$). Simple comparisons revealed significant differences in TP_e between five pairs of conditions, as listed in Table 7.8(b). Figure 7.8(a) visualizes the significant comparisons between conditions for each target depth.

As TP_e is also dependent on W_e, we performed a multi-factor Aligned Rank Transform (ART) ANOVA. The results revealed significant main effects for *Depth* ($F^1_{145} = 474, p < .001, \eta^2_p = .0.77$), but no significant interaction effects between *Hand-Ray Technique* and *Depth* were found.

Additionally, in Figure 7.9, we plotted sample movement paths of the Fitts' Law selection task for one participant to inspect the movement in each *Hand-Ray Technique*. The movement paths for *Baseline* appear more jittery and unstable close to the target spheres. For *RayToTarget*, the blue, tracked movement paths also show the same jittering, but the purple, redirected path illustrates the stabilized movement close to the target spheres. The movement paths for *RayToGaze* reveal that this particular participant seemed to overshoot or slip off the targets for some of the selections.



(a) (b) (c)

(d) (e) (f)

Figure 7.9: Selection task sample movement paths of a participant with ID_{DP} 3.7: (a) Baseline–1.3m, (b) RayToTarget–1.3m, (c) RayToGaze–1.3m, (d) Baseline–4m, (e) RayToTarget–4m, and (f) RayToGaze–4m. Blue Path: tracked Hand-Ray position, Purple Path: redirected Hand-Ray position, and Orange Sphere: gaze position.

7.4.2 NASA RAW TLX

For testing H2, we considered *Hand-Ray Technique* as our independent variable and all TLX (sub)scales of the NASA-TLX questionnaire as dependent variables. As none of the six TLX subscales fulfilled the normality assumption, we performed one Friedman test for each of them. For the *Total Workload* (TW), residuals were normally distributed, and we therefore conducted a one-way ANOVA. For better comparison of the measured workload across all (sub)scales, we normalized the TW value by dividing the total score by six, which corresponds to the number of TLX subscales.

Overall, *Baseline* (Mdn= 37.5, Q1-Q3= $[27.1-48.1]$) has the highest measured TW, followed by *RayToGaze* (Mdn= 29.6, Q1-Q3= $[21-39.2]$), and *RayToTarget* (Mdn= 27.5, Q1-Q3= $[16-37.1]$). *Baseline* also consistently received the highest median score on all TLX subscales, followed by *RayToGaze*, while *RayToTarget* received the lowest scores. All TLX subscales except for *Effort* have a measured task load below 40 for all study conditions. The detailed summary of the descriptive statistics is listed in Table 7.2.

The Friedman tests indicated significant differences for the following subscales: *Mental Demand* (χ^2(2) = 11.3, p = .004), *Effort* (χ^2(2) = 14.5, p < .001), *Performance* (χ^2(2) = 19.5, p < .001), and *Frustration* (χ^2(2) = 7.6, p = .022). Post hoc pair-wise Wilcoxon signed-rank tests with Holm-Bonferroni correction revealed that measured *Mental Demand* ($V = 330.5, p = .002$), *Effort* ($V = 244, p < .001$), and *Frustration* ($V = 279, p = .025$) are significantly lower for *RayToTarget* compared to *Baseline*. *Performance* is significantly worse (inverted scale in NASA TLX, where a higher score represents poor performance) for *Baseline* compared to *RayToTarget* ($V = 330.5, p < .001$) and *RayToGaze* ($V = 330.5, p = .001$). For TW, the ANOVA results indicated a significant difference between the *Hand-Ray Techniques* (F(2, 75) = 9.16, p < .001, η^2_G=.063). A post hoc paired samples t-test showed that TW is significantly lower for *RayToTarget* ($t(29) = 3.938, p = .001$) and *RayToGaze*($t(29) = 2.6279, p = .027$) compared to *Baseline*. The barplots in Figure 7.10 visualize the TLX scales that have significant statistical differences between conditions. A plot containing all TLX scales can be found in Appendix A4.1.

Table 7.2: Descriptive statistics for NASA RAW TLX subscales and Total Workload.

	Baseline		**RayToTarget**		**RayToGaze**	
Scale	M±SD	Mdn[Q1-Q3]	M±SD	Mdn[Q1-Q3]	M±SD	Mdn[Q1-Q3]
Mental	31.7±23.2	25[15-45]	21.8±18.6	15[10-25]	23.7±22.0	17.5[10-25]
Physical	36.3±22.7	35[15-53.8]	27.2±17.8	25[15-35]	34.3±22.3	30[16.2-48.8]
Temporal	37.7±20.8	37.5[20-53.8]	34.8±20.1	30[15-50]	38.5±20.5	35[21.2-53.8]
Performance	38.8±17.4	35.[30-48.8]	27.0±17.4	22.5[16.2-35]	27.3±14.4	25[16.2-38.8]
Effort	47.0±23.8	52.5[26.2-68.8]	34.7±23.4	35[16.2-45]	41.7±22.5	40[20-63.8]
Frustration	30.3±21.9	22.5[15-43.8]	21.5±20.9	17.5[5-28.8]	23.2±16.2	22.5[10-30]
Total	37.0±15.6	37.5[27.1-48.1]	27.8±14.1	27.5[16-37.1]	31.4±4.5	29.6[21-39.2]

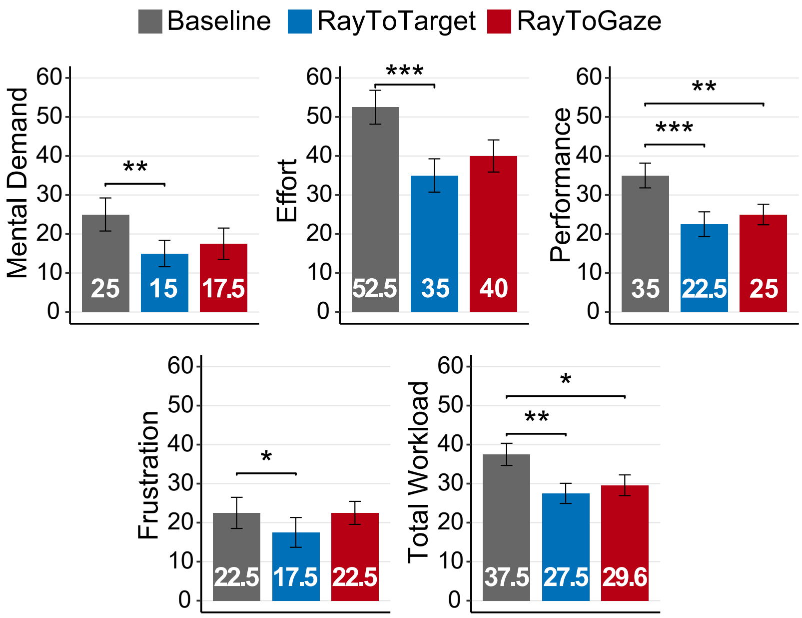


Figure 7.10: Barplots for the significant scales of the Nasa RAW TLX.

7.4.3 Sense of Agency

We did not find any negative effects on the perceived sense of agency for any of the three conditions (H3). In Figure 7.12, we plotted and visually inspected agreement-disagreement Likert scales for each of the four agency statements. As visible in the plotted scales, the agreement rate for all four statements is very high, with at least $>= 93\%$ in all conditions. Around a third of the participants ($27\% - 43\%$) strongly agreed with each statement. No participant expressed strong disagreement with the statements in any of the conditions. Only very few participants *disagree* or *somewhat disagree* with the statements, with overall $n <= 2$ for each condition.

7.4.4 Subjective Ranking of Preference

We matched the participants' voting for their subjective preference for each study block to the respective conditions, as visualized in Figure 7.11. The stacked bar plot reveals different preferences regarding selection with versus without redirection (H4), and differences in preference between the two redirection techniques (H6).

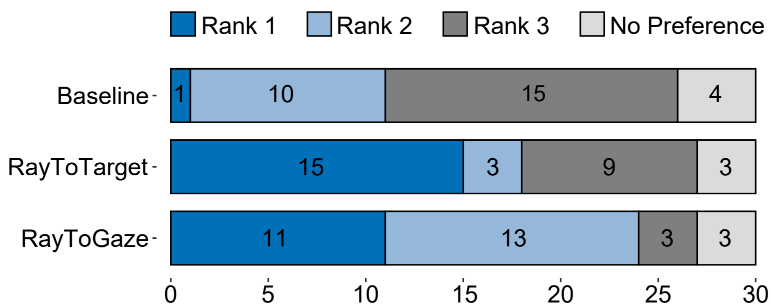


Figure 7.11: Subjective ranking for each *Hand-Ray Technique*.

In total, 26 out of 30 participants ranked all three conditions from most (Rank 1) to least preferred (Rank 3). Even if participants stated that they did not notice or could not describe the perceived differences, most of them still had a clear preference based on their experience and selection performance during the study. Only 2 participants had no preference at all (6 votes). One participant stated that they prefer *RayToGaze* the most, but had no preference between *RayToTarget* and *Baseline*. Another participant preferred *RayToTarget* the least, with no difference between *RayToTarget* and *Baseline*. This results in a total of 10 votes with no preference. Looking at the individual ranks, for Rank 1, *RayToTarget* is the most preferred *Hand-Ray Technique* with 15 votes, followed by *RayToGaze* with 11 votes. Only one person preferred *Baseline* over both redirection conditions. For Rank 2, the number of votes for *RayToGaze* (13 votes) and *Baseline* (10 votes) is close. The votes for rank 3 clearly show that *Baseline* (15 votes) is least preferred, followed by *RayToTarget* (9 votes), while *RayToGaze* only received 3 votes.

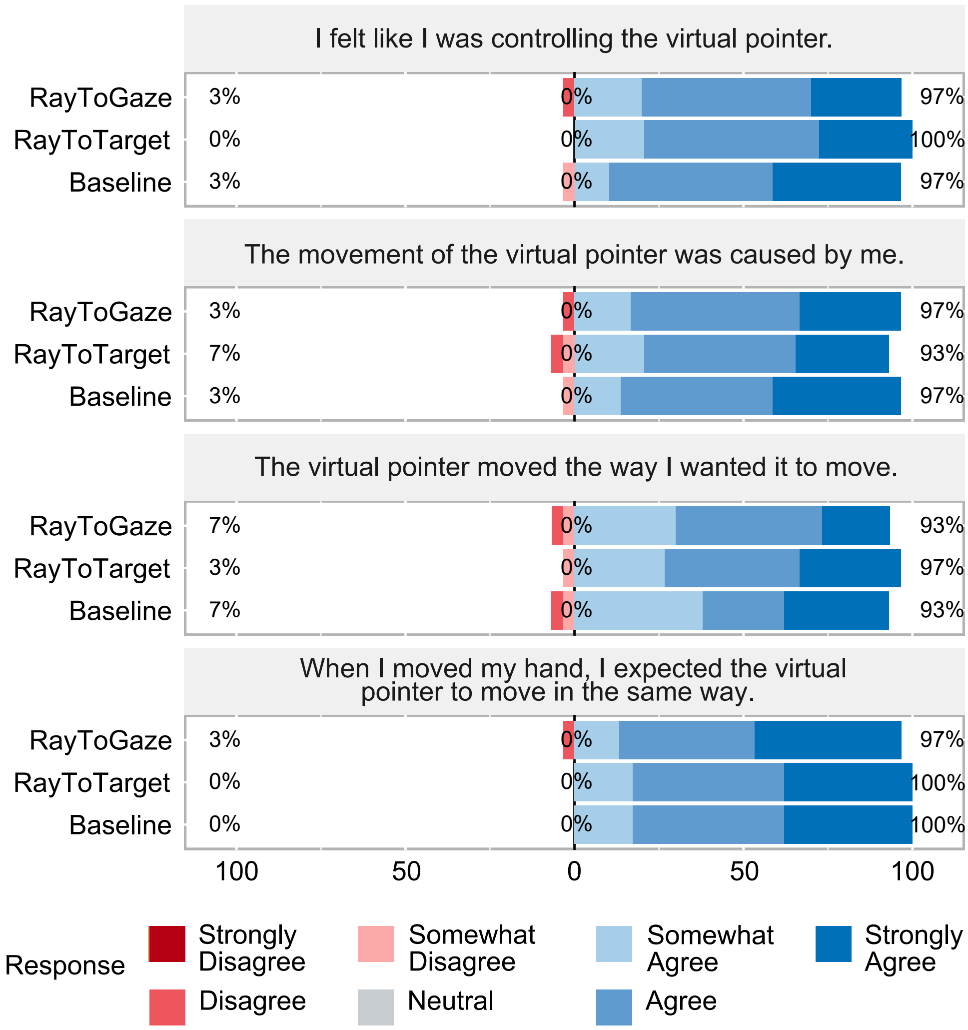


Figure 7.12: Agreement-disagreement Likert-scales for sense of agency statements.

7.4.5 Qualitative User Feedback

For the qualitative feedback, we followed a similar approach as in our previous study on Controller-Ray redirection (cf. Section 6.4.5). We also took inspiration from exploratory, inductive content analysis [Arm17]. We started with sorting, grouping, and categorizing the qualitative feedback. Corresponding statements within categories were then paraphrased and the number of occurrences was counted. Please note that only occurrences of explicitly mentioned statements are reflected by this type of qualitative content analysis. While it provides insights into aspects that might have been more relevant or memorable for the specific participants, it does not represent their full perspectives and opinions. An overview of statements with occurrences $n >= 3$ for each condition is listed in Table 7.3. As an exception, two individual statements with lower occurrences were included in the table for the *Agency* category of the *RayToGaze* method, to reflect the contradictory nature of these statements between to participants. The main findings based on the qualitative feedback are summarized by category as follows:

General

In total, 23 out of the 30 participants (76%) reported that they noticed a difference in the selection interactions. 15 of the participants stated that the selection in all three study blocks (conditions) was different. Only one participant explicitly reported that they noticed the gaze targeting during the selection. For the two conditions with Hand-Ray redirection, many participants mentioned that the pointer was snapped or pulled towards the target ($n = 11$, 36%) and described it as sticky or magnetic ($n = 8$, 26%). Two participants also noted that there was *aim assist*. Three participants explained that they did not always keep looking at the target while they were performing the pinch gesture for confirming their selection. Instead, they were already looking for the next target in the Fitts' Law selection task sequence. In contrast, one participant stated that they always focused on the current target during the whole selection process before moving on to the next target. Two participants wondered whether their performance would have been better if they had been made aware of the gaze-based redirection beforehand.

Performance

Many participants stated that the *Baseline* condition was more difficult, while selection was considered easier for both Hand-Ray redirection techniques. 5 participants reported that they made the most mistakes in *Baseline*, while one participant said that they made the most mistakes in *RayToTarget* because they felt overconfident due to the fast movement.

Agency and Pointer

Several participants reported that the selection in the *Baseline* condition without redirection felt most like their real movement, but it was also mentioned that the pointer was jittery and slipped off the target. Some participants expressed that *RayToTarget* provided too much assistance and that the pointer moved too fast. In one case, *RayToGaze* received contrasting feedback from two individuals: While one participant felt they had less autonomy, another said that they had the most control in the *RayToGaze* condition.

User Experience

Various positive and negative sentiments were stated by participants. Negative sentiments for *Baseline* were: exhausting, having the least control, more mental load, and less supportive. Positive sentiments for *RayToTarget* and for*RayToGaze* were: satisfying, more comfortable, and less frustrating. Specific positive sentiments were: motivating for *RayToTarget*, and great and pleasant for *RayToGaze*.

Table 7.3: Qualitative analysis of the user feedback with paraphrased statements.

	Baseline	**RayToTarget**	**RayToGaze**
Performance	- most mistakes (5) - more difficult (11)	+ fastest (2) + easier selection (10)	+ easier selection (10)
Agency	+ real movement (4)	- too much assistance (3)	- less autonomy (1) + most control (1)
Pointer	- jitter (4) - slipping off target (4)	- too fast (4)	- jitter (4)
UX	- negative sentiments (4)	+ supportive (5) + positive sentiments (5)	+ supportive (4) + positive sentiments (5)

7.5 Discussion

We proposed two different redirection techniques based on assistive redirection of the raycast for facilitating spatial interactions with Hand-Rays. In our user study, we compared both redirection techniques to classic Hand-Ray as a baseline to investigate selection performance and qualitative aspects. In the following, we will discuss our findings and elaborate on the differences between the techniques. Based on our findings, we will also address implications for designing and using Hand-Ray redirection techniques.

RQ1: Does assistive Hand-Ray redirection improve object selection performance and UX compared to classic Hand-Ray?

Selection Performance

Our Fitts' Law results on MT, error rate, and TP_e confirmed our H1a, H1b, and H1c. We found that both Hand-Ray redirection techniques had better selection performance than *Baseline*. The error rate for *Baseline* also has more variability with a large SD (see Figure 7.6(a)), which indicates that several participants had difficulties with performing successful selections. This is not surprising, since classic Hand-Ray pinch is more prone to input-induced errors due to the Heisenberg Effect of Spatial Interaction [Bow+01] (see Section 2.3). Figure 7.9 illustrates this effect with the more unstable and shaky movement path of *Baseline* close to each target sphere in the Fitts' Law task setup. If we compare the Hand-Ray study results with our previous work (c.f. Chapter 6), Hand-Ray interaction is generally less efficient than Controller-Ray. The movement times and error rates are higher, and the effective throughput is lower for Hand-Ray – we assume that this is due to the required pinch gesture. However, while assistive Controller-Ray redirection had no significant effect on the error rate, assistive Hand-Ray redirection significantly reduced the error rate by around 4%. This comparison suggests that classic Hand-Ray interaction might be more challenging than classic Controller-Ray interaction, which supports the use of assistive redirection for facilitating hand-tracking based spatial interaction.

Perceived Workload

The NASA Raw TLX results support our H2. We found that both Hand-Ray redirection conditions have lower perceived *Total Workload* and higher perceived *Performance* compared to *Baseline*. This also matches our Fitts' results for selection performance. For the subscales *Mental Demand*, *Effort*, and *Frustration*, only *RayToTarget* received significantly lower scores compared to *Baseline*. We will look into possible explanations when discussing RQ2.

Sense of Agency

Using Hand-Ray redirection did not negatively affect the participants' sense of agency. Both redirection techniques received a very high agreement rate of >=93% for each agency statement, which validates our H3. 76% of the participants ($n = 23$) noticed a difference between the conditions, and many mentioned snapping or pulling ($n = 11$) and a sticky or magnetic ($n = 8$) behavior of the pointer. This suggests that participants still felt in control of their movements and the controlled Hand-Ray, even though they noticed the redirection. We expected similar results to our previous study on Controller-Ray redirection (see Section 6.4.3), since we used the same redirection algorithm for Hand-Ray redirection.

User Preferences and Qualitative Feedback

Having Hand-Ray redirection was preferred by the majority of participants ($n = 26$). 15 participants ranked *Baseline* last (Rank 3), while only 1 person ranked it as most preferred (Rank 1). This makes Hand-Ray the least preferred technique based on the subjective ranking. This result confirms our H4. Furthermore, participants either preferred *RayToTarget* over *RayToGaze* or vice versa, which supports our hypothesis H6.

The qualitative feedback also matches our H1a, H1b, H1c, H2, and H4. Overall, both Hand-Ray redirection techniques received mainly positive statements. They were described as supportive, more comfortable, and less frustrating. In contrast, most statements for *Baseline* were negative (Table 7.3). Participants felt that they made the most mistakes in *Baseline*, which is confirmed by the measured Fitts' results for error rate. They stated that selection was more difficult, exhausting, and less supportive. This is also reflected in the significantly higher TLX subscales and perceived Total Workload for *Baseline*. Moreover, the reported pointer jitter and the pointer slipping off the selection target could have contributed to higher error rates. This can actually be seen in the plotted sample movement paths in Figure 7.9. In our discussion of RQ2, we will address potential reasons why the pointer slipped off the target for some participants.

RQ2: Are there differences in selection performance and UX between the two Hand-Ray redirection techniques?

Selection Performance

We did not find any significant differences in the Fitts' Law measurements' main effects between *RayToTarget* and *RayToGaze*. However, our statistical analysis revealed significant interaction effects for $TP_e \times Depth$. The redirection techniques have higher TP_e compared to *Baseline*. Interestingly, at a *Depth* of $4m$, TP_e visibly decreases for *RayToGaze* and is significantly lower compared to *RayToTarget* (Figure 7.8(a)). Moreover, TP_e slightly deceases for *Baseline* at $4m$, while it does not change for *RayToTarget*. These results do not support our H5a and H5b, but they confirm H5c, as our findings suggest that TP_e for *RaytoGaze* is negatively affected by increasing target depth, while TP_e for *RayToTarget* remains robust regardless of *Depth*. This effect of *Depth* on TP_e would explain the results of the NASA TLX, where *RayToGaze* did not have significantly lower perceived workload in several subscales compared to *Baseline*, while *RayToTarget* did. The qualitative feedback provides insights about possible causes for this inconsistency.

User Preferences and Qualitative Feedback

The qualitative data revealed that the ranking for subjective preference was more divided for *RayToTarget* compared to *RayToGaze* (see Figure 7.11). For *RayToTarget*, some participants reported that it was the fastest technique, selection was easy, and it was supportive. Others expressed that *RayToTarget* had too much assistance, and that the pointer moved too fast. This might explain the contrasting subjective ranking for *RayToTarget*. One participant in particular also stated that they made more mistakes because they became overconfident with *RayToTarget* due to the fast selection in this condition. *RayToGaze* received some contrasting feedback as well. Many participants said that the selection was easy, whereas others reported pointer jitter. One participant felt that they had less autonomy, while another thought they had the most control with *RayToGaze*. These opposing views further support our H6.

Based on the participants' feedback, we assume that individual visual search and selection strategies appear to play a role in gaze-based redirection. Some participants only fixated on the target during the targeting phase, but not during confirmation of the selection. This is consistent with the issue of *Early and Late Triggers* that Kumar et al. [Kum+08] and Pfeuffer et al. [PGG24] identified for multi-modal interaction, where the eyes leave the target before the manual selection command is registered. This might have contributed to the interaction effects for $TPe \times Depth$. If participants looked away from the target before the pinch gesture was completed, the redirection was no longer applied. This means that in these instances, no assistance was provided during the selection confirmation. This could explain the reported pointer jitter for *RayToGaze* and the decrease in TPe at a target depth of 4m. We hypothesize that participants might look more closely at smaller and closer targets compared to larger, more distant ones. Since we used the same angular sizes at depths $1.3m$ and $4m$ in our Fitts' Law tasks, the targets had a much larger actual target width W at depth $4m$. We speculate that the performance of *RayToGaze* could potentially be improved by revealing the gaze-based redirection to users. If this assistive redirection is known in advance, users could focus their gaze more consciously on the targets during the selection.

To summarize, our qualitative insights suggest that there are differences in individual preferences and in the way participants approached the selection task. This influences the UX and might also affect the selection performance between *RayToTarget* and *RayToGaze*. Both Hand-Ray redirection techniques had specific aspects that particular participants liked and disliked. For the practical application of Hand-Ray redirection, we suggest that the redirection parameters, such as gain factor and redirection zone, should be adjustable or adaptive to accommodate individual preferences and different types of users.

7.6 Limitations and Future Work

Since we used an abstract standardized Fitts' Law selection task in our study, the practical application of Hand-Ray redirection should be further investigated for other, realistic use cases.

Moreover, *RayToTarget* has limitations in cluttered environments since each object needs a larger, pre-defined redirection zone, which would lead to intersections with nearby targets. For example, when typing on a virtual keyboard, the keys would usually be placed too close to each other to use *RayToTarget* effectively. For these use cases, *RayToGaze* would be a suitable alternative. As our study revealed, *RayToGaze* has the same benefits as *RayToTarget* in terms of most performance parameters and UX measurements. The application and performance of *RayToTarget* in cluttered environments should be evaluated together with other existing target disambiguation techniques in future work.

Our findings indicate that users might apply different visual target search and selection strategies. This should be further investigated and validated in the context of gaze-assisted and hand-tracking-based interaction techniques.

Our study had only 2 left-handed participants. Furthermore, we mainly tested younger users with a mean age of 26 years, and out of our 30 participants, only 8 had never used hand tracking in VR before. A Spearman's correlation test did not indicate a correlation between error rate and participant age ($r_s(28) = -.114, p = .549$). According to the Kendall's tau-b correlation there was also no correlation between error rates and previous VR experience ($\tau_b = -.040, p = .771$). Nevertheless, future evaluations of assistive Hand-Ray redirection should include a more diverse user group.

In this chapter, we investigated the performance and UX for raycast redirection as an assistive technique for improving Hand-Ray interaction. Our two proposed redirection techniques should be tested against other established hand- and gaze-based interaction techniques in a comparative study. Since we only used a fixed redirection zone of 2 $\times$ the target size, it would also be interesting to examine different redirection zone parameters and adaptive resizing of the redirection zone. In theory, a redirection zone of 1 $\times$ the target size should still stabilize the pointer during the selection confirmation gesture. Comparisons with controller-based raycast redirection (see Chapter 6), direct snapping like *Sticky Ray* (see Section 2.4.1), and with regular targets that simply have larger hitboxes (3D colliders) could also be considered in the future.

We also cannot rule out that the HMD used for the study could have caused hand- and gaze-tracking inaccuracies that might have impacted the study results. Therefore, this empirical study should be replicated on other HMDs to evaluate the robustness of redirected Hand-Rays independently of the hardware used.

7.7 Conclusion

This chapter's research looked into assistive Hand-Ray redirection based on hand- and eye-tracking. Both input modalities remain highly relevant for spatial interaction, as built-in hand- and eye-tracking is becoming increasingly available on various modern HMDs. In our work, we proposed and evaluated two assistive Hand-Ray redirection techniques for facilitating far interactions for object selection: (i) *RayToTarget*, which redirects the Hand-Ray towards the target center, and (ii) the gaze-assisted *RayToGaze*, which redirects the Hand-Ray towards the user's gaze position on the target. We conducted a comparative Fitts' Law selection study with both redirection techniques and classic Hand-Ray as a baseline. We measured the selection performance, sense of agency, perceived workload, subjective preference, and UX. Our results show that Hand-Ray redirection provides the same benefits as Controller-Ray redirection (see Chapter 6). The redirection significantly improves selection performance and reduces perceived workload, while the sense of agency is not negatively affected. However, performance for *RayToGaze* significantly decreases with increased target depth, while it remains robust for *RayToTarget*. Our qualitative analysis revealed that Hand-Ray redirection was generally preferred, but individual preferences and differences in target acquisition strategies could lead to distinct and contrasting UX between participants. These aspects might also impact the performance of both redirection techniques.

This chapter builds on our previous work on assistive Controller-Ray redirection. We confirmed that raycast redirection can also be effectively applied for Hand-Rays. Additionally, we further extended Hand-Ray redirection by adding gaze as input modality for defining the target position of the redirection. Our work uncovered advantages, but also challenges and potential issues that need to be answered for the development and practical application of assistive raycast redirection for spatial interaction. It further highlights the importance of combining quantitative measurements with subjective insights for assessing novel interaction techniques. With our work, we hope to encourage further research on assistive interaction techniques and raycast redirection in combination with different input modalities for improving spatial interaction.

CHAPTER 8

GAZE&BLINK: HANDS-FREE SPATIAL INTERACTION

The content of this chapter is based on a preprint [Tim+25]. Parts of the text have been rephrased and revised, and additional figures and tables have been added.

8.1 Motivation

In this chapter, we further investigated eye-tracking-based techniques for spatial interactions in order to answer our RQ-V:

RQ-V
Can we develop a gaze-based hands-free spatial interaction technique for discrete and continuous interaction tasks?

With our work on eye-tracking-based interaction, we aimed to contribute to the diversity of interaction techniques by proposing a novel, hands-free solution for spatial interaction on HMDs. So far, our work focused on discrete selection tasks, which consist of simple selections of target objects. However, interacting with UIs on modern HMDs also requires performing various continuous interaction tasks, such as drag-and-drop and scrolling. While existing hands-free techniques, such as *Gaze&Dwell* [Han+03; Mot+17; MBS21] and *Outline Pursuits* [Sid+20], support discrete selection, solutions for handling continuous interaction tasks have not yet been established and formally evaluated. Nowadays, interest in gaze-based interaction techniques has increased due to the availability of commercially available HMDs with built-in eye-tracking, such as the Meta Quest Pro, the Vive Focus Vision, the Apple Vision Pro (APV), and the Varjo XR-4. Thus, we were motivated to fill this research gap by proposing *Gaze&Blink*, a new eye-tracking focused interaction technique for performing hands-free discrete and continuous spatial interaction tasks.

We were inspired by related work and the potential of combining gaze- and blink-based input for implementing discrete and continuous spatial interaction tasks. Performing a blink only requires approximately 120 ± 2 ms [Duc17]. In contrast, existing hands-free techniques, such as *Gaze&Dwell*, rely on a dwell threshold, where 300ms is considered a low threshold, and *Outline Pursuits* requires even more time for executing the smooth pursuit with the eye movement. For the implementation of continuous interactions, we also took inspiration from existing work of Nukarinen et al. [Nuk+16], who combined gaze with head rotation for desktop system control.

Another focus of our work was to investigate whether our proposed spatial interaction technique would be a viable, hands-free alternative to *Gaze&Pinch* [Pfe+17]. Although *Gaze&Pinch* is an established technique that supports discrete and continuous spatial interaction tasks, it might not be suitable in certain scenarios. These include, for example, public, confined spaces, where pinch gestures and arm movements might be inappropriate, or cases where users have limited mobility or restrictions in performing hand and arm movements. Therefore, our aim was to provide a hands-free solution with *Gaze&Blink* that would offer comparable performance to *Gaze&Pinch*.

For evaluating our gaze-blink-based interaction technique, we examined the following research questions:

RQ1 How can we implement an interaction technique that covers both discrete and continuous interaction tasks through *Gaze&Blink*?

RQ2 How does *Gaze&Blink* compare to *Gaze&Pinch* for different discrete and continuous UI interactions?

RQ3 What are the advantages and challenges of *Gaze&Blink*?

For answering these questions, we implemented our *Gaze&Blink* technique (RQ1) and conducted two user studies (RQ2, RQ3). In the first study, we evaluated our technique against *Gaze&Pinch* in a realistic UI setup featuring familiar UI layouts and elements. Although *Gaze&Blink* had similar performance in terms of speed and UX, participants triggered unintended selections with involuntary blinks. Based on these findings, we formulated another research question to investigate potential improvements:

RQ4 Are there optimization methods to reduce involuntary blinks for *Gaze&Blink* that could improve this interaction technique?

To address RQ4, we created a deep-learning model for classifying voluntary and involuntary blinks. We implemented *Gaze&BlinkPlus* by extending *Gaze&Blink* with the trained model for blink classification. Our second study evaluated *Gaze&Pinch*, *Gaze&Blink*, and our extended technique *Gaze&BlinkPlus*. To summarize, our contributions in this chapter are as follows:

C1 *Gaze&Blink* and its extension *Gaze&BlinkPlus* as novel hands-free interaction techniques for discrete and continuous interaction tasks.

C2 Two comprehensive user studies with realistic menu UIs and interaction tasks, comparing *Gaze&Pinch* with our proposed techniques.

C3 Comprehensive insights into interaction performance, perceived workload, qualitative aspects, and challenges for integrating blinks in spatial interaction techniques.

8.2 Gaze&Blink Interaction Technique

This section describes the interaction design, including the input mechanism and interaction states, of our novel hands-free interaction technique *Gaze&Blink*. For examining RQ1, we combined gaze- and blink-based input for performing discrete selections, based on concepts in existing work (cf. Section 2.5.4 and Section 2.5.5). The implementation of hands-free eye-tracking-based continuous interactions is more challenging because it requires continuous movement tracking. Since saccadic eye movements are not suited for smooth continuous movement input, we designed an interaction technique that integrates head-tracking with gaze- and blink-based input. For our interaction design, we took inspiration from existing approaches that use eye openness and head interactions (cf. Section 2.5.5 and Section 2.5.6). To map these multi-modal input requirements, we defined the following five interaction states for *Gaze&Blink*:

- **Default:** When both eyes are open
- **Selection:** When both eyes are closed
- **Drag start:** Marked by closing of only one eye
- **Drag update:** When one eye is closed and the head is moved
- **Drag end:** When the eye is opened again

These five states for our interaction technique are described in the state graph in Figure 8.1. If both eyes are open, we are in the default state, where no interaction is performed. As soon as both eyes are closed below a threshold, which is pre-calibrated for each user, a selection confirmation is triggered. To prevent unintended drag interactions while in the selection state, we can only return to the default state. If only one eye is closed, the drag interaction is initiated from the default state. While head movement is detected, a drag update is performed. The drag interaction ends when both eyes are opened or fully closed.

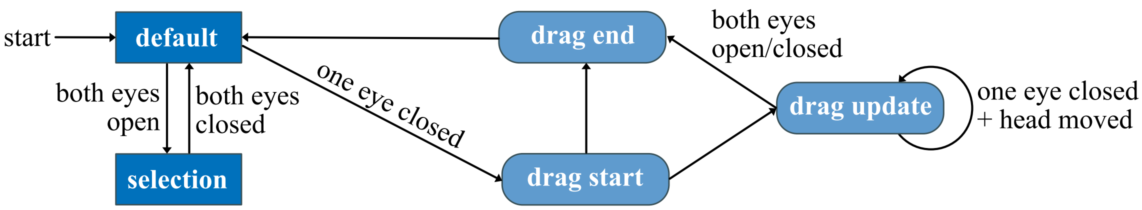


Figure 8.1: State graph for our Gaze&Blink interactions: (rectangles) discrete and (rounded rectangles) continuous interaction.

These five states are used to define the two types of interaction for *Gaze&Blink*:

(D) Discrete interaction for selection

(C) Continuous interaction for scroll and drag

8.2.1 (D) Discrete Selection

Figure 8.2(a) illustrates the discrete *Gaze&Blink* interaction task and compares it with the equivalent *Gaze&Pinch* interaction [Pfe+17]. *Gaze&Blink* selection is composed of target indication by looking at the target and selection confirmation by blinking, as pictured in Figure 8.3:

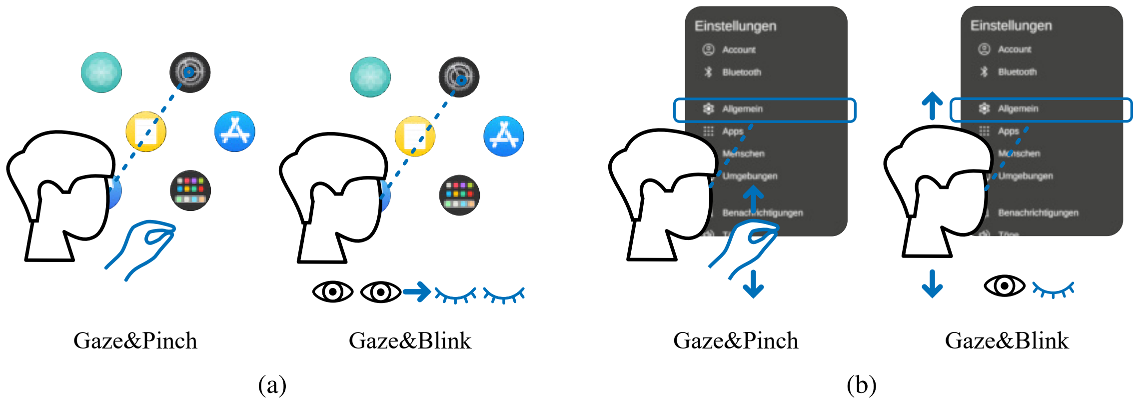


Figure 8.2: Comparison of discrete and continuous interactions for *Gaze&Pinch* and *Gaze&Blink*: (a) Discrete interaction: selection and (b) continuous interaction: scroll and drag

1. Default state: both eyes open ⇒ target indication
2. Selection state:
   - i Start: both eyes closed ⇒ target selection
   - ii End: both eyes open ⇒ return to default state

The blink input is detected with a specified eye openness threshold value (see Section 8.3.1). Blinks can be performed nearly instantly by users (120 ± 2ms) [Duc17]. This allows for a fast update of the blink selection confirmation state and makes it suitable for interaction tasks with a high input frequency, such as keyboard input. In contrast, *Gaze&Dwell* has higher selection times and lower throughput than *Gaze&Pinch* [MBS21] due to the required dwell activation threshold (300ms for a low threshold) [Ami+25; MBS21]. Therefore, we assumed that *Gaze&Blink* would be a more suitable alternative to *Gaze&Pinch* in terms of interaction performance.

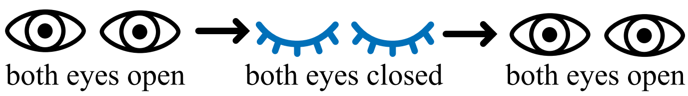


Figure 8.3: Illustrated *Gaze&Blink* selection sequence.

8.2.2 (C) Continuous Scroll and Drag

For continuous interaction tasks, we designed *Gaze&Blink* for supporting scrolling and drag-and-drop. As illustrated in the input sequence in Figure 8.4, the continuous movement is performed by keeping one eye closed while moving the head in the desired direction. Since the movement is controlled based on the head rotation and not the gaze position, it is smooth and stable. In addition, this enables users to visually inspect the movement and their surroundings while performing the interaction. The sequence for this continuous interaction is defined as follows:

1. Default state: both eyes open ⇒ target indication
2. Drag state:
   - i Start: one eye closed ⇒ target locked

ii Update: one eye closed & head rotation ⇒ target movement

iii End: both eyes open ⇒ target released, return to default state

The rotation-based movement is calculated based on raycasting onto a plane, where the intersection point is used as the origin position. Depending on the delta between the current and previous point (angular movement), the UI object is then moved by the rotational gain amount. This approach results in a non-linear Control-Display (C/D) ratio. The movement follows an angular mapping, similar to raycasting with a pointer. This means that the C/D ratio between head-rotation and angular gain of the object movement increases with distance from the head to the target object. For distant targets, a small head-rotation would lead to a larger movement compared to a close target.

Since continuous interactions can only be triggered by closing one eye (see Figure 8.1), *Gaze&Blink* does not require a movement threshold for distinguishing between performing discrete and continuous interactions such as *Gaze&Pinch*. This continuous interaction matches scrolling or drag-and-drop interaction tasks performed with *Gaze&Pinch* (see Figure 8.2(b)).

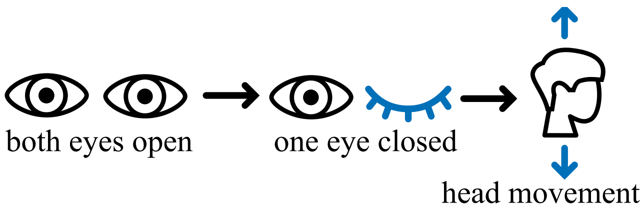


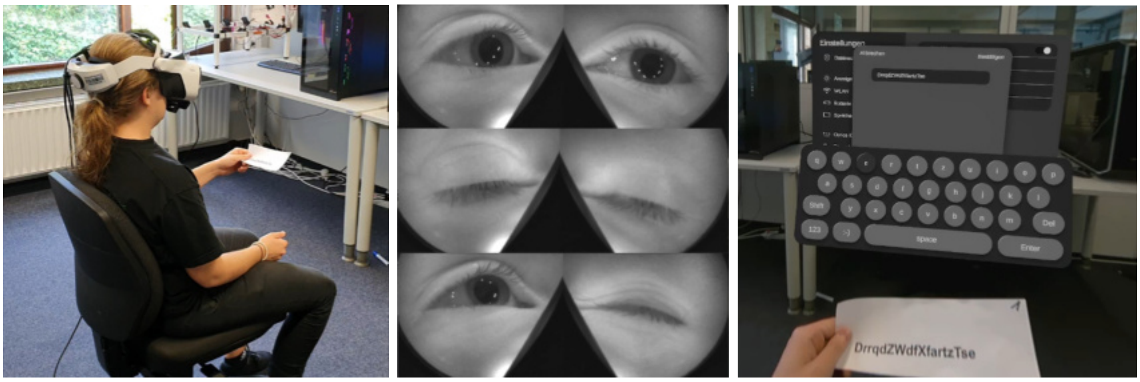
Figure 8.4: Illustrated *Gaze&Blink* scroll sequence.

In summary, *Gaze&Blink* implements both discrete and continuous interactions with comparable functionality to the *Gaze&Pinch* interaction technique. To our knowledge, at the time of developing and investigating *Gaze&Blink*, no other techniques for performing continuous scroll and drag-and-drop interaction tasks with corresponding functionality to hand-tracking-based interactions have been proposed for hands-free eye-tracking-based spatial interactions in XR. Furthermore, existing hands-free, eyes-only spatial interactions have lower performance compared to blink-based techniques [Lu+20; PZ23]. Our objective was to design an interaction technique for efficient task execution that is also suitable for fast and frequent interaction tasks. Based on these considerations, we believe that *Gaze&Pinch* is a suitable baseline for testing the performance and UX of our proposed approach. An overview comparing the similarities and differences between *Gaze&Pinch* and *Gaze&Blink* is shown in Table 8.1.

Table 8.1: Comparison of properties for *Gaze&Pinch* and *Gaze&Blink* interaction.

	Gaze&Pinch	**Gaze&Blink**
Target Indication	gaze	gaze
Selection Confirmation	pinch with index and thumb	blink with both eyes
Continuous Movement	hold pinch and drag with hand	close one eye and rotate head
Modalities	gaze, hand	gaze, blink, head

8.3 User Study 1 – Gaze&Blink



(a) (b) (c)

Figure 8.5: *Gaze&Blink* interaction: (a) User wearing the Varjo XR-4 HMD, (b) example of eye tracking for a blink and a left eye wink, and (c) spatial UI with keyboard input task from our study.

8.3.1 Methodology

Study design

We conducted Study 1 to test the viability, evaluate the performance, and uncover potential usability issues of *Gaze&Blink*. Since we intended to conduct a comparison against the already established *Gaze&Pinch* in a realistic use case scenario, we designed a sequence of tasks, which included common UI menu and system control interaction tasks (see Section 8.3.1). We choose a menu interface that is inspired by the visual appearance, layout, and behavior of the AVP HMD (see Section 8.3.1), since it is the most recent and well known consumer-grade device with *Gaze&Pinch* as its default main interaction technique.

For our study, we used a Varjo XR-4[1], which is a PC MR HMD with a cable-tethered connection to a Windows computer. The desktop computer had an Intel Core i7 14700k CPU and an NVIDIA RTX 4080 graphics card with 32 GB RAM. For gaze- and blink-tracking, we used the integrated eye-tracker of the Varjo XR-4. An example of the eye-tracking input using blinking and winking is shown in Figure 8.5(b). We set our render resolution to 3840×3744 pixels and used a field of view (FoV) of $120° \times 105°$. For hand-tracking, we used the Ultraleap Leap Motion Controller 2[2] with the custom mounting kit from Varjo.

Study Tasks

Since this is our first evaluation of *Gaze&Blink*, we focused on basic UI interactions and did not consider more complex 3D object manipulation tasks such as scaling or rotation. We chose the following commonly used discrete and continuous UI menu and system control interaction tasks, which had to be completed in a fixed sequence:

[1] https://varjo.com/products/xr-4/
[2] https://www.ultraleap.com/products/

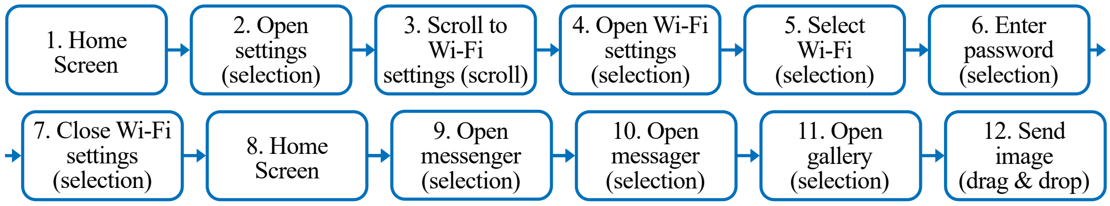


Figure 8.6: Sequence of interactions for completing one study task trial.

1. **Selection task:** Select the correct menu item (icons) from the main menu (first Settings, later Messenger menu).
2. **Scroll task:** Scroll through the settings menu to find Wi-Fi settings, select Wi-Fi settings, select toggle button.
3. **Text input task:** Select the correct keys on the virtual keyboard to input the Wi-Fi password, select confirm button.
4. **Drag-and-drop task:** Drag-and-drop an image from a folder into a messenger window.

We designed a scenario where each participant had to connect to a specified Wi-Fi network and then send a picture through the messaging application (see Figure 8.6). The task starts on the home screen of the interface without a Wi-Fi connection. Users were then tasked with establishing an internet connection by finding the settings menu on the home screen (*selection task*). Afterward, they needed to search for the Wi-Fi setting in the settings menu list (*scroll task*) and enable the Wi-Fi connection (*selection task*). Participants were then asked to enter the correct Wi-Fi password handed out on a piece of paper (*text input task*). Participants could clearly see their surroundings and the printed Wi-Fi password in the color passthrough of the HMD, as pictured in Figure 8.5(c). Once the correct password was successfully entered, users had to navigate back to the messaging app (*selection task*) and drag-and-drop a specific image into the text field (*drag-and-drop task*). The full sequence of interactions is visualized in Figure 8.6.

Based on our research questions, we formulated the following hypotheses to examine whether there are significant differences in performance and UX between *Gaze&Pinch* and *Gaze&Blink*:

H1a Task performance, measured as task completion time and error rate, differs significantly between *Gaze&Pinch* and *Gaze&Blink*.

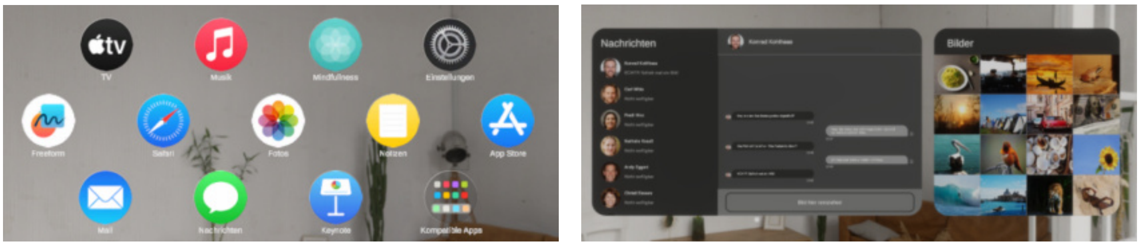
H2a Perceived workload differs significantly between *Gaze&Pinch* and *Gaze&Blink*.

H3a UX differs significantly between *Gaze&Pinch* and *Gaze&Blink*.

Technical Implementation and Assumptions

Due to the AVP's privacy restriction[3], which does not allow access to the raw eye tracking position, we conducted all our studies on the Varjo XR-4 that provides similar specifications to the AVP HMD. For the gaze interaction, we used the built-in eye-tracker of the Varjo XR-4 to perform gaze-based (invisible) raycasting for the interaction with the UI elements. To provide realistic interaction tasks involving familiar UI layouts, we decided to implement our UI design based on Apple's VisionOS UI. It offers standardized and familiar UI menus and elements based on

[3]https://www.apple.com/privacy/docs/Apple_Vision_Pro_Privacy_Overview.pdf



(a) (b)

Figure 8.7: Sample views of different menu interfaces used during our study: (a) Home screen with multiple application icons and (b) messenger with chat history and an image gallery window.

established UI patterns. The VisionOS UI provides common tasks and applications, such as a messenger, a settings menu, a photo gallery, and a virtual keyboard, as illustrated in Figure 8.7. We built the application in the Unity engine[4], version 2022.3.3f1, and implemented the UI elements with the Nova UI Framework[5] using VisionOS App Icons from Figma[6]. The UI is placed at 2.5m distance from the user according to guidelines for spatial interfaces [Ber+21b; Mic24]. To ensure good usability and a comfortable interaction, the sizes of the UI elements were chosen according to guidelines and based on parameters in existing work [Ber+21b; Met24]. Overall, the angular sizes of the used UI elements are between 1.06° and 22.23°. The detailed angular sizes for each specific selectable UI element are listed in Table 8.2.

While the reconstruction of missing data from eye tracking systems is quite common [GWM24], we did not interpolate with future eye-tracking data to avoid additional input latency during blinks in our interaction technique. In addition, since eye movements during eyelid closure [Kir+23] might cause unintended, additional interactions, we save and use the position of the last raycast if the openness of the eye is below a threshold of 0.7. The discrete and continuous interactions for *Gaze&Blink* are implemented according to the five states and our state graph in Figure 8.1. The eye openness for both eyes and when only one eye is open are based on a threshold, which is individually calibrated for each participant at the start of the study. To capture the gaze positions, we used the highest provided sampling rate of 200 Hz supported by the Varjo XR-4 and captured the data independently of the rendering frame rate of the application. Additionally, we provided auditory feedback whenever a participant interacted with a UI element.

For the pinch gesture in our *Gaze&Pinch* condition, we used the hand tracking system provided by Varjo based on the Leap Motion Controller 2 kit. It recognizes a gesture once thumb and index finger are held together below the recommended threshold of 0.8, as specified by Ultraleap[7]. To distinguish a selection confirmation from a drag gesture in *Gaze&Pinch*, we checked if the hand moved by a certain distance while performing a pinch. Here, we found that the best minimum distance is 7cm and the best minimum pinch duration is 300ms.

[4] https://unity.com/
[5] https://novaui.io/
[6] "VisionOS App Icons • Free Resource" by David Akhmedbayev (figma.com/@david_akh), CC BY 4.0
[7] https://docs.ultraleap.com/xr-and-tabletop/xr/unity/plugin/features/pinch-and-grab-detection.html

Table 8.2: Table with angular sizes of the UI elements.

Selectable UI Elements	Width	Height
Main UI Elements		
Main Menu Button	3.95°	3.95°
Close Window Button	2.05°	2.05°
Settings Menu		
Settings List Elements (WLAN, ...)	13.14°	1.96°
Wi-Fi Network Menu Bar	20.25°	2.02°
Wi-Fi Toggle Button	3.38°	1.56°
Password Input Field	18.06°	2.02°
Cancel / Confirm Button	3.93°	1.06°
Keyboard		
Keyboard Letter Key	2.71°	2.71°
Keyboard Shift Key	3.79°	2.71°
Keyboard Space Key	16.43°	2.71°
Message Menu		
Message Contact List Element	14.91°	3.07°
Send Image Button	1.66°	1.66°
Image Thumbnail	5.52°	5.52°
Drag-and-Drop Input Field	22.23°	2.33°

Questionnaires

We used a pre-study questionnaire with questions on demographics and XR experience, and we recorded the baseline SSQ [Bro+96]. After each study condition was completed, we again administered the SSQ and in addition, participants had to fill out the NASA RAW TLX [Har06a], SUS [Bro+96], and the short version of the User Experience Questionnaire (UEQ-S) [STH17]. At the end of the study, we collected qualitative feedback with two open questions, as specified in Table 8.3:

Table 8.3: Open questions used in our questionnaires. We used DeepL[8] for translation.

	German Question	Translated English Question
Q1	Welche der beiden Eingabemethoden hat Dir mehr gefallen? Bitte begründe Deine Antwort kurz.	Which of the two input methods did you like more? Please provide a brief reason for your answer.
Q2	Was ist Dir bei den Eingabemethoden besonders aufgefallen?	What did you particularly notice about the input methods?

[8] https://www.deepl.com/en/translator

Study Procedure and Participants

We recruited our participants through our university's study platform. The platform is primarily used by HCI and psychology students, who received study credits for their participation. We specified the following exclusion criteria for participants: No pacemaker or past cases of epilepsy, following Tobii eye-tracking safety guidelines[9], and no long, artificial finger nails with nail art that might cause pinch gesture tracking issues. All participants had normal vision or wore contact lenses. To ensure optimal eye-tracking performance, wearing glasses was not allowed. Since instructions during the study were given in German, all participants were required to have sufficient German language skills. We performed an a priori power analysis, which indicated a minimum sample size of $n = 15$ to detect large effects (Cohen's $d = .8$) with 80% power for a paired-samples t-test with an alpha level of 0.05.

The study was conducted in a stationary and seated MR setup, as pictured in Figure 8.5(a). First, participants signed a consent form and filled out the pre-study questionnaire. Next, they spent at least 150 seconds inside a small playground environment to familiarize themselves with the interactions and the interface. The starting condition was alternated between *Gaze&Pinch* and *Gaze&Blink* for each participant. The participants followed a fixed sequence of tasks, as specified in Figure 8.6. For each condition, the task sequence was repeated five times. For task number 6 (*Enter password*), we generated a total of five passwords for the five repeated task sequences, as listed in Table 8.4. The passwords were generated at random using a password generator. The password order for each participant was balanced using a Latin square design. The passwords are listed in Table 8.4. The table also includes the number of shift key inputs required and the summed up unit Euclidean distance between the keys for each password, where adjacent characters on the keyboard have a distance of 1. We calibrated the eye-tracker before each trial and we had the option to re-calibrate the eye-tracking on request at any time during the study. For optimal blink detection, we performed a manual eye-openness threshold calibration for the left and right eye on each participant. After each condition, participants took off their HMD to rest their eyes. They also filled out the questionnaires specified in Section 8.3.1 on a tablet. Finally, after completing all conditions and standardized questionnaires, they answered the two open questions. The study took approximately 60 minutes in total.

8.3.2 Results

For the statistical analysis, we first tested the normality of our collected data, except for the SSQ data, by performing Shapiro-Wilk tests on the differences between our paired measurements for the two interaction techniques: *Gaze&Pinch* and *Gaze&Blink*. In case the normality assumption was violated, we applied appropriate transformations zo the data to achieve a normal distribution. Afterward, we performed paired samples t-tests. We recorded SSQ data at three points in time – once pre-study and after each of the two conditions was completed. We tested the SSQ data for normality of residuals with the Shapiro-Wilk test. Appropriate data transformations were used if

[9] https://help.tobii.com/hc/en-us/articles/212372449-Safety-guidelines

Table 8.4: The Wi-Fi passwords used in task number 6 (Enter password, see Figure 8.6) with required amount of shift key inputs and unit Euclidean key distance on the keyboard.

Password	Shifts	Euclidean Distance
DrrqdZWdfXfartzTse	7 Shifts	63.82
wUKFKGAdrQNgcXRyjt	7 Shifts	74.51
YHJTyyOnGGfGfbiici	7 Shifts	82.31
gkSiEnVxRiRmtgMFfC	12 Shifts	129.56
rhshuLBIPNkepdmvol	3 Shifts	87.05

the normality assumption was violated. If normal distribution of the data could be assumed, we performed one-way repeated measures ANOVAs and confirmed sphericity using Mauchly's test. Otherwise, we used the non-parametric Friedman test. For plotting the significant differences, we used the following indicators for the significance levels: $* < .05, ** < .01, *** < .001$.

Demographics and Prior HMD Experience

In total, we collected data from $n = 16$ participants (women=13, men=3), aged from 19 to 39 years (M=24.9, SD=5.05). Only one participant was left-handed, while the rest were right-handed. 9 participants had normal and 7 had corrected to normal vision (contact lenses only). Generally, participants were infrequent users of HMDs with reported usage of: $n = 1$ once a month, $n = 4$ once a quarter, $n = 2$ once every six months, $n = 1$ once a year, and $n = 8$ less than once a year (including $n = 2$ who have never used an HMD before).

Task Performance Measurements

We looked at the overall trial measurements, consisting of completion time and selection error rate, and also analyzed the collected data for each interaction task (see Section 8.3.1) within the task sequence of a trial to investigate task-specific differences.

For calculating the selection error rate, we considered all selection interactions as errors where the selection confirmation was performed on any UI element that was not the correct one for progressing the task. We did not count blinks as error selections if the gaze was not directed at any UI element.

Overall Trial – Completion Time and Selection Error Rate

For mean trial completion time, we found no statistically significant difference between *Gaze&Pinch* (M=151, SD=58.8) and *Gaze&Blink* (M=157, SD=47.5). However, we found a significantly lower mean selection error rate for *Gaze&Pinch* (M=3.03, SD=1.76) compared to *Gaze&Blink* (M=8.70, SD=4.03) with t(15) = $-5.23, p < .001$. As shown in Figure 8.8, the mean selection error rate for *Gaze&Blink* is almost three times as high as for *Gaze&Pinch*, while their mean trial completion time is very similar.

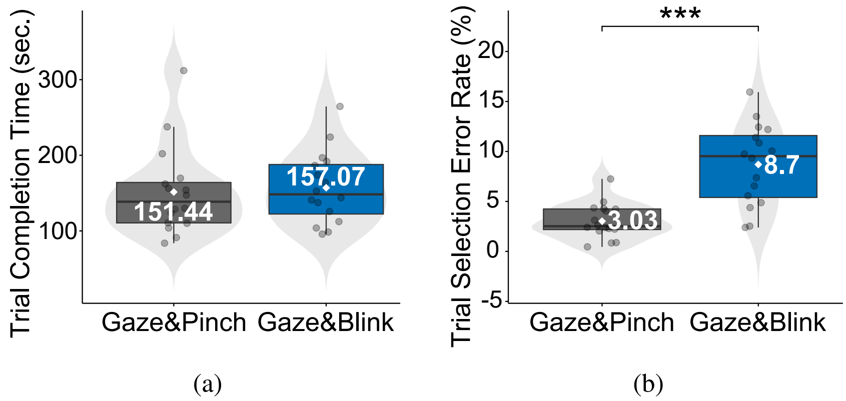


Figure 8.8: Overall trial measurements, Study 1: (a) Trial completion time (sec.) and (b) trial selection error rate (%).

Settings Menu – Scrolling

For the settings menu task, we found no significant differences between mean scrolled distance (in pixels) for *Gaze&Pinch* (M=2262, SD=710) compared to *Gaze&Blink* (M=2369, SD=623). However, the mean scroll interaction count is significantly higher ($t(15) = 2.826, p = .013$) for *Gaze&Pinch* (M=4.62, SD=2.34) compared to *Gaze&Blink* (M=2.61, SD=1.06). The scroll interaction count is defined as the amount of unique, new continuous scroll interactions that were started during a trial, which is based on each completed pinch-and-drag (returned to default state) that was performed on the settings UI. Although the mean scrolled distance in pixels is similar, the mean scroll interaction count for *Gaze&Pinch* is almost twice as high compared to scrolling with *Gaze&Blink*, as visualized in Figure 8.9. Moreover, we found a significant difference for mean selection error count for *Gaze&Pinch* (M=2.78, SD=1.85) and *Gaze&Blink* (M=9.09, SD=4.07) with $t(15) = -5.72, p < .001$. As shown in Figure 8.9, *Gaze&Blink* has a mean selection error rate that is three times higher than *Gaze&Pinch*.

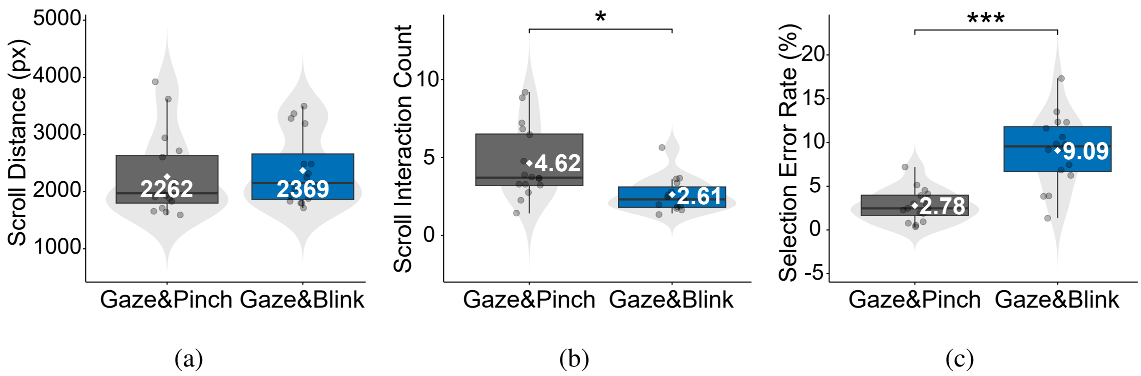
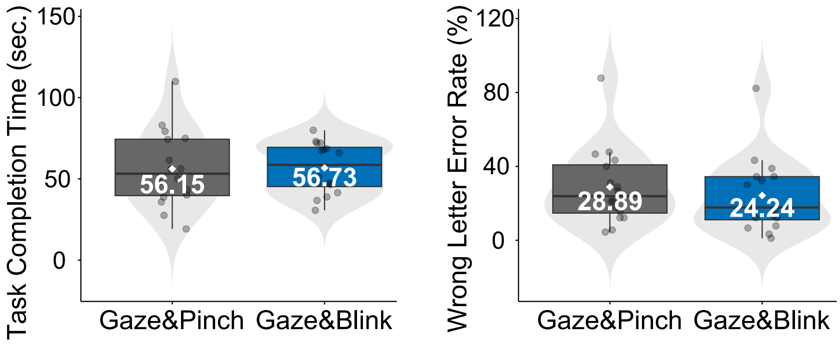


Figure 8.9: Settings menu task measurements, Study 1: (a) Scroll distance (px), (b) scroll interaction count, and (c) selection error rate (%).

Wi-Fi Password – Text Input

As illustrated in Figure 8.10, the mean keyboard input task completion times almost the same (*Gaze&Pinch* M=56.15, SD=23.5; *Gaze&Blink* M=56.73, SD=15.7) and the mean wrong letter selection error rates are very close to each other (*Gaze&Pinch* M=28.89, SD=21.0; *Gaze&Blink* M=24.24, SD=20.5) for both conditions. Our statistical testing revealed no significant differences between our two conditions for text input task completion time and wrong letter selection error rate. The wrong letter selection error rate is calculated based on the amount of wrong letter selections in a given password sequence.

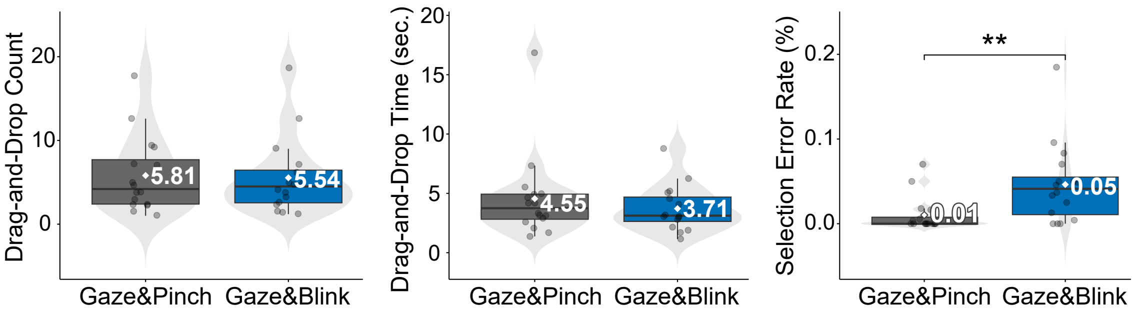


(a) (b)

Figure 8.10: Wi-Fi password keyboard input task measurements, Study 1: (a) Task completion time (sec.) for keyboard input and (b) wrong letter error rate (%).

Messenger – Drag-and-Drop

We found no significant differences for mean drag-and-drop interaction count (*Gaze&Pinch* M=5.81, SD=4.57; *Gaze&Blink* M=5.54, SD=4.6) and mean drag-and-drop interaction time (*Gaze&Pinch* M=4.55, SD=3.63; *Gaze&Blink* M=3.71, SD=1.93). As shown in Figure 8.11, both measurements are similar to each other in terms of means and distribution of the data. The selection error rate was very low ($< 1\%$) for both techniques. Still, *Gaze&Pinch* (M=0.01, SD=0.02) has a significantly lower mean error rate compared to *Gaze&Blink* (M=0.05, SD=0.05) with t(15) = $-3.99, p = .0012$.



(a) (b) (c)

Figure 8.11: Messenger menu drag-and-drop task measurements, Study 1: (a) Drag-and-drop count, (b) drag-and-drop time (sec.), and (c) selection error rate (%).

Simulator Sickness Questionnaire

The mean and SD of the SSQ subscales and the total score are listed in Table 8.5. We did not find any significant differences for the subscale *Oculomotor Disturbance* and the *Total Simulator Sickness*. The Friedman tests indicated no significant differences for the subscales *Nausea* and *Oculomotor*, but significant differences for *Disorientation* ($\chi^2(2, n = 16) = 16.24, p < .001$). The post hoc Wilcoxon signed-rank tests, corrected with holm-bonferroni, revealed that *Disorientation* was significantly lower in the measured baseline than for both study block 1 ($p = .011$) and study block 2 ($p = .012$). This increase is also clearly visible in Figure 8.12.

Table 8.5: SSQ scores mean and SD, Study 1.

	Nausea	Oculomotor	Disorientation	Total SSQ
Prestudy Baseline	8.94 ± 10.1	10.4 ± 14.9	6.09 ± 21.4	95.2 ± 125
Study Block 1	12.5 ± 13.8	23.2 ± 21.2	28.7 ± 24.1	241.0 ± 198.0
Study Block 2	13.1 ± 10.4	33.6 ± 21.1	33.1 ± 29.1	299.0 ± 205.0

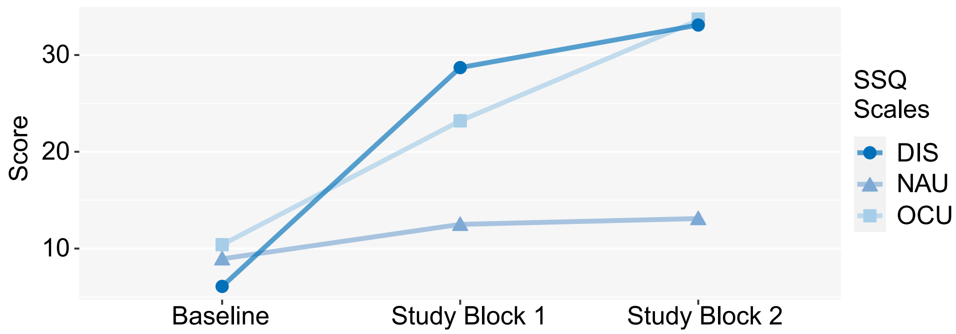


Figure 8.12: Line graph with mean SSQ subscale scores, Study 1.

NASA RAW TLX

We did not find any significant differences between the two interaction techniques for all subscales and the normalized total workload of the NASA-TLX questionnaire. The mean and SD for each scale are listed for comparison in Table 8.6.

System Usability Scale

Gaze&Pinch (M=63.4, SD=16.8) and *Gaze&Blink* (M=63.6, SD=14.7) received very similar SUS ratings with means slightly below the average SUS score of 68 [LS18]. We found no statistically significant differences in SUS scores when comparing both conditions.

User Experience Questionnaire

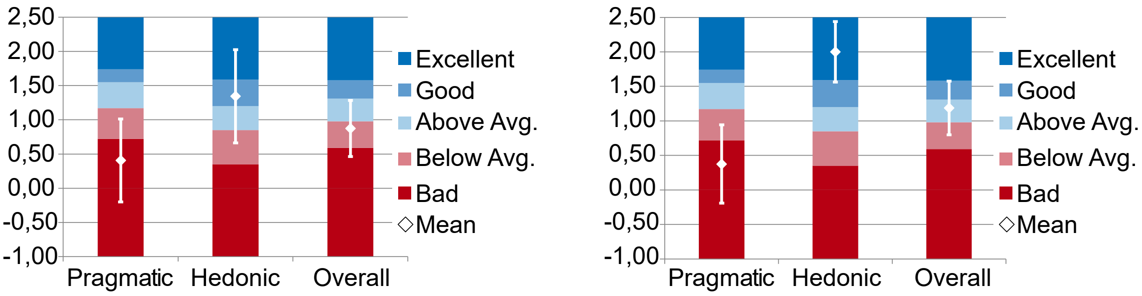
We analyzed the results with the official UEQ-S analysis tool[10]. As depicted in the UEQ-S benchmarks in Figure 8.13, pragmatic quality is rated as *bad* for both *Gaze&Pinch* (M=0.406,

[10] https://www.ueq-online.org

Table 8.6: NASA RAW TLX mean and SD, Study 1.

	Gaze&Pinch	**Gaze&Blink**
Mental Demand	25.9 ± 22.7	28.8 ± 24.1
Physical Demand	24.4 ± 28.0	26.9 ± 25.0
Temporal Demand	23.4 ± 21.5	21.2 ± 23.6
Performance	42.2 ± 25.0	33.1 ± 21.6
Effort	37.2 ± 27.4	40.0 ± 24.3
Frustration	31.9 ± 27.5	38.4 ± 31.6
Total Workload	31.9 ± 20.0	30.3 ± 18.5

SD=1.238) and *Gaze&Blink* (M=0.375, SD=1.158). The hedonic quality is rated as *good* for *Gaze&Pinch* (M=1.344, SD=1.390) and *excellent* for *Gaze&Blink* (M=2.000, SD=0.894). This results in an overall rating of *below average* for *Gaze&Pinch* (M=0.875, SD=0.838) and *above average* for *Gaze&Blink* (M=1.188, SD=0.795). We also performed a statistical analysis of the overall UEQ-S rating, which did not reveal any significant differences between the two techniques.



(a) (b)

Figure 8.13: UEQ-S benchmark graphs, Study 1: (a) *Gaze&Pinch* and (b) *Gaze&Blink*.

Qualitative User Feedback

We followed an approach based on exploratory qualitative content analysis to evaluate the user feedback [Arm17]. Two researchers sorted, grouped, and categorized the user feedback. The grouped statements were then paraphrased. It is important to note that the results of this qualitative analysis only provide insights about explicitly stated opinions and do not reflect the complete perspectives and opinions of the participants. 7 participants stated that they preferred *Gaze&Pinch* while 9 stated that they preferred *Gaze&Blink*. Generally, participants had contrasting opinions regarding performance and UX for both interaction techniques. Some participants found that *Gaze&Pinch* was faster, caused fewer errors, and was easier to use, and they felt that *Gaze&Blink* was more error-prone. However, others were of the opposite opinion. Regarding the UX, some participants reported that *Gaze&Pinch* was more familiar to them. There was also some criticism, including the hand not being tracked, complicated hand-eye coordination, and tense fingers from continuous pinching. *Gaze&Blink* was described as more natural, pleasant, and intuitive. Other participants did not agree and instead expressed that *Gaze&Blink* was unintuitive, difficult to learn, and more exhausting. An overview of the qualitative user feedback is given in Table 8.7.

Table 8.7: Overview of the qualitative user feedback, Study 1.

	Gaze&Pinch	**Gaze&Blink**
Performance	+ Less errors, more control faster, easier – More error-prone	+ Less errors, more accurate, more consistent, easier to use, fast – Error prone, extremely unreliable
UX	+ More familiar – Hand often outside tracked area, hand-eye coordination complicated, fingers tense afterward, cramping, frustrating	+ More natural, pleasant, interesting, more intuitive – Not intuitive, difficult to learn, more exhausting

8.3.3 Summary

The trial completion time of *Gaze&Blink* does not seem to be negatively impacted by its higher error rate compared to *Gaze&Pinch* (hypothesis H1a). This is also reflected in the performance dimension of the NASA-TLX. The settings menu and drag-and-drop task also show higher error rates for *Gaze&Blink*, while the interaction time does not significantly differ from *Gaze&Pinch*. Therefore, we can only partially confirm H1a. We suspect that involuntary blinks might have caused the higher error rates in *Gaze&Blink*. For more details on the classification of voluntary and involuntary blinks, refer to our preprint [Tim+25]. Paired with insights from the qualitative feedback, we speculate that performing multiple consecutive blink interactions may be faster than using pinch gestures, where participants reported hand tracking area limitations and finger strain. These issues could have had a negative impact on the trial completion and interaction times for *Gaze&Pinch*. As a result, it was not significantly faster than the more error-prone *Gaze&Blink* in our study. Since the keyboard input times and error rates are comparable, the participants seemed to type equally well using both techniques, *Gaze&Blink* and *Gaze&Pinch*.

There were no differences in the perceived workload for both interaction techniques (H2a) and in the UX (H3a). While the SUS and UEQ-S scores were not significantly different, the UEQ-S benchmarks (see Figure 8.13) revealed that the hedonic quality is rated higher for *Gaze&Blink* than for *Gaze&Pinch*. The SSQ scores for *Nausea* and *Oculomotor Disturbance* did not increase significantly during the study. Only *Disorientation* increased significantly while participants used the HMD compared to the pre-study baseline. This could be attributed to the several high ratings for *Fullness of Head*, potentially caused by the fit and relatively high weight of the HMD.

In conclusion, *Gaze&Blink* seems to be a viable alternative to *Gaze&Pinch* in terms of interaction speed, overall perceived workload, and UX. Our evaluation revealed that this applies to both discrete and continuous *Gaze&Blink* interactions. Another notable finding is that the task completion times were not significantly higher for *Gaze&Blink* compared to *Gaze&Pinch*, although the selection error rates for *Gaze&Blink* were significantly higher. We assumed that *Gaze&Blink* might compensate for more errors through faster interaction times. This finding motivated our following attempt to reduce the high error rates for *Gaze&Blink*, which would increase the reliability and effectiveness of the interaction technique. The qualitative feedback was highly varied, with opposing statements regarding perceived errors, speed, and ease of use for both techniques. Different interaction strategies, previous experience with spatial interaction, and individual preferences could have led to differing UX, resulting in contrasting feedback. Moreover, possible hand- and eye-tracking issues could have negatively impacted the UX.

Interestingly, *Gaze&Pinch* has a higher scroll interaction count than *Gaze&Blink*, which means that the total number of scroll interactions during task number 3 (*Scroll to Wi-Fi settings*, see Figure 8.6) was higher when participants used *Gaze&Pinch*. This suggests that for *Gaze&Pinch*, participants tend to perform several smaller pinch-scroll gestures instead of larger hand or arm movements. Fewer head rotations for scrolling are used for *Gaze&Blink*, but both interaction techniques have similar scroll distances. This indicates that rotational movement mapping for scrolling with *Gaze&Blink* does not lead to movement mapping issues or increase interaction difficulty, and achieves comparable performance as *Gaze&Pinch*.

8.4 User Study 2 – Gaze&BlinkPlus

Our findings in Study 1 revealed that participants performed more accidental interactions when using *Gaze&Blink* compared to *Gaze&Pinch*. As involuntary blinks occur quite frequently at around 17 blinks per minute, this causes accidental interactions with *Gaze&Blink* [Tim+25]. To mitigate this problem, we propose *Gaze&BlinkPlus*, an extension to *Gaze&Blink* that aims to reduce accidental blinks by training a deep-learning model using recorded eye-tracking data for filtering out accidental blinks. This model is then used in real-time eye-tracking during the study to decide whether a performed blink was voluntary. To the best of our knowledge, at the time of conducting our research, blink classification has not been used in the context of spatial interaction design for blink-based interaction techniques. Although there is existing work on blink classification [Abe+13; Sat+15; Sat+17], they rely on eye images, which can be a privacy concern. In contrast, our approach directly classifies blinks using the raw gaze signal along with other eye markers, such as the gaze position and pupil size. As explaining the technical details of the deep-learning model for blink-classification would go beyond the scope of this dissertation, please refer to our preprint *A Hands-free Spatial Selection and Interaction Technique using Gaze and Blink Input with Blink Prediction for Extended Reality* [Tim+25]. It contains further information on the methodology and the model training process, including the full set of classification features, the setup of the deep-learning architecture, and details on the model layers. The preprint also includes a description of the additional data collection study for obtaining training data for the deep-learning model.

8.4.1 Methodology

Study Design and Tasks

We followed the previous study procedure for Study 1, as described in Section 8.3.1 and included *Gaze&BlinkPlus* as a third condition. We also used the same study task sequence as described in Section 8.3.1. For Study 2, we formulated the following hypotheses to investigate the task performance, perceived workload, and UX between the three conditions *Gaze&Pinch*, *Gaze&Blink*, and *Gaze&BlinkPlus*:

H1b Task Performance, measured as task completion time and error rate, differs significantly between the three conditions.

H2b Perceived workload differs significantly between the three conditions.

H3b UX differs significantly between the three conditions.

Technical Implementation and Assumptions

We used the same implementation and experimental setup as in Study 1, but added *Gaze&BlinkPlus* with voluntary and involuntary blink classification as a third condition (cf. Section 8.3.1). For communication with the deep-learning model, we also implemented a client-server model that sends the data of the eye-tracker from the Unity application to the server over TCP. Our Unity application continuously listens for results from the server for a classified voluntary blink. The detailed description of the implementation of the client-server communication and the model prediction on the server can be found in our preprint [Tim+25].

Study Procedure and Participants

The recruiting followed the same procedure as in Study 1 (see Section 8.3.1). However, participants could choose to be either compensated with study credits or money. We also excluded participants who were part of the data collection study for obtaining data to train the deep-learning model. Our a priori power analysis indicated a minimum sample size of $n = 12$ to detect large effects ($\eta_p^2 = .4$) with 80% power for an ANOVA (repeated measures, within factors) with an alpha level of .05. To match the number of participants in Study 1, we performed our evaluation with $n = 17$ participants, resulting in a slightly lower effect size of $\eta_p^2 = 0.33$.

For Study 2, we used the same procedure, the same number of repetitions with 5 trials per condition, and the same questionnaires as in Study 1. We used a balanced Latin-square design for determining the order of our three conditions *Ganze+Pinch*, *Ganze+Blink*, and *Ganze+BlinkPlus* for each participant. The study took approximately 90 minutes to complete.

8.4.2 Results

For the statistical analysis of the collected data, we followed the same procedure for normality testing as in Study 1 and chose the appropriate parametric or non-parametric statistical tests and post hoc tests accordingly.

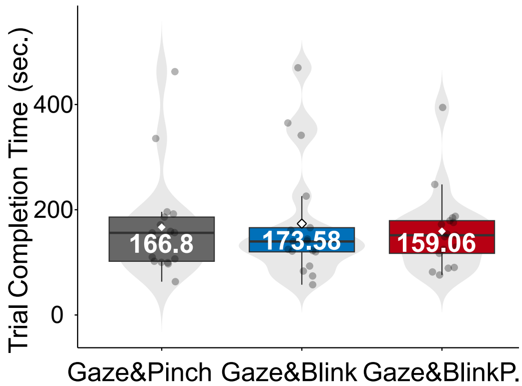
Demographics and Prior HMD Experience

We collected data from 17 participants (women=8, men=9), aged between 20 and 61 years (M=28.71, SD=8.73). Among them, only 2 participants were left-handed, while the others were right-handed. 9 participants had normal vision and 8 corrected to normal vision (contact lenses only). Most participants were infrequent users of HMDs with reported usage of: $n = 3$ more than once per week, $n = 1$ once a month, $n = 2$ once every quarter, $n = 3$ once every six months, $n = 4$ once a year, and $n = 4$ less than once a year.

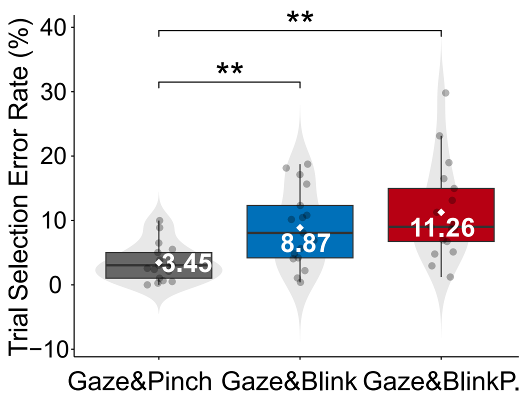
Task Performance Measurements

Overall Trial – Completion Time and Selection Error Rate

The trial completion time was log-10-transformed to fulfill the normality assumption. A repeated measures ANOVA did not indicate a significant difference. The boxplots in Figure 8.14 also show that the trial completion times for all three conditions are very similar. For the overall error rate, a repeated measures ANOVA revealed a significant effect of the selection technique ($F(2,32) = 15.20, p < .001, \eta_p^2 = .487$). Pairwise comparisons with Bonferroni-adjusted p-values showed significant differences between *Gaze&Pinch* and both *Gaze&Blink* ($p = .002$) and *Gaze&BlinkPlus* ($p = .001$), but not between the two blink interaction techniques, as pictured in Figure 8.14.



(a)



(b)

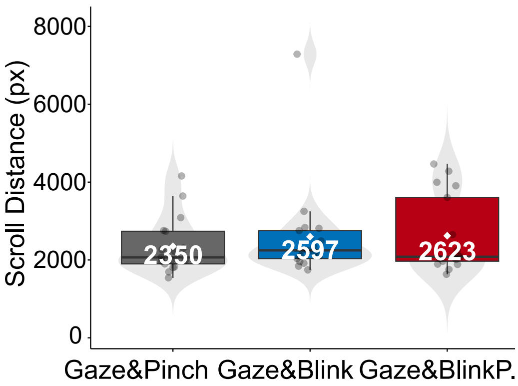
Figure 8.14: Overall trial measurements, Study 2: (a) Trial completion time (sec.) and trial selection error rate (%).

Settings Menu – Scrolling

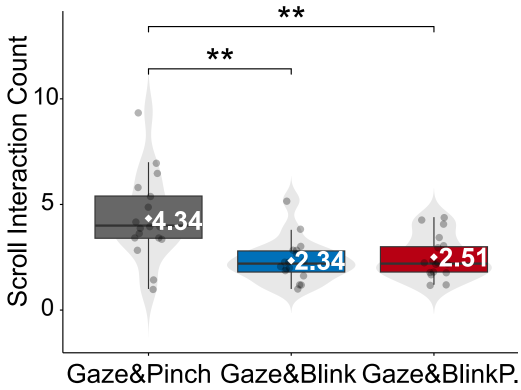
The scroll distance was not significantly different between all three interaction techniques according to a Friedman test ($\chi^2(2) = 5.77, p = .056$). Figure 8.15 shows that all three conditions have similar mean scroll distances. However, *Gaze&Pinch* has a visually larger upper box compared to the other two conditions, meaning that the data is positively skewed. If we look at the individual data points, we see that several participants had a scroll distance about twice as high as the others.

The scroll interaction count and selection error rate were both sqrt-transformed. Scroll interaction count was additionally corrected using Greenhouse-Geisser because the Mauchly's test indicated a violation of the sphericity assumption. A repeated measures ANOVA indicated a significant effect of condition on scroll count ($F(1.458, 23.334) = 12.99, p < .001, \eta_p^2 = .448$). Post hoc tests with Bonferroni correction revealed significant differences for: *Gaze&Pinch–Gaze&Blink* ($p = .003$) and *Gaze&Pinch–Gaze&BlinkPlus* ($p = .005$). Figure 8.15 shows that the mean scroll interaction count for *Gaze&Pinch* is almost twice as high compared to the other two conditions.

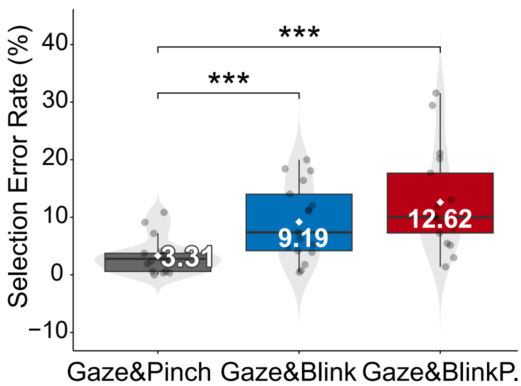
The statistical test results also indicated that the selection error rate for the scrolling task is significantly affected by the condition ($F(2, 32) = 22.33, p < .001, \eta_p^2 = .583$). Again, significant differences were found between *Gaze&Pinch–Gaze&Blink* ($p = .002$), as well as between *Gaze&Pinch—Gaze&BlinkPlus* ($p < .001$) (see Figure 8.15).



(a)



(b)



(c)

Figure 8.15: Settings menu task measurements, Study 2: (a) Scroll distance (px), (b) scroll interaction count, and (c) selection error rate (%).

Wi-Fi Password – Text Input

The task completion times for the keyboard input of the passwords and the wrong letter selection error rate were transformed using sqrt and log-10, respectively. Repeated measures ANOVAs did not reveal any significant differences for both measurements. As Figure 8.16 shows, all three conditions had comparable mean values for both measurements. *Gaze&BlinkPlus* has a visually slightly higher task completion time in the boxplot compared to the other two conditions, but the lowest error rate with the least spread in the data.

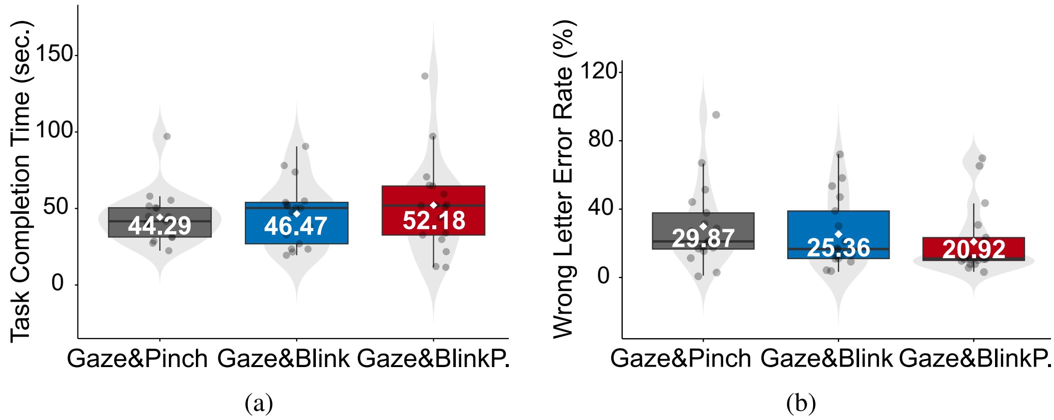


Figure 8.16: Wi-Fi password keyboard input task measurements, Study 2: (a) Task completion time (sec.) for keyboard input and (b) wrong letter error rate (%).

Messenger – Drag-and-Drop

For the drag-and-drop task, no significant differences were found for all three performance measures: interaction count, interaction time, and selection error rate. Figure 8.17 visualizes the similar mean values between all three conditions and measurements. *Gaze&BlinkPlus* has a slightly wider spread for drag-and-drop time, while *Gaze&Blink* has a slightly wider spread for wrong letter error rate.

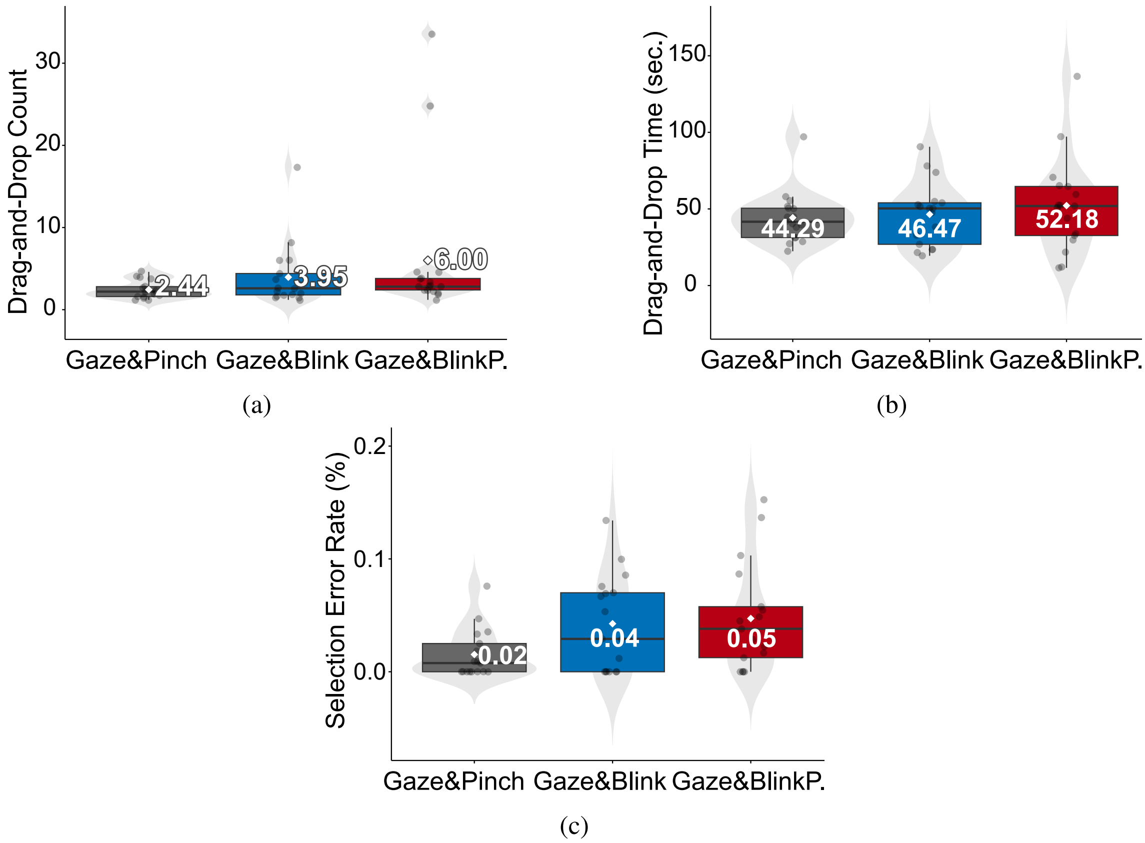


Figure 8.17: Messenger menu drag-and-drop task measurements, Study 2: (a) Drag-and-drop count, (b) drag-and-drop time (sec.), and (c) selection error rate (%).

Simulator Sickness Questionnaire

We used the same approach as in Study 1 (see Section 8.3.2) and did not find any significant differences for the three subscales and the total SSQ score. Compared to the baseline. a visual increase in the subscales during the study can be observed in Figure 8.18, especially for *Disorientation* and *Oculomotor*. The mean and SD for each subscale and the total score are listed in Table 8.8.

Table 8.8: SSQ scores mean and SD, Study 2.

	Nausea	Oculomotor	Disorientation	Total SSQ
Prestudy Baseline	9.50 ± 12.2	20.1 ± 21.8	13.1 ± 27.2	160.0 ± 214.0
Study Block 1	20.2 ± 27.4	29.9 ± 38.4	29.5 ± 58.6	298.0 ± 453.0
Study Block 2	22.4 ± 31.4	38.8 ± 42.1	33.6 ± 63.6	355.0 ± 495.0
Study Block 3	23.0 ± 34.4	38.8 ± 43.0	35.2 ± 68.0	363.0 ± 521.0

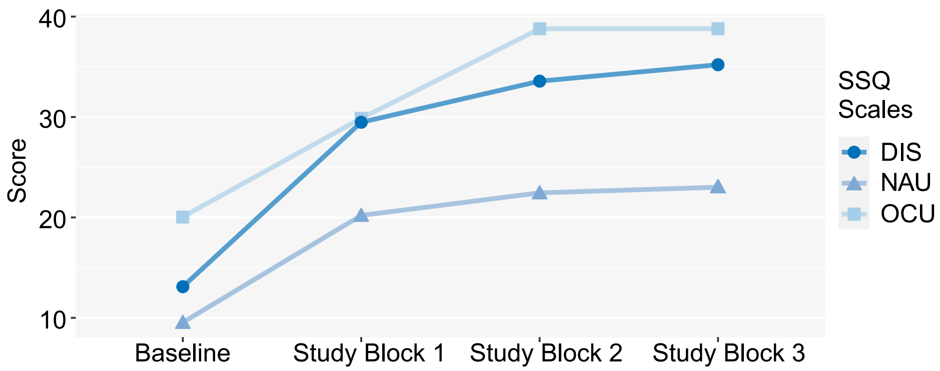


Figure 8.18: Line graph with mean SSQ subscale scores, Study 2.

NASA RAW TLX

Same as in Study 1 (see Section 8.3.2), we did not find any significant differences between the three interaction techniques for all NASA-TLX subscales and the normalized total workload. For comparison. we listed the mean and SD for each NASA-TLX scale in Table 8.9.

Table 8.9: NASA RAW TLX mean and SD, Study 2.

	Gaze&Pinch	Gaze&Blink	Gaze&BlinkPlus
Mental Demand	44.7 ± 29.2	41.8 ± 26.0	50.0 ± 29.3
Physical Demand	37.4 ± 26.3	40.3 ± 28.1	49.1 ± 30.4
Temporal Demand	42.9 ± 34.0	35.6 ± 32.8	42.6 ± 34.6
Performance	30.3 ± 22.1	27.9 ± 26.0	31.8 ± 25.9
Effort	54.7 ± 30.9	50.0 ± 30.9	62.6 ± 25.0
Frustration	48.8 ± 36.5	38.2 ± 30.2	49.7 ± 32.1
Total Workload	41.1 ± 23.2	39.0 ± 21.8	47.6 ± 20.8

System Usability Scale

The mean SUS scores for *Gaze&Pinch* (M=61.6, SD=16.6), *Gaze&Blink* (M=56.9, SD=17.8), and *Gaze&BlinkPlus* (M=55.4, SD=17.0) are similar in each condition. All mean scores are below the average SUS score of 68 [LS18]. The test results did not indicate any statistically significant differences for the SUS scores between the conditions.

User Experience Questionnaire

We analyzed the results with the UEQ-S analysis tool. As visualized in the benchmark graphs in Figure 8.19, the pragmatic quality is rated as *bad* for all three conditions, *Gaze&Pinch* (M=0.30, SD=0.994), *Gaze&Blink* (M=0.250, SD=1.275), and *Gaze&BlinkPlus* (M=0.132, SD=1.125). The hedonic quality is rated *below average* for *Gaze&Pinch* (M=0.779, SD=1.208), *good* for *Gaze&Blink* (M=1.309, SD=0.990), and *above average* for *Gaze&BlinkPlus* (0.985, SD=1.191). This results in an overall rating of *bad* for *Gaze&Pinch* (M=0.544, SD=0.725), *below average* for *Gaze&Blink* (M=0.779, SD= 0.911), and *below average* for *Gaze&BlinkPlus* (M=0.559, SD=0.921). We did not find any significant differences between the three techniques based on the statistical analysis of the overall UEQ-S rating.

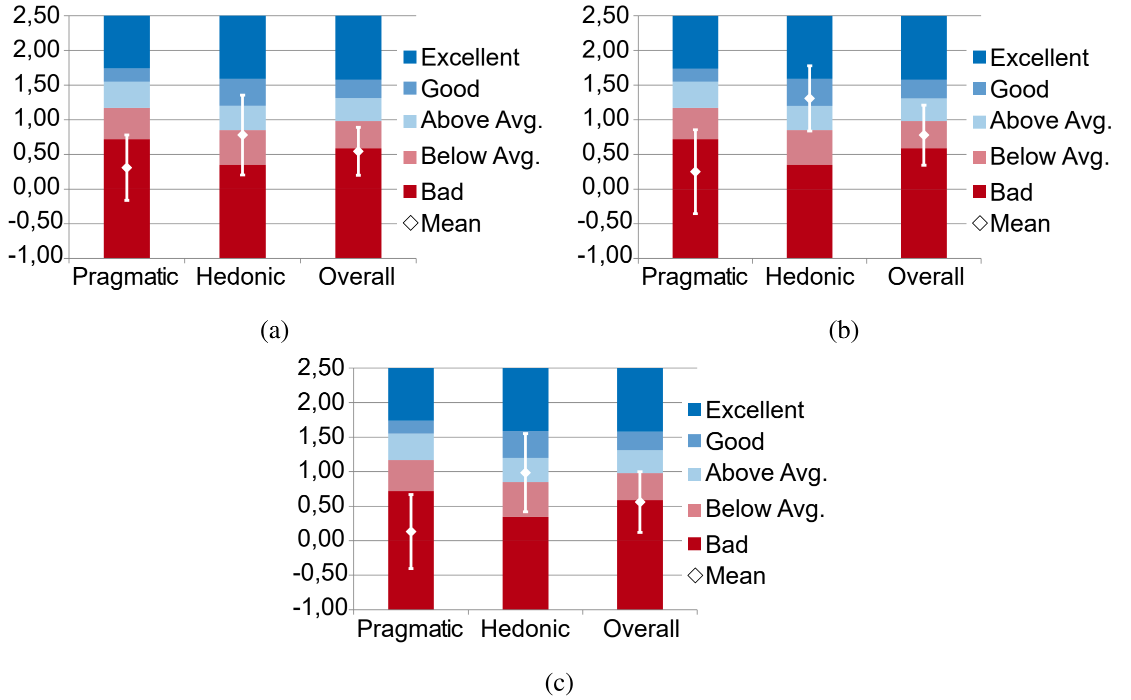


Figure 8.19: UEQ-S benchmark graphs, Study 2: (a) *Gaze&Pinch*, (b) *Gaze&Blink*, and (c) *Gaze&BlinkPlus*.

Qualitative User Feedback

Following the same approach as in Study 1 (see Section 8.3.2), we performed an exploratory analysis of the statements by sorting, categorizing, and subsequently paraphrasing them. Again, the results only contain explicitly expressed opinions and do not represent the comprehensive thoughts and opinions of the participants. 7 participants stated a preference for *Gaze&Pinch*, while 8 favored one or both blink-based techniques (*Gaze&Blink* $n = 6$, *Gaze&BlinkPlus* $n = 3$). 2 participants said that they liked both interaction techniques, with one preferring *Gaze&Pinch* and one *Gaze&BlinkPlus*. 2 other participants reported that they perceived no difference between *Gaze&Blink* and *Gaze&BlinkPlus*. One participant expressed that they would have liked a mixture of blink-based and pinch-based interactions.

Similar to Study 1, participants had opposing views regarding the interaction techniques. *Gaze&Pinch* was described as easier and faster, but some also stated that it was more error-prone and slow. Some participants expressed that *Gaze&Pinch* is more pleasant, but others found it annoying, more tedious, and physically demanding. In contrast, both blink interaction techniques were also described as more pleasant and even more intuitive. However, they were also criticized as inconsistent, more tedious, and annoying. An overview of the reported categorized and paraphrased statements is listed in Table 8.10.

Table 8.10: Overview of qualitative user feedback, Study 2.

	Gaze&Pinch	**Gaze&Blink / Gaze&Blink Plus**
Performance	+ Easier, faster – More error-prone, slow	+ Easier – More error-prone
UX	+ More pleasant – More eye strain, annoying, unusual, more tedious physically demanding	+ More pleasant, more intuitive – Inconsistent, more tedious, annoying

8.4.3 Summary

The results of Study 2 confirmed our assumption regarding blink interactions. Although *Gaze&Blink* and *Gaze&BlinkPlus* have higher error rates per trial compared to *Gaze&Pinch*, this did not seem to negatively impact the trial completion time (hypothesis H1b). This is consistent with our findings in Study 1 and also reflected in the NASA-TXL performance ratings that did not significantly differ from each other between the conditions.

In accordance with Study 1, we found that the settings menu had higher error rates for both blink-based techniques compared to *Gaze&Pinch*. We could not find any significant differences between *Gaze&Pinch* and our blink-based techniques for the drag-and-drop task (see Figure 8.17). Therefore, we can only partially confirm our hypothesis H1b. For *Gaze&Blink* and *Gaze&BlinkPlus*, we suspect that participants might have performed a selection of a menu entry immediately before starting the scrolling due to the nature of our proposed blink-based interaction techniques and the design of the settings menu. We speculate that participants might have looked and blinked at one of the menu entries in the list of settings, thereby making a selection, before performing the scrolling task. This would lead to a rise in unwanted but not involuntary selections for the task. However, we recommend conducting further research on this issue and testing other types of *UI*s menus and scrolling tasks. In line with Study 1, we found that *Gaze&Blink* and *Gaze&BlinkPlus* significantly improve the scroll interaction count, making them suitable hands-free alternatives to *Gaze&Pinch* for performing scrolling tasks.

In the qualitative feedback, participants repeated some similar opinions and sentiments as reported by participants in Study 1. This especially includes the reported tracking area limitations and finger strain. These issues could have negatively impacted the trial completion and interaction times for *Gaze&Pinch*. Participants seemed to type equally well using *Gaze&Blink*, *Gaze&BlinkPlus*, and *Gaze&Pinch*, since all interaction techniques have comparable keyboard input times and error rates. Moreover, there were no significant differences in perceived workload (hypothesis H2b). We also did not find any significant differences between the three interaction techniques based on the SUS questionnaire and UEQ-S ratings. Therefore, we have to reject our hypothesis H3b. However, in the UEQ-S benchmarks (see Figure 8.19), we can visually see that *Gaze&Blink* received a higher rating for the hedonic quality and thus, an overall higher rating than *Gaze&Pinch* and *Gaze&BlinkPlus*. Regarding the UEQ-S ratings, we have to note that our within-participants design may have indirectly added a bias. Because we presented both *Gaze&Blink* and *Gaze&BlinkPlus* to the participants, they might have given the seemingly identical conditions lower ratings for the hedonic quality items in the questionnaire.

Although the SSQ scores increased during the study, which is expected when using HMDs, we found no significant differences between the scores in each study block compared to the measured pre-study baseline simulator sickness score. This indicates that participating in our study and using the interaction techniques did not cause a significant increase in simulator sickness for the participants.

In conclusion, based on our findings, we believe that *Gaze&BlinkPlus* and *Gaze&Blink* can be viable hands-free alternatives to *Gaze&Pinch* in terms of speed and overall perceived workload and UX. However, the high error rates during the scroll task warrant a more in-depth analysis. While *Gaze&BlinkPlus* resulted in the lowest number of errors during the keyboard interaction, we think that this could be further improved with additional, more extensive data collection and a different deep-learning architecture. The qualitative feedback was also highly varied, with opposing statements for each interaction technique. There is no clear preference for one technique based on the relatively evenly distributed preferences for blink-based interaction versus *Gaze&Pinch*, as stated in Section 8.4.2. Similar to Study 1, individual differences, preferences, and experiences, as well as tracking issues, might have negatively impacted the UX and led to the contrasting user feedback for the interaction techniques.

8.5 General Discussion

Finally, we would like to discuss the collected results of our work in relation to the research questions that we formulated at the beginning of this chapter.

RQ1: How can we implement an interaction technique that covers both discrete and continuous interaction tasks through Gaze&Blink?

For answering RQ1, we designed *Gaze&Blink* to support discrete and continuous interaction tasks by taking inspiration and combining ideas from existing work. In addition, we extended *Gaze&Blink* with a deep-learning model for classification of voluntary and involuntary blinks. We thoroughly evaluated our proposed interaction techniques in two user studies. *Gaze&Blink* seems to have robust performance for discrete and continuous interaction tasks. Users had no major issues with intuitively understanding and coordinating the multi-modal input and executing the interaction state transitions (see Figure 8.1). This is also supported by our NASA RAW TLX and SUS results, which did not reveal any significant difference between *Gaze&Blink* and *Gaze&Pinch*.

RQ2: How does Gaze&Blink compare to Gaze&Pinch for different discrete and continuous UI interactions?

Reflecting on RQ2, our user study results indicate that our blink-based techniques achieve comparable performance to *Gaze&Pinch* when tested with realistic interaction tasks and a UI that is based on established operating system designs (in our case, particularly reminiscent of the UI in Apple's VisionOS). We found significant differences for some metrics of the evaluated interaction techniques, but there were also comparable measurements with no statistical differences between the conditions for other metrics. Hypothesis H1a and H1b can only be partially confirmed in terms of selection error rates. While the measured error rates in Study 2 were slightly higher for the *Gaze&Blink* and *Gaze&BlinkPlus* conditions, this had no negative impact on the trial completion time compared to *Gaze&Pinch*, as there were no significant differences between the conditions. Moreover, our results do not support our hypotheses H2a, H2b, H3a, and H3b. Our findings suggest that *Gaze&Blink* and *Gaze&BlinkPlus* generally match *Gaze&Pinch* in terms of performance, perceived workload, and UX. These results support that our proposed *Gaze&Blink* interaction technique is suitable for discrete and continuous interaction tasks.

We believe that blink-based interaction techniques can be a viable hands-free alternative and provide an additional option for facilitating spatial interactions. This might be especially useful for HMD usage in confined spaces or for users that are unable to perform a pinch gesture.

RQ3: What are the advantages and challenges of Gaze&Blink?

Based on our findings for RQ2, we can report regarding RQ3 that *Gaze&Blink* can be a viable alternative as a hands-free interaction technique. The qualitative user feedback revealed positive sentiments for using *Gaze&Blink*. It was described as intuitive, pleasant, and natural in its usage, whereas *Gaze&Pinch* was described as more physically demanding and led to finger strain. However, not all participants agreed with this. They had contrasting views and criticized that *Gaze&Blink* was more error-prone, inconsistent, and annoying to use. These findings emphasize that individual preferences have to be taken into account when designing new interaction techniques, as they influence user perception and UX. Consequently, it might be beneficial to provide users with multiple techniques for spatial interaction that they can choose from.

RQ4: Are there optimization methods to reduce involuntary blinks for Gaze&Blink that could improve this interaction technique?

We suspect that involuntary blinks are a major reason behind the negative user feedback, which leads us to RQ4. We tried to reduce accidental interactions caused by involuntary blinks with a deep-learning model. While our model achieved an accuracy of 75% on unseen data without user calibration, this still does not seem to be sufficient to effectively reduce the negative impact of involuntary blinks. Additionally, the blink classification in our *Gaze&BlinkPlus* technique may have inadvertently filtered out some voluntary interactions by participants, which might have caused some level of frustration during the interaction. Yet, we would like to point out that *Gaze&BlinkPlus*, although not statistically significant, had the lowest error rate during the selection tasks. This suggests that an improved deep-learning model that is capable of reliably filtering out involuntary blinks could further enhance the overall interaction experience.

8.5.1 Limitations and Future Work

Complex Interactions

Since the goal of our research in this chapter was to design and conduct initial empirical evaluations on our novel hands-free interaction technique *Gaze&Blink*, we did not delve into more complex 3D object manipulation tasks such as scaling or rotation. However, we believe that it is important to extend our proposed technique for 3D manipulation tasks to create a more comprehensive interaction technique.

Standardized Evaluation

We tested our proposed technique with common UI menu interaction tasks, which provided us with valuable insights into the performance and UX in a realistic use case scenario. Our results revealed that task completion times were not negatively affected by higher error rates for the blink-based techniques. This leads us to the assumption that blink-based interactions are, in general, potentially faster than pinch-based interactions using hand-tracking. This hypothesis should be investigated in future studies with a standardized setup that could also include additional performance metrics such as the effective throughput.

User Calibration

Our deep-learning model reached an accuracy of 95% on the validation set and an accuracy of 75% on the test set. This means that it can be used with a 75% accuracy without extensive retraining or user calibration. Since we did not fine-tune our model for each participant, we were able to predict voluntary and involuntary blinks in our study without requiring additional time or computational effort. However, this might have considerably impacted the model's predictive performance for some participants. The resulting increase in overall error rate could have lowered the user acceptance of our proposed technique. Future work should address this issue by improving the model. We consider a predictive performance of 75% on the test set as solid basis for further model development.

Accessibility

Although our proposed interaction technique was not designed with a focus on accessibility, it is an important property for creating a user-friendly interaction technique. We think this warrants future studies and further development of our technique regarding accessibility. This should be approached carefully, and respective target user groups should be directly involved in the process.

Inability to Close One Eye

During our studies, we discovered that a small percentage of people have a *unilateral apraxia of eyelid closure* and are therefore unable to close one eye (wink). This can also affect only one of the eyes, which is interestingly often the dominant eye [LK11]. Possible causes are stated in our preprint [Tim+25], as discussing them would go beyond the scope of this dissertation. These people where excluded from participating in our studies. Since our proposed interaction technique relies on unilateral eye closure (winking) for continuous interactions, some users might find it more challenging or even impossible to use. We suggest further research on extended and alternative blink-based techniques that do not require winking for performing continuous scrolling and drag-and-drop interactions.

Further Limitations

There are further limitations regarding blinking and deep-learning that might have negatively influenced the performance of our proposed interaction technique. One of them is the so-called *blink suppression* [RT97], a mechanism where blinking suppresses visual processing and reduces visual perception. Other limitations are the high computational demands of our currently used neural network and that we only used eye-tracking features (e.g., pupil size, openness) for the blink classification. These topics and their technical details go beyond the scope of this dissertation, and thus, further information can be found in our preprint [Tim+25].

8.6 Conclusion

In this chapter, we introduced two novel gaze-based interaction techniques, *Gaze&Blink* and its extension *Gaze&BlinkPlus*, which facilitate hands-free spatial interactions in public and confined environments. Our findings suggest that our proposed interaction techniques are a viable hands-free alternative to *Gaze&Pinch* with comparable performance, perceived workload, and UX. Moreover, an advantage of *Gaze&Blink* is that it can be implemented on the same type of HMDs that currently support *Gaze&Pinch* without requiring additional tracking hardware. In the future, a combination of these techniques could support further improvements in privacy, safety, and accessibility for using HMDs and make them more accessible to a wider range of users.

Our empirical evaluation revealed that *Gaze&Blink* has a comparable overall trial completion time to *Gaze&Pinch* but exhibits a slightly higher error rate. To address this issue, we developed a novel machine learning-based blink classification system capable of distinguishing between voluntary and involuntary blinks. We used the model to extend our technique and create *Gaze&BlinkPlus*. Study 2 revealed similar performance for metrics on trial and task completion time, letter selection error rate, and interaction count for *Gaze&Blink*, *Gaze&BlinkPlus*, and *Gaze&Pinch*. While there are still some significant differences in selection error rate in favor of *Gaze&Pinch*, we also found that our techniques performed well in other metrics, such as the scroll interaction count. Qualitative feedback from participants showed no definitive preference between the interaction techniques examined. Specifically, 41% of participants expressed a preference for either *Gaze&Blink* or *Gaze&BlinkPlus*, while 47% favored *Gaze&Pinch*. Further, 12% of participants indicated that they had no preference between the blink-based techniques and *Gaze&Pinch*.

Overall, this leads us to conclude that our proposed interaction techniques represent a viable hands-free alternative to *Gaze&Pinch*. Our research offers valuable insights into gaze- and blink-based spatial interactions and highlights their potential as well as revealed limitations and challenges. We believe that our work can serve as a solid foundation for further investigations in the area of gaze- and blink-based spatial interaction techniques.

CHAPTER 9

CONCLUSION

9.1 Summary of Key Contributions and Findings

In this dissertation, we explored the design and implementation of interactive IIS and spatial interaction techniques in the context of WA, based on our overarching research question:

> **Main Research Question**
>
> How can we design interactive, immersive VR applications and spatial interactions to improve the spatial understanding of WA?

For exploring this question, we implemented two VR applications and investigated different aspects for improving spatial interaction techniques and facilitating spatial selection. Guided by five corresponding sub-research questions, we created various interactive features for our IIS and proposed several novel spatial interaction techniques in this dissertation. Through our empirical evaluations, we obtained promising results and valuable insights. In the following, we will summarize our key contributions and findings, structured based on the sub-research questions.

> **RQ-I**
>
> How can we use immersive technologies to create interactive IIS that provide new perspectives for exploring WA?

Chapter 4: In our work, we were faced with a new, unfamiliar context with unknown requirements, for which we developed interactive IIS within the scope of two cross-disciplinary research projects involving WA. Therefore, it was important to adopt a UCD approach and closely work with the researchers from the other research fields in our projects to understand their research practices and derive requirements and objectives for viewing the WA in a spatial context. We implemented various interactive features based on our identified requirements (see Chapter 4) and jointly collected ideas for potential features that could support the researchers in their study of the WA. As this is a novel application context, we followed an exploratory approach and iteratively designed and developed the interactive features. The researchers stated that the most important advantages of the IIS are the virtual visualizations of the spatial and temporal context for WA. Therefore, we focused on integrating new perspectives and virtual content that cannot be experienced in the real world. This includes the visualization of the spatial distribution and clustering of WA, spatially located information on WA, reconstructions of former structures in combination with time of day simulation to reveal light and shadows, and inaccessible overview spots. These interactive features enabled the researchers to view and examine the WA in a new way. The IIS serve as extensions to the real-world locations and the resulting insights have already contributed to raising new research questions.

RQ-II

Which spatial UI layouts and controls are suitable for reading and navigating through long texts in immersive VR applications?

Chapter 5: We designed four different spatial text panel layouts for displaying long texts in VR based on established desktop UI patterns. They consist of two discrete variants using pagination and two continuous variants using a scrollbar. Our findings indicate that the scrollbar and pagination pattern can both be applied to 3D spatial UIs. We did not find any significant differences in reading performance between our two discrete and two continuous variants for reading and interacting with long texts. However, the classic scroll pattern without any additional button controls showed poor usability due to the sensitive control of the scrollbar in VR. In contrast, the classic horizontal pagination pattern received the highest UEQ pragmatic (goal-oriented) rating. This suggests that the control mapping and usability of 2D desktop UI patterns can be negatively affected by spatial interaction techniques when transferred to 3D spatial UIs.

RQ-III

How can we implement an assistive raycasting interaction technique that redirects the raycast towards an intended target to facilitate distant selection of spatial content?

Chapter 6: We demonstrated that the concept of gain-based movement redirection can be successfully used in the context of spatial interaction. We developed a proximity- and rotational gain-based assistive Controller-Ray redirection and implemented three different visualizations for the visible, redirected raycast: (i) *RayRotation* using a straight pointer, (ii) *RayCurve* using a curved pointer, and (iii) *RayCurveRotation* that combines both approaches. The evaluation of the raycast redirection techniques revealed that they did perform significantly better compared to non-assistive classic raycasting, with faster movement time, higher effective throughput, and a lower perceived workload. Different raycast visualizations do not appear to influence the selection performance.

As we intended, the implementation of a subtle, assistive raycast redirection did not negatively affect the users' sense of agency during the interaction. However, due to individual preferences and experiences with interacting in VR, participants had different, even contrasting opinions, on the three redirected raycast visualizations. While some preferred the curved visualization, others did not like it and even found it distracting. This suggest that, besides performance metrics, individual preferences should to be considered for designing novel spatial interaction techniques.

RQ-IV
Can we integrate hand-tracking and eye-tracking with raycast redirection to improve spatial selection?

Chapter 7: We extended our previous work on assistive Controller-Ray redirection and successfully applied the concept of assistive raycast redirection to Hand-Rays. Our findings indicate that Hand-Ray redirection offers the same advantages as Controller-Ray redirection in terms of improved selection performance and lowered perceived workload. Moreover, the sense of agency was also not negatively impacted by the redirection. We created two Hand-Ray redirection techniques. (i) *RayToTarget* is the Hand-Ray equivalent to our Controller-Ray redirection. It redirects the raycast towards the target center based on a predefined redirection zone. (ii) *RayToGaze* adds gaze-assistance and redirects the Hand-Ray towards the user's gaze position. This gaze-assisted approach can be used for target disambiguation and is more flexible than *RayToTarget* since it does not require prior knowledge of the selection targets or target-based redirection zone setup in the VE. Interestingly, we found that performance for *RayToGaze* significantly decreases with increased target depth, while it remains robust for *RayToTarget*. We hypothesize that this is caused by a mismatch of gaze position and selection timing due to differences in users' visual search and targeting strategies. Some users would already look at the next target while they were still performing the pinch gesture for selection confirmation. In these cases, the gaze-assisted redirection would only be applied during the targeting phase but would no longer be active for the selection confirmation. Therefore, the advantage of raycast redirection for compensating input jitter would not apply. Although Hand-Ray redirection was generally perceived as supportive, qualitative feedback suggests that individual preferences have considerable influence on VR, as participants had contrasting opinions on both Hand-Ray redirection techniques in terms of perceived difficulty, amount of assistance, and selection speed.

RQ-V
Can we develop a gaze-based hands-free spatial interaction technique for discrete and continuous interaction tasks?

Chapter 8: We created a novel hands-free spatial interaction technique for discrete and continuous spatial interaction tasks by combining gaze for target indication with blinking, winking, and head-rotation. Our comprehensive user studies confirmed that *Gaze&Blink* and its extension

Gaze&BlinkPlus are viable alternatives to *Gaze&Pinch* with comparable speed, perceived workload, and UX. There was also no clear subjective preference between *Gaze&Pinch* and our two interaction techniques, suggesting that participants were comfortable using all three interaction techniques. Our user studies with a realistic spatial UI and common menu interaction tasks demonstrated that selection confirmation, scrolling, and drag-and-drop, as well as frequent consecutive interactions like keyboard input, can be successfully performed with our hands-free technique.

Our findings revealed that *Gaze&Blink* has a comparable overall trial completion time to *Gaze&Pinch* but exhibits a slightly higher error rate. Hence, we hypothesized that *Gaze&Blink* interactions could actually be faster than *Gaze&Pinch*, since the higher number of errors, which required additional user interactions for completing each interaction task, did not negatively affect the task completion time for *Gaze&Blink*. Moreover, we assumed that the higher error rate might have been caused by involuntary blinking during *Gaze&Blink* interaction. While we attempted to reduce involuntary blink inputs by applying deep-learning for blink prediction to *Gaze&BlinkPlus*, the blink classification should be further improved to increase the efficiency of our technique. Nevertheless, our proposed interaction techniques demonstrated comparable performance to *Gaze&Pinch* and could be implemented as alternative, hands-free interaction techniques on the same HMDs.

To summarize, our work in this dissertation investigated how we can design interactive IIS and improve spatial interactions for supporting spatial understanding for the study of WA. Based on this overarching main research question, we derived sub-research questions and proposed different solutions for reading long texts and facilitating spatial selection of small and distant targets in VEs. This dissertation is a first exploration that lays the initial groundwork for using immersive technologies in the novel context of WA. We developed various interactive features that provide advantages and enable new perspectives for virtually accessing the WA with spatial and temporal context. Our user studies offer valuable insights and showed promising results for assisting users with spatial interactions. We demonstrated that different desktop UI patterns can be applied to VEs for creating spatial UIs. Moreover, we also validated the viability and performance of our concepts for assistive redirection and hands-free interactions. These results provide relevant implications for the design and practical application of our proposed techniques. Our work serves as a starting point for establishing the concept of assistive spatial interaction techniques and proposes potential directions for future work on using redirection in spatial interaction design. Our evaluations revealed limitations and challenges for spatial interactions that should be addressed in the future.

While we primarily developed and evaluated our IIS and spatial interaction techniques on VR HMDs within the scope of this dissertation, we created several AR and MR prototypes. In addition, we implemented and tested our hands-free interaction technique with a MR HMD. Our proposed spatial interaction techniques are not limited to a specific HMD or platform, as the same techniques can generally be used on any XR platform or in any VE that supports spatial interactions.

The cross-disciplinary context of integrating immersive technologies for facilitating spatial understanding of WA in the humanities presents specific spatial interaction requirements and challenges due to the extensive amount of research data and the spatial distribution of the WA in the IIS. This opens up numerous possibilities for further research on 3D visualizations, interactive features, and spatial interaction techniques.

9.2 Technical Contributions

This section describes our technical contributions to spatial interaction techniques and interactive features for the IIS. We also introduce a pilot project for an inscribed space, which demonstrates that our concepts and interactive features can be applied to other, novel use cases involving WA.

9.2.1 Improved Spatial Interaction Techniques

In this dissertation, we proposed and evaluated improvements to spatial interaction techniques based on the requirements of our two applications for IIS. Based on our findings, we integrated several improvements into our VR applications for RFB05 and RFB02.

Spatial UI for Reading Long Texts

Our findings in Chapter 5 indicate that established desktop UI patterns can be applied to spatial UI. Therefore, we continued using them for our spatial information panels (see Section 4.3.3) to ensure that the interface would be familiar to new users without VR experience. Our study found that users have different preferences in terms of using buttons or the scrollbar for navigating in long texts. To accommodate individual preferences, both UI control options should be provided in the spatial interface. Since we have to display a particularly large amount of text inside the spatial information panels for RFB05, we revised their design and layout based on these conclusions. We implemented a combination of scrollbar and scroll-buttons (corresponding to panel variant C.BUT in Section 5.2) as UI controls for the long description texts, as shown in Figure 9.1(a).

(a) (b)

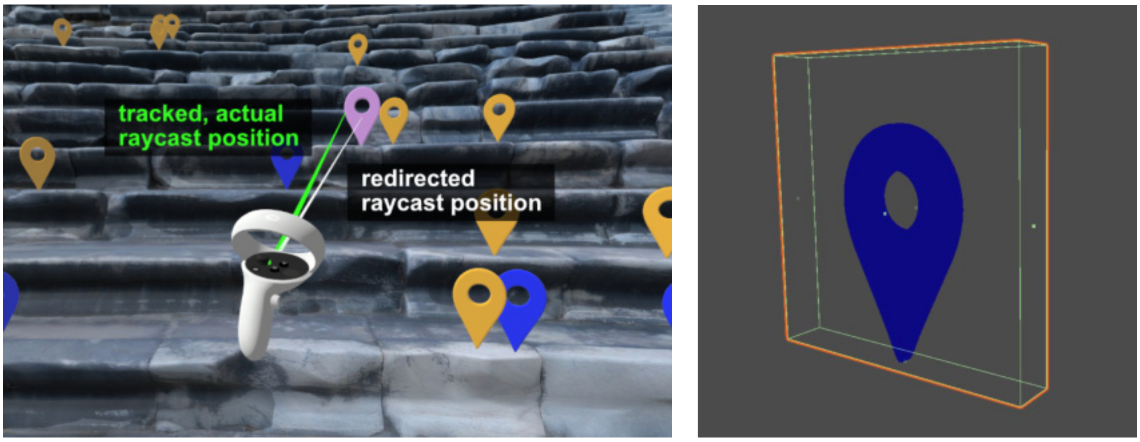
Figure 9.1: Spatial information panel for RFB05: (a) Default tabbed view mode and (b) alternative side-by-side view mode.

Since the information, including images and photos, resulted in panels that were quite wide and large, which caused too much overlap with the environment and with other spatial panels, we made them adaptive based on the feedback from the humanities researchers. As pictured in Figure 9.1(b), the panel is set up in a default, compact tabbed mode in portrait format that contains three pages with different content in each tab. In order to see and compare the entire content of a spatial information panel, they can be switched to an alternative, side-by-side view. In this expanded mode, the three pages are arranged as a curved display in landscape format.

We informally tested the updated spatial information panels with the researchers, who liked the controls and the flexibility of the revised adaptive panel layout and setup.

Redirected Raycast Selection for Written Artefacts

With our work in Chapter 6 we demonstrated that the concept of gain-based, assistive redirection can be applied to spatial interaction and that this facilitates raycast selection. Since different raycast visualizations had no effect on the selection performance, while the curved raycast visualizations were distracting for some users, we decided to use the simple, straight raycast pointer (corresponding to the *RayRotation* technique in Section 6.2). We added raycast redirection into our VR application for RFB02 for selecting the spatial WA markers (see Section 4.3.2), as pictured in Figure 9.2(a). This large-scale outdoor setting with scattered and clustered WA makes raycasting-based selection of the WA markers especially challenging. For integrating the raycast redirection, we modified the MRTK pointer prefab and setup redirection zones around the 3D pins by adding additional invisible GameObjects with box colliders, as shown in Figure 9.2(b). The redirection zone is defined by a box collider that is roughly double the size of the 3D pins. We also increased the maximum raycasting distance from 10m (MRTK default setting) to 60m. This covers nearly half of the width of the theater (130m) and reduces the need for frequent repositioning and teleporting for the selection of multiple WA markers.



(a) (b)

Figure 9.2: (a) Comparison of tracked raycast position (green pointer, hidden from user) and the redirected raycast shown to the user (white pointer) and (b) 3D pin with redirection zone.

9.2.2 Support for Understanding Written Artefacts

The interactive features that we developed for our IIS provide different advantages that support researchers in their work. As listed in Table 9.1, they can be grouped into four main aspects:

Table 9.1: Interactive features for supporting research on Written Artefacts

	Interactive Features
Spatial Context	The spatial representation and placement of the WA within the true-to-scale digitized 3D models of the IS and the WA filtering functions visualize their spatial distribution and clustering into different groups and themes.
Restoring Visibility	Non-invasive restoration provides important temporal context through digital 3D reconstructions of no longer preserved structures and blended overlays that visualize hidden former layers.
Unique Perspectives	Various overview spots offer new and unique perspectives that are inaccessible or impossible to see at the real-world locations.
Simulate Conditions	The simulation of a day-night-cycle in combination with 3D reconstructions of former structures can reveal the original visibility and shading of WA throughout the day. Filling the space with neutral agents can help with visualizing the former seating capacity.

9.2.3 Novel Use Case – Immersive Jewish Cemetery of Altona

The concept of IIS including, the developed features, interaction techniques, insights, and learnings from this dissertation, can be applied to other IS. One example is the research of WA at the Jewish Cemetery of Altona in Hamburg, Germany.

Background

The Jewish Cemetery of Altona[1] was actively used between 1616-1868. As the oldest Jewish cemetery in Hamburg (Germany) and the oldest Jewish-Portugues cemetery in Northern Europe, it is an important heritage site and testament of early modern Jewish culture in Europe. The cemetery is divided into two sections, Sephardic and Ashkenazic. It houses a large collection of tombstones from Sephardic and Ashkenazi Jews with multilingual inscriptions in Hebrew, Portuguese, and German, as well as a wide range of iconographic depictions. In total, there are around 7,000 completely or partially preserved tombstones.

This setting at the cemetery has new, unique properties compared to the two IS in our research projects. Although the cemetery is a large-scale outdoor space like RFB02 Miletus, it has completely different spatial properties, including the layout, presentation, quantity, and distribution of the WA.

[1] https://www.jüdischer-friedhof-altona.de

(a)

(b)

Figure 9.3: Different types of tombstones in the Sephardic and Ashkenazic sections of the cemetery: (a) Sephardic tombstones and (b) Ashkenazic tombstones.

As depicted in Figure 9.3, the tombstones of the Sephardic Jews are either horizontal, flat-lying stone slabs with a rectangular shape or tent-shaped stones. Some of them have partially sunken into the ground or are covered by vegetation. In contrast, most traditional Ashkenazic tombstones are vertical stone slabs with a rounded or pointed top. They are mostly densely placed next to each other and arranged in narrow rows (see Figure 9.3). This leads to difficult access to parts of the WA. Combined with the large area (2 hectares) of the cemetery and the vast amount of WA, this makes on-site research more challenging. Since this IS is a cemetery, another concern is the ethical on-site exploration and research of the WA. Furthermore, the tombstones are exposed to environmental influences and weathering-related decay.

Potential of creating an Immersive Inscribed Space

Based on our work for this dissertation, we believe that this IS would also greatly benefit from a virtual research environment with appropriate data visualizations and interactive tools. An IIS could digitally preserve the tombstones and provide unrestricted, virtual access to the Jewish Cemetery of Altona. Comparable to our two existing projects, the digitized tombstones could be visualized with spatial context and different temporal layers. In addition, existing and new research data could be integrated and interactively explored and investigated within the VE. Research data could include, for example, transcriptions of inscriptions, typologies of iconographic features, biographical and bibliographical data, photographs showing an earlier, better-preserved state of tombstones, or material analysis results. These data visualizations in the IIS could provide additional perspectives and support the on-site work of the researchers.

Pilot Project

We carried out an initial pilot project with researchers from the fields of Jewish studies, epigraphy, mineralogy, and HCI, and with the support of Michael Studemund-Halévy and the Stiftung Demkmalpflege Hamburg (foundation for the care and preservation of historical monuments)[2].

[2] https://denkmalstiftung.de/

(a) (b)

Figure 9.4: Example view of exported 3D model with mesh surface and texture: (a) Detailed surface structure and (b) high-resolution, realistic texture.

In addition, we worked with the same external 3D scanning providers that supported us in RFB05 in Lucklum (see Section 3.1.2). With our experience from digitizing the two previous IS, we were able to quickly specify our requirements for 3D scanning and processing small areas and selections of tombstones at the Jewish Cemetery of Altona. For the pilot, we choose to digitize tombstones that are well preserved and easily accessible. These include detailed high-resolution scans of a group of four tombstones and a prominent tombstone made of marble in the Sephardic section, as well as a small tombstone in the Ashkenazic section. Additionally, a small, less detailed area scan of the surroundings of the small tombstone was captured. Given the size of the area and the large number of tombstones, capturing high-resolution scans of the whole cemetery might not be feasible. Therefore, a combination of selected high-resolution scans combined with lower-resolution area scans could be a possible solution. The results of our pilot provide valuable insights about this mixed approach to help with future decision-making. After the tombstones were cleaned, they were digitized with a combination of photogrammetry and structured-light 3D scan data for the detail scans, and with LIDAR-scan data for the area scans. The data was processed with RealityCapture[3]. Figure 9.4 shows the exported 3D models with detailed, high-quality textures that are suited for use in real-time rendering with the Unity engine and on standalone XR devices.

We developed a small, initial prototype for visualizing the digitized group of tombstones with spatial WA markers and information panels in the Unity engine (see Figure 9.5). I was responsible for the technical implementation, which was done based on the developed tools in our two existing projects, while the other researchers provided research data and photographs for selected inscriptions. We used the same pin-shaped, color-coded 3D markers that we created for RFB02, as this is also an outdoor setting. The different markers indicate inscriptions in two different languages, iconographic elements, and damage to WA on the tombstones. The setup and layout for the information panel followed the ones developed for RFB02 and feature the same bill-boarding functionality. Moreover, changes in real-time lighting can be used inside the IIS to emphasize shadows and contrasts on the mesh surface for improving the readability of inscriptions,

[3]https://www.capturingreality.com/

as illustrated in Figure 9.6. The results of our pilot project demonstrate the transferability and applicability of the features that we developed. Of course, this novel use case also poses new requirements and different challenges that raise new research questions. From an HCI perspective, the dense and narrow placement of tombstones challenges the design of straightforward spatial interaction techniques that facilitate target disambiguation and selection. Since Hebrew text is written and read from right to left, this raises questions regarding the appropriate UI layout, presentation, and readability of spatial information in XR.

Our pilot project shows that other IS can utilize and build on our work for creating IIS and understanding WA in cross-disciplinary research settings. Furthermore, it demonstrates that each scenario can bring new, unique requirements and challenges that encourage further research on spatial interaction techniques.

Figure 9.5: Spatial UI markers and information panel displaying research data.

Figure 9.6: Example view of two different lighting settings for inscriptions in the 3D model: (left) Directional light from the front and (right) from the side.

9.3 Learnings and Recommendations

Based on our main findings, we provide learnings and recommendations for practical application in the development and evaluation of new spatial interaction techniques.

9.3.1 Adapting 2D UI Patterns

Established 2D desktop UI patterns can be applied to spatial interfaces. However, differences in the spatial input modalities and input mapping can affect the usability negatively. It might be necessary to adapt a 2D UI pattern to make it more suitable for use in spatial interfaces. Multiple patterns can be combined, and alternative controls should be integrated to account for individual user preferences.

9.3.2 Using Raycast Redirection

Assistive raycast redirection can be used to facilitate spatial interaction without compromising the users' sense of agency. While gaze-assisted raycast redirection supports target disambiguation and does not require prior target setup, it might not be as robust as target center-based raycast redirection with pre-defined redirection zones. Since users can employ different target acquisition and selection strategies, gaze-assisted raycast redirection does not perform equally well for all users. The appropriate redirection technique should be chosen depending on the use case and the distribution of targets inside the VE. Users should be clearly informed about the usage of assistive raycast redirection for spatial interaction in applications. They should also be provided with the option to disable and enable the redirection at any time.

9.3.3 Consider Individual Preferences

Individual preferences and differing levels of experience with spatial interaction can noticeably influence the UX and perceived behavior of the interaction. We therefore recommend offering several different options for spatial interaction, whether in terms of spatial UI controls or selection interaction techniques. This would also increase the general accessibility of the application for a wider audience.

9.3.4 Combine Quantitative and Qualitative Methods

For investigating spatial interaction techniques in a novel domain, such as in our case for WA, a user-centered approach can facilitate understanding of the target user group, the context, the unique requirements, and the interaction challenges. For gaining in-depth understanding and comprehensive insights into new or adapted spatial interaction techniques, a combination of standardized quantitative performance measurements, questionnaires, and qualitative user feedback should be used. As we stated in the previous paragraph, it is important to consider individual preferences. By applying qualitative methods and examining subjective aspects, users' pain points and preferences can be uncovered. Addressing them accordingly can aid in improving the user acceptance and UX for novel spatial interaction techniques.

9.4 Outlook

In this dissertation, we introduced the concept of IIS and implemented IIS for two research projects involving WA. For facilitating interaction within these immersive applications, we proposed and evaluated various spatial interaction techniques. The findings in our work provide insights and directions for future improvements and evaluation of assistive spatial interaction techniques. The implementation for the raycast redirection should be further investigated and refined. Moreover, the concept of assistive redirection could be integrated with other existing interaction techniques or adapted for other input modalities.

Our work contributes to novel applications of immersive technologies in the field of virtual heritage and serves as the initial groundwork for the future development and usage of IIS. The interactive features and integrated content for our IIS could be further extended for use in academic contexts and beyond for the general public. In the future, the implementation should also be extended to other immersive technologies and XR platforms. As our pilot project showed, the concept of IIS can also be applied to other IS. These spaces could even have different properties than our projects and pose new requirements and challenges for the technical implementation and the spatial interactions. In the following, we give a more detailed outlook on potential future work.

9.4.1 Assistive Multi-Modal Interaction Techniques

The results of our work demonstrated the potential of using multi-modal techniques with hand-, gaze-, and head-tracking for improving spatial interactions. Especially with more widespread XR adoption and an increasingly broader target user group, we should provide alternative interaction techniques for users with different levels of spatial interaction experience, individual preferences, and physical capabilities. Besides the performance, we need to consider the ease of use and UX of interaction techniques. Our findings in Chapter 6 and Chapter 7 indicate that subtle, assistive raycast redirection facilitates spatial selection without negatively affecting the user's sense of agency. In this context, assistive redirected raycasting could be further refined by adjusting the redirection algorithm and parameters. The concept of redirection could also be applied to other input modalities, such as head movements, for supporting continuous interaction tasks. Since blinking causes a minimal downward shift of the gaze (see Section 8.5), redirection could also be implemented to stabilize the gaze direction in gaze-blink-based interaction techniques. These approaches could be used to improve our proposed *Gaze&Blink* technique in Chapter 8. Other existing interaction techniques and concepts, such as GHA (see Section 2.5.3), could also be further extended and improved by integrating the concept of assistive redirection. Generally, future research on novel assistive interaction techniques could further facilitate spatial interaction and make XR applications more accessible to a wider audience.

9.4.2 Immersive Inscribed Spaces in Mixed Reality

With further technological advancements, IIS could be implemented for future standalone MR HMDs. While we already created some initial MR prototypes for testing out our ideas in Section 4.4, our available consumer hardware had a very limited field of view and did not provide enough performance for complex, high-quality visualizations – especially in outdoor settings with varying lighting conditions, such as in RFB02. However, blending the physical and digital worlds by placing virtual 3D content into the real physical IS could provide spatial guidance and support on-site research of WA. Our developed interactive features could be used to provide valuable additional context and information to the researchers on-site. Since the spatial WA markers are placed at their actual location in the true-to-scale 3D models, their placement could be mapped to the real IS with MR. This could assist users in finding WA at their real-world location. The filtering function could provide on-site visualization of the distribution and clustering of different types of WA. Moreover, showing virtual reconstructed structures and hidden layers could greatly help users to visualize former properties of the physical IS.

9.4.3 Multi-User And Collaborative Scenarios

The UWA Cluster brings together researchers from different fields who are involved in the research on WA. Considering that the researchers might not always be in the same physical location, future development should include multi-user features to enable virtual co-located collaboration directly inside the IIS. This could, for example, include collaborative tools for creating spatial 3D annotations and tags or a tool for defining and sharing custom overview spots in the IIS. For the technical implementation, the current VR applications would have to be extended with networking functions, and the required infrastructure for setting up and maintaining client-server networking would have to be created.

9.4.4 Science Communication

Another important aspect for research on understanding WA and an objective of the UWA cluster is science communication and knowledge transfer to the general public. As we anecdotally experienced and observed in our numerous public presentations, the immersive and interactive experience of the applications easily caught the visitors' interest. The IIS were always very positively received. Especially students and children were enthusiastic about using HMDs and exploring both spaces. We believe that IIS of cultural heritage sites can support science communication to promote interest and raise public awareness for the research on WA and the importance of cultural heritage. Since we developed the applications for two research projects, future implementations could include new features that are targeted towards non-academic users and interested users from the general public. The applications could offer an additional mode that focuses on the introduction and explanation of the IIS and its WA. This could even be further expanded to include age-appropriate educational content and explanations. With our implementation of *Miletus Touch* (see Section 4.4.3), we already integrated some of these concepts by including more introductory information and a guided tour of the WA inside the ancient theater of Miletus for RFB02. This demonstrated the potential for further implementations and interactive concepts that are tailored to science communication. With future development and additional use cases for IIS, new requirements and challenges could emerge, which will require further research in the area of spatial interaction. For this context, the ease of use of the IIS and its interactive features remain crucial. This emphasizes and confirms the relevance of the work in this dissertation for improving and advancing spatial interaction techniques.

BIBLIOGRAPHY

[W3C20] World Wide Web Consortium (W3C). *1.4.3 Contrast (Minimum)*. Website. (accessed January 10, 2023). 2020. URL: https://www.w3.org/TR/UNDERSTANDING-WCAG20/visual-audio-contrast-contrast.html (cited on page 79).

[Abe+13] Kiyohiko Abe, Hironobu Sato, Shogo Matsuno, Shoichi Ohi, and Minoru Ohyama. "Automatic Classification of Eye Blink Types Using a Frame-Splitting Method." In: *Engineering Psychology and Cognitive Ergonomics. Understanding Human Cognition*. Edited by Don Harris. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pages 117–124. ISBN: 978-3-642-39360-0. DOI: 10.1007/978-3-642-39360-0_13 (cited on page 151).

[Alg15] Mike Alger. *Visual Design Methods for Virtual Reality*. (accessed: November 4, 2024). Sept. 2015. URL: https://aperturesciencellc.com/vr/VisualDesignMethodsforVR_MikeAlger.pdf (cited on page 99).

[Ami+25] 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 2025 CHI Conference on Human Factors in Computing Systems*. (CHI). Yokohama, Japan: Association for Computing Machinery, 2025. ISBN: 9798400713941. DOI: 10.1145/3706598.3713623 (cited on pages 33, 138).

[Ann21] Daniel Meßner und Anna Priebe. "Forschen in der Ferne." In: *Neunzehn – Magazin der Universität Hamburg*. 17th edition. Mittelweg 177, 20148 Hamburg, Germany: Universität Hamburg, Oct. 2021, pages 28–32. URL: https://www.uni-hamburg.de/newsroom/presse/publikationen/19neunzehn-2-2021.pdf (cited on pages 76, 78, 217).

[AA13] Ferran Argelaguet and Carlos Andujar. "A survey of 3D object selection techniques for virtual environments." In: *Computers & Graphics* 37.3 (2013), pages 121–136. ISSN: 0097-8493. DOI: https://doi.org/10.1016/j.cag.2012.12.003 (cited on pages 3, 20, 22).

[Arg+16] Ferran Argelaguet, Ludovic Hoyet, Michael Trico, and Anatole Lecuyer. "The role of interaction in virtual embodiment: Effects of the virtual hand representation." In: *2016 IEEE Virtual Reality (VR)*. Greenville, SC, USA: IEEE, Mar. 2016, pages 3–10. DOI: 10.1109/VR.2016.7504682 (cited on pages 100, 119).

[Arm17] Andreas Armborst. "Thematic Proximity in Content Analysis." In: *Sage Open* 7.2 (2017), pages 1–11 (cited on pages 107, 128, 149).

[AK21] Afraa Z. Attiah and Enas F. Khairullah. "Eye-Blink Detection System for Virtual Keyboard." In: *2021 National Computing Colleges Conference (NCCC)*. 2021, pages 1–6. DOI: 10.1109/NCCC49330.2021.9428797 (cited on page 28).

[Azi+24] Samantha Aziz, Dillon J Lohr, Lee Friedman, and Oleg Komogortsev. "Evaluation of Eye Tracking Signal Quality for Virtual Reality Applications: A Case Study in the Meta Quest Pro." In: *Proceedings of the 2024 Symposium on Eye Tracking Research and Applications*. (ETRA). Glasgow, United Kingdom: Association for Computing Machinery, 2024. ISBN: 9798400706073. DOI: 10.1145/3649902.3653347 (cited on page 26).

[Azm+16] Mahdi Azmandian, Mark Hancock, Hrvoje Benko, Eyal Ofek, and Andrew D. Wilson. "Haptic Retargeting: Dynamic Repurposing of Passive Haptics for Enhanced Virtual Reality Experiences." In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, 2016, pages 1968–1979. ISBN: 9781450333627. DOI: 10.1145/2858036.2858226 (cited on pages 29, 30).

[Bak03] James R. Baker. "The Impact of Paging vs. Scrolling on Reading Online Text Passages." In: *Usability News*. Edited by Barbara S. Chaparro. Volume 5. (Editor Barbara S. Chaparro). Wichita State University: Software Usability Research Laboratory (SURL), 2003 (cited on pages 76, 89).

[BPC19] Marc Baloup, Thomas Pietrzak, and Géry Casiez. "RayCursor: A 3D Pointing Facilitation Technique based on Raycasting." In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. (CHI). Glasgow, Scotland Uk: Association for Computing Machinery, 2019, pages 1–12. ISBN: 9781450359702. DOI: 10.1145/3290605.3300331 (cited on pages 22, 23).

[BKM09] Aaron Bangor, Philip Kortum, and James Miller. "Determining what individual SUS scores mean: adding an adjective rating scale." In: *J. Usability Studies* 4.3 (May 2009), pages 114–123. URL: https://uxpajournal.org/determining-what-individual-sus-scores-mean-adding-an-adjective-rating-scale/ (cited on page 69).

[BGH04] Martin S. Banks, Tandra Ghose, and James M. Hillis. "Relative image size, not eye position, determines eye dominance switches." In: *Vision Research* 44.3 (2004), pages 229–234. ISSN: 0042-6989. DOI: 10.1016/j.visres.2003.09.029 (cited on page 26).

[BBM17] Loris Barbieri, Fabio Bruno, and Maurizio Muzzupappa. "User-centered design of a virtual reality exhibit for archaeological museums." In: *International Journal on Interactive Design and Manufacturing (IJIDeM)* 12 (2017), pages 561–571. DOI: 10.1007/s12008-017-0414-z (cited on page 17).

[Bar13] Angelos Barmpoutis. "Digital Epigraphy Toolbox." In: *Humanities Commons* (Feb. 2013), page 11. DOI: http://dx.doi.org/10.17613/M64W9R (cited on page 18).

[BS19a] Mayra Donaji Barrera Machuca and Wolfgang Stuerzlinger. “The Effect of Stereo Display Deficiencies on Virtual Hand Pointing.” In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. (CHI). Glasgow, Scotland Uk: Association for Computing Machinery, 2019, pages 1–14. ISBN: 9781450359702. DOI: 10.1145/3290605.3300437 (cited on page 34).

[Bat+11] Scott Bateman, Regan L. Mandryk, Tadeusz Stach, and Carl Gutwin. “Target assistance for subtly balancing competitive play.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. (CHI). Vancouver, BC, Canada: Association for Computing Machinery, 2011, pages 2355–2364. ISBN: 9781450302289. DOI: 10.1145/1978942.1979287 (cited on page 30).

[Bat+22] Anil Ufuk Batmaz, Moaaz Hudhud Mughrabi, Mayra Donaji Barrera Machuca, and Wolfgang Stuerzlinger. “Effect of Stereo Deficiencies on Virtual Distal Pointing.” In: *Proceedings of the 28th ACM Symposium on Virtual Reality Software and Technology*. (VRST). Tsukuba, Japan: Association for Computing Machinery, 2022. ISBN: 9781450398893. DOI: 10.1145/3562939.3565621 (cited on pages 33, 35).

[Bat+23a] Anil Ufuk Batmaz, Moaaz Hudhud Mughribi, Mine Sarac, Mayra Barrera Machuca, and Wolfgang Stuerzlinger. “Measuring the Effect of Stereo Deficiencies on Peripersonal Space Pointing.” In: *2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*. 2023, pages 1–11. DOI: 10.1109/VR55154.2023.00063 (cited on pages 33, 35).

[Bat+21] Anil Ufuk Batmaz, Mohammad Rajabi Seraji, Johanna Kneifel, and Wolfgang Stuerzlinger. “No Jitter Please: Effects of Rotational and Positional Jitter on 3D Mid-Air Interaction.” In: *Proceedings of the Future Technologies Conference (FTC) 2020, Volume 2*. Edited by Kohei Arai, Supriya Kapoor, and Rahul Bhatia. Cham: Springer International Publishing, 2021, pages 792–808. ISBN: 978-3-030-63089-8 (cited on pages 22, 33).

[BS19b] Anil Ufuk Batmaz and Wolfgang Stuerzlinger. “Effects of 3D Rotational Jitter and Selection Methods on 3D Pointing Tasks.” In: *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 2019, pages 1687–1692. DOI: 10.1109/VR.2019.8798038 (cited on pages 22, 33, 114).

[BS20] Anil Ufuk Batmaz and Wolfgang Stuerzlinger. “Effect of Fixed and Infinite Ray Length on Distal 3D Pointing in Virtual Reality.” In: *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. (CHI EA). <conf-loc>, <city>Honolulu</city>, <state>HI</state>, <country>USA</country>, </conf-loc>: Association for Computing Machinery, 2020, pages 1–10. ISBN: 9781450368193. DOI: 10.1145/3334480.3382796 (cited on pages 33, 35).

[BS23] Anil Ufuk Batmaz and Wolfgang Stuerzlinger. "Rotational and Positional Jitter in Virtual Reality Interaction in Everyday VR." In: *Everyday Virtual and Augmented Reality*. Edited by Adalberto Simeone, Benjamin Weyers, Svetlana Bialkova, and Robert W. Lindeman. Cham: Springer International Publishing, 2023, pages 89–118. ISBN: 978-3-031-05804-2. DOI: 10.1007/978-3-031-05804-2_4 (cited on page 22).

[Bat+23b] Anil Ufuk Batmaz, Rumeysa Turkmen, Mine Sarac, Mayra Donaji Barrera Machuca, and Wolfgang Stuerzlinger. "Effect of Grip Style on Peripersonal Target Pointing in VR Head Mounted Displays." In: *2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 2023, pages 425–433. DOI: 10.1109/ISMAR59233.2023.00057 (cited on page 33).

[Bau] Alessandro Bausi. *Beta masaheft: Manuscripts of Ethiopia and Eritrea*. Website. (Hiob Ludolf Centre for Ethiopian Studies, Universität Hamburg, accessed June 1, 2021). DOI: 10.25592/BetaMasaheft (cited on page 47).

[Bau+23] Alessandro Bausi, Dmitry Bondarev, Nadine Bregler, Sara Chiarini, Giovanni Ciotti, Janine Droese, Eliana Dal Sasso, Michael Friedrich, Agnieszka Helman-Ważny, Michael Kohs, Leah Mascia, Ann Lauren Osthof, Malena Ratzke, Ondřej Škrabal, Szilvia Sövegjártó, Thies Staack, and Hanna Wimmer. "Definition of 'Written Artefact'." (This paper is an occasional paper. Its results are based on exchanges, reflections, and discussions of the Theory and Terminology working group.). Nov. 2023. DOI: 10.25592/uhhfdm.13836 (cited on pages vi, viii, 2).

[Ber+21a] 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 2021 CHI Conference on Human Factors in Computing Systems*. (CHI). Yokohama, Japan: Association for Computing Machinery, 2021. ISBN: 9781450380966. DOI: 10.1145/3411764.3445193 (cited on pages 33, 34, 100, 118).

[Ber+21b] 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 2021 CHI Conference on Human Factors in Computing Systems*. (CHI). Yokohama, Japan: Association for Computing Machinery, 2021. ISBN: 9781450380966. DOI: 10.1145/3411764.3445193 (cited on page 142).

[BO24] Christof Berns and Ann Lauren Osthof. "Streets as contested spaces in ancient Miletus." English. In: *História da arquitetura – Perspectivas temáticas (III). A rua na estrutura urbana.* Edited by Manuel Joaquim Moreira da Rocha and Nuno Resende. Porto, Portugal, 2024, pages 131–141. ISBN: 978-989-8970-73-2. URL: https://ler.letras.up.pt/uploads/ficheiros/20075.pdf (cited on page 6).

[Bos21] Sebastian Bosch. *X-Ray Fluorescence Imaging Data for the Reconstruction of Emblematic Inscriptions in the Interior of the Church Lucklum*. Sept. 2021. DOI: 10.25592/uhhfdm.14457 (cited on page 57).

[Bow+01] Doug Bowman, Chadwick Wingrave, Joshua Campbell, and Vinh Ly. *Using pinch gloves for both natural and abstract interaction techniques in virtual environments*. Technical report. Department of Computer Science, Virginia Polytechnic Institute & State University, Jan. 2001. URL: http://hdl.handle.net/10919/20058 (cited on pages 22, 99, 114, 129).

[Bow99] Doug A. Bowman. "Interaction Techniques for Common Tasks in Immersive Virtual Environments: : Design, Evaluation, and Application." PhD thesis. Georgia Institute of Technology, 1999 (cited on pages 3, 9, 21, 22).

[BH97] Doug A. Bowman and Larry F. Hodges. "An Evaluation of Techniques for Grabbing and Manipulating Remote Objects in Immersive Virtual Environments." In: *Proceedings of the 1997 Symposium on Interactive 3D Graphics*. (I3D). Providence, Rhode Island, USA: Association for Computing Machinery, 1997, 35–ff. ISBN: 0897918843. DOI: 10.1145/253284.253301 (cited on page 20).

[BC64] George E. P. Box and David R. Cox. "An Analysis of Transformations." In: *Journal of the Royal Statistical Society: Series B (Methodological)* 26.2 (July 1964), pages 211–243. ISSN: 0035-9246. DOI: 10.1111/j.2517-6161.1964.tb00553.x (cited on page 121).

[BB17] Eleni Bozia and Angelos Barmpoutis. "Augmented Reality for Epigraphy: How to bring holograms of inscriptions to your classrooms." In: *CIEGL17: 15th International Congress of Greek and Latin Epigraphy*. Aug. 2017 (cited on page 18).

[BBW14] Eleni Bozia, Angelos Barmpoutis, and Robert S. Wagman. "Open-Access Epigraphy: Electronic Dissemination of 3D digitized Archaeological Material." In: *In Proceedings of the International Conference on Information Technologies for Epigraphy and Digital Cultural Heritage in the Ancient World (EAGLE 2014), Sapienza Università Editrice* (2014), pages 421–435. DOI: 10.13133/978-88-98533-42-8 (cited on page 18).

[BL20] Khosrow Bozorgi and Zack Lischer-Katz. "Using 3D/VR for Research and Cultural Heritage Preservation: Project Update on the Virtual Ganjali Khan Project." In: *Preservation, Digital Technology & Culture* 49.2 (2020), pages 45–57. DOI: 10.1515/pdtc-2020-0017 (cited on page 18).

[Bro+96] John Brooke et al. "SUS-A quick and dirty usability scale." In: *Usability evaluation in industry* 189.194 (1996), pages 4–7 (cited on pages 68, 143, 211).

[Bro+90] Frederick P. Brooks, Ming Ouh-Young, James J. Batter, and P. Jerome Kilpatrick. "Project GROPEHaptic displays for scientific visualization." In: *Proceedings of the 17th Annual Conference on Computer Graphics and Interactive Techniques*. (SIGGRAPH). Dallas, TX, USA: Association for Computing Machinery, 1990, pages 177–185. ISBN: 0897913442. DOI: 10.1145/97879.97899 (cited on page 1).

[Bur+24] Dipima Buragohain, Yahui Meng, Chaoqun Deng, Qirui Li, and Sushank Chaudhary. "Digitalizing cultural heritage through metaverse applications: challenges, opportunities, and strategies." In: *Heritage Science* 12.1 (Aug. 2024), pages 1–16. DOI: 10.1186/s40494-024-01403-1 (cited on page 16).

[CC] Center for the Study of Islam and Society (PPIM) and Centre for the Study of Manuscripts Culture (CSMC). *Digital Repository of Endangered and Affected Manuscripts in Southeast Asia (DREAMSEA)*. https://dreamsea.co/. URL: https://drive.google.com/file/d/1OuGvzOuPrJPJm6CRkXHO-5lkgm-Beabf/view (cited on page 47).

[Cena] Centre for the Study of Manuscript Cultures. *Cultural Heritage*. Website. (accessed January 10, 2025). URL: https://www.csmc.uni-hamburg.de/cultural-heritage.html (cited on page 16).

[Cenb] Centre for the Study of Manuscript Cultures (CSMC). *About - Understanding Written Artefacts*. Website. (accessed 5 August, 2024). URL: https://www.csmc.uni-hamburg.de/about.html (cited on page 2).

[Cenc] Centre for the Study of Manuscript Cultures (CSMC). *Research Field Inscribing Spaces*. Website. (accessed 5 August 2024). URL: https://www.csmc.uni-hamburg.de/research/cluster-projects/field-b.html (cited on pages 2, 4, 5).

[CGC13] Eugene Ch'ng, Vincent Gaffney, and Henry Chapman, editors. *Visual Heritage in the Digital Age*. London: Springer, 2013. ISBN: 978-1-4471-5534-8. DOI: 10.1007/978-1-4471-5535-5 (cited on page 16).

[CM13] Yeonjoo Cha and Rohae Myung. "Extended Fitts' law for 3D pointing tasks using 3D target arrangements." In: *International Journal of Industrial Ergonomics* 43.4 (2013), pages 350–355. ISSN: 0169-8141. DOI: https://doi.org/10.1016/j.ergon.2013.05.005 (cited on page 34).

[Che+23] Di Laura Chen, Marcello Giordano, Hrvoje Benko, Tovi Grossman, and Stephanie Santosa. "GazeRayCursor: Facilitating Virtual Reality Target Selection by Blending Gaze and Controller Raycasting." In: *Proceedings of the 29th ACM Symposium on Virtual Reality Software and Technology*. (VRST). Christchurch, New Zealand: Association for Computing Machinery, 2023. ISBN: 9798400703287. DOI: 10.1145/3611659.3615693 (cited on pages 23, 27, 116).

[CMS88] Michael Chen, S. Joy Mountford, and Abigail Sellen. "A study in interactive 3-D rotation using 2-D control devices." In: *ACM SIGGRAPH Computer Graphics* 22.4 (June 1988), pages 121–129. ISSN: 0097-8930. DOI: 10.1145/378456.378497 (cited on page 20).

[Che+17] Lung-Pan Cheng, Eyal Ofek, Christian Holz, Hrvoje Benko, and Andrew D. Wilson. "Sparse Haptic Proxy: Touch Feedback in Virtual Environments Using a General Passive Prop." In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, 2017, pages 3718–3728. ISBN: 9781450346559. DOI: 10.1145/3025453.3025753 (cited on pages 29, 30).

[Col+24] Mark Colley, Beate Wanner, Max Rädler, Marcel Rötzer, Julian Frommel, Teresa Hirzle, Pascal Jansen, and Enrico Rukzio. "Effects of a Gaze-Based 2D Platform Game on User Enjoyment, Perceived Competence, and Digital Eye Strain." In: *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. (CHI). Honolulu, HI, USA: Association for Computing Machinery, 2024. ISBN: 9798400703300. DOI: 10.1145/3613904.3641909 (cited on page 28).

[Cro57] Edward R. F. W. Crossman. *The speed and accuracy of hand movements*. ([unpublished report]). 1957 (cited on page 32).

[CSD93] Carolina Cruz-Neira, Daniel J. Sandin, and Thomas A. DeFanti. "Surround-screen projection-based virtual reality: the design and implementation of the CAVE." In: *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques*. (SIGGRAPH). Anaheim, CA: Association for Computing Machinery, 1993, pages 135–142. ISBN: 0897916018. DOI: 10.1145/166117.166134 (cited on pages 17, 62).

[Del14] Ben Delaney. *Sex, Drugs and Tessellation: The Truth About Virtual Reality, as Revealed in the Pages of CyberEdge Journal*. CreateSpace Independent Publishing Platform, 2014. ISBN: 9781500893293 (cited on page 1).

[Dil92] Andrew Dillon. "Reading from paper versus screens: a critical review of the empirical literature." In: *Ergonomics* 35.10 (1992), pages 1297–1326. DOI: 10.1080/00140139208967394 (cited on page 19).

[DKO18] Tilman Dingler, Kai Kunze, and Benjamin Outram. "VR Reading UIs: Assessing Text Parameters for Reading in VR." In: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. (CHI EA). Montreal QC, Canada: Association for Computing Machinery, 2018, pages 1–6. DOI: 10.1145/3170427.3188695 (cited on pages 19, 20, 75, 77).

[Doc] Oculus Developer Documentation. *Distance Hand Grab Interactors*. Website. (accessed August 10, 2023). URL: https://developer.oculus.com/

documentation/unity/unity-isdk-distance-hand-grab-interaction/ (cited on pages 22, 94).

[Duc17] Andrew T. Duchowski. *Eye tracking methodology: Theory and practice*. Springer, 2017. DOI: 10.1007/978-1-84628-609-4 (cited on pages 136, 138).

[Ebe16] Manuel Eberl. "Fisher–Yates shuffle." In: *Archive of Formal Proofs* (Sept. 2016). (Formal proof development, accessed January 20, 2023). ISSN: 2150-914x. URL: https://isa-afp.org/entries/Fisher_Yates.html (cited on pages 101, 120).

[Eri+20] Austin Erickson, Kangsoo Kim, Gerd Bruder, and Gregory F. Welch. "Effects of Dark Mode Graphics on Visual Acuity and Fatigue with Virtual Reality Head-Mounted Displays." In: *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. Atlanta, USA: IEEE, 2020, pages 434–442. DOI: 10.1109/VR46266.2020.00064 (cited on pages 20, 75, 77).

[EBR20] Shaghayegh Esmaeili, Brett Benda, and Eric D. Ragan. "Detection of Scaled Hand Interactions in Virtual Reality: The Effects of Motion Direction and Task Complexity." In: *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. Atlanta, GA, USA: IEEE, Mar. 2020, pages 453–462. DOI: 10.1109/VR46266.2020.00066 (cited on page 29).

[FFK13] Ribel Fares, Shaomin Fang, and Oleg Komogortsev. "Can we beat the mouse with MAGIC?" In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. (CHI). Paris, France: Association for Computing Machinery, 2013, pages 1387–1390. ISBN: 9781450318990. DOI: 10.1145/2470654.2466183 (cited on page 27).

[Fis+88] Scott S. Fischer, Elizabeth M. Wenzel, Clayton Coler, and Michael McGreevy. "Virtual Interface Environment Workstations." In: *Proceedings of the Human Factors Society Annual Meeting* 32.2 (1988). (Aerospace Human Factors Research Division NASA Ames Research Center), pages 91–95. DOI: 10.1177/154193128803200219 (cited on page 1).

[Fis+87] Scott S. Fisher, Michael McGreevy, James Humphries, and Warren Robinett. "Virtual environment display system." In: *Proceedings of the 1986 Workshop on Interactive 3D Graphics*. (I3D). Chapel Hill, North Carolina, USA: Association for Computing Machinery, 1987, pages 77–87. ISBN: 0897912284. DOI: 10.1145/319120.319127 (cited on page 1).

[Fit54] Paul M Fitts. "The information capacity of the human motor system in controlling the amplitude of movement." In: *Journal of experimental psychology* 47.6 (1954), page 381 (cited on page 31).

[FP64] Paul M Fitts and James R Peterson. “Information capacity of discrete motor responses.” In: *Journal of experimental psychology* 67.2 (1964), page 103. DOI: 10.1037/h0055392 (cited on page 31).

[Fle48] Rudolph Flesch. “A new readability yardstick.” In: *Journal of Applied Psychology* 32.3 (1948), pages 221–233. DOI: 10.1037/h0057532 (cited on page 78).

[FHZ96] Andrew Forsberg, Kenneth Herndon, and Robert Zeleznik. “Aperture Based Selection for Immersive Virtual Environments.” In: *Proceedings of the 9th Annual ACM Symposium on User Interface Software and Technology*. (UIST). Seattle, Washington, USA: Association for Computing Machinery, 1996, pages 95–96. ISBN: 0897917987. DOI: 10.1145/237091.237105 (cited on pages 10, 21, 22).

[Foua] Foundation of the Hellenic World. *A walk through ancient Miletus*. Website. (Research Project ARCHIMED 2000-2006, accessed November 6, 2024). URL: https://www.fhw.gr/choros/miletus/en/index.php (cited on page 17).

[Foub] Foundation of the Hellenic World. *Kivotos (CAVE system)*. Website. (Hellenic Cosmos Cultural Centre, accessed November 6, 2024). URL: https://www.hellenic-cosmos.gr/en/kivotos (cited on page 17).

[Fouc] Foundation of the Hellenic World. *Mediterranean harbors, ships and sea: the invisible routes*. Website. (Research Project ARCHIMED 2000-2006, accessed November 6, 2024). URL: https://harbors.ime.gr/en/ (cited on page 17).

[Foud] Foundation of the Hellenic World. *Virtual Cinema*. Website. (Hellenic Cosmos Cultural Centre, accessed November 6, 2024). URL: https://www.hellenic-cosmos.gr/en/virtual-cinema (cited on page 17).

[Foue] Foundation of the Hellenic World. *Website of the Foundation of the Hellenic World*. Website. (Foundation, accessed November 6, 2024). URL: https://www.ime.gr/fhw/ (cited on page 17).

[Gab+21a] Jenny Gabel, Christof Berns, Sebastian Bosch, Jost Eickmeyer, Kaja Harter-Uibopuu, Nathalie Martin, Ann Lauren Osthof, Johann Anselm Steiger, and Frank Steinicke. “User-Centered Design of Immersive Research Applications for Understanding Written Artefacts.” In: *2021 International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments*. (ICAT-EGVE). Sankt Augustin, Germany: The Eurographics Association, 2021, pages 31–35. ISBN: 978-3-03868-142-7. DOI: 10.2312/egve.20211325 (cited on pages 13, 47).

[Gab+22] Jenny Gabel, Christof Berns, Sebastian Bosch, Jost Eickmeyer, Kaja Harter-Uibopuu, Nathalie Martin, Ann Lauren Osthof, Johann Anselm Steiger, and Frank Steinicke. “Immersive Inscribed Spaces – Bringing Virtuality to Written Artefacts for Humanities.” In: *i-com Journal of Interactive Media* 21.1 (2022), pages 7–18. DOI: 10.1515/icom-2022-0012 (cited on pages 12, 47).

[Gab+21b] Jenny Gabel, Sukran Karaosmanoglu, Celeste Mason, Sebastian Rings, and Frank Steinicke. "Corona Beat - Kicking The Sedentary Habit Induced By Prolonged Social Distancing." In: *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops*. (VRW). (3DUI Contest). Virtual Conference, 2021, pages 687–688. DOI: 10.1109/VRW52623.2021.00226 (cited on page 14).

[GLS23] Jenny Gabel, Melanie Ludwig, and Frank Steinicke. "Immersive Reading: Comparison of Performance and User Experience for Reading Long Texts in Virtual Reality." In: *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. (CHI EA). Hamburg, Germany: Association for Computing Machinery, 2023, pages 1–8. ISBN: 9781450394222. DOI: 10.1145/3544549.3585895 (cited on pages 13, 75).

[Gab+25] Jenny Gabel, Ann Lauren Osthof, Christof Berns, Jost Eickmeyer, Kaja Harter-Uibopuu, Johann Anselm Steiger, and Frank Steinicke. "Exploring Written Artefacts in Virtual Reality – Potential and Challenges for Creating Immersive Tools and Applications." In: *2025 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops*. (VRW). Saint-Malo, France, 2025, pages 25–28. DOI: 10.1109/VRW66409.2025.00013 (cited on pages 13, 47).

[GRS23] Jenny Gabel, Sebastian Rings, and Frank Steinicke. "Ancient Theatre of Miletus: A Virtual Reality Research Tool for Studying Written Artefacts." In: *Proceedings of the 2023 ACM Symposium on Spatial User Interaction*. (SUI). Sydney, NSW, Australia: Association for Computing Machinery, 2023, pages 1–3. ISBN: 9798400702815. DOI: 10.1145/3607822.3618021 (cited on pages 13, 47).

[Gab+23] Jenny Gabel, Susanne Schmidt, Oscar Ariza, and Frank Steinicke. "Redirecting Rays: Evaluation of Assistive Raycasting Techniques in Virtual Reality." In: *Proceedings of the 29th ACM Symposium on Virtual Reality Software and Technology*. (VRST). Christchurch, New Zealand: Association for Computing Machinery, 2023, pages 1–11. ISBN: 9798400703287. DOI: 10.1145/3611659.3615716 (cited on pages 13, 93, 118, 119).

[Gab+24] Jenny Gabel, Susanne Schmidt, Ken Pfeuffer, and Frank Steinicke. "Guiding Handrays in Virtual Reality: Comparison of Gaze- and Object-Based Assistive Raycast Redirection." In: *Proceedings of the 2024 ACM Symposium on Spatial User Interaction*. (SUI). (Honorable Mention for Best Paper Award). Trier, Germany: Association for Computing Machinery, 2024, pages 1–12. ISBN: 9798400710889. DOI: 10.1145/3677386.3682080 (cited on pages 13, 113).

[Ger] German Archaeological Institute (DAI). *iDAI.world*. Website. (accessed: August 12, 2024). URL: https://idai.world/ (cited on pages 18, 47).

[Gib33] James J. Gibson. "Adaptation, after-effect and contrast in the perception of curved lines." In: *Journal of Experimental Psychology* 16.1 (Feb. 1933), pages 1–31. DOI: 10.1037/h0074626 (cited on page 30).

[Gig93] Michael A. Gigante. "1 - Virtual Reality: Definitions, History and Applications." In: *Virtual Reality Systems*. Edited by R.A. Earnshaw, M.A. Gigante, and H. Jones. Boston: Academic Press, 1993, pages 3–14. ISBN: 978-0-12-227748-1. DOI: 10.1016/B978-0-12-227748-1.50009-3 (cited on page 16).

[GL17] Mar Gonzalez-Franco and Jaron Lanier. "Model of Illusions and Virtual Reality." In: *Frontiers in Psychology* 8 (June 2017). DOI: 10.3389/fpsyg.2017.01125 (cited on page 30).

[Goo] Google Arts & Culture. *Open Heritage*. Website. (accessed May 10, 2021). URL: https://artsandculture.google.com/project/openheritage (cited on page 18).

[GM96] Evan D. Graham and Christine L. MacKenzie. "Physical versus virtual pointing." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. (CHI). Vancouver, British Columbia, Canada: Association for Computing Machinery, 1996, pages 292–299. ISBN: 0897917774. DOI: 10.1145/238386.238532 (cited on page 31).

[GWM24] Jesse W. Grootjen, Henrike Weingärtner, and Sven Mayer. "Uncovering and Addressing Blink-Related Challenges in Using Eye Tracking for Interactive Systems." In: *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. (CHI). Honolulu, HI, USA: Association for Computing Machinery, 2024. ISBN: 9798400703300. DOI: 10.1145/3613904.3642086 (cited on page 142).

[GB05] Tovi Grossman and Ravin Balakrishnan. "A probabilistic approach to modeling two-dimensional pointing." In: *ACM Transactions on Computer-Human Interaction (TOCHI)* 12.3 (Sept. 2005), pages 435–459. ISSN: 1073-0516. DOI: 10.1145/1096737.1096741 (cited on page 34).

[GB06] Tovi Grossman and Ravin Balakrishnan. "The Design and Evaluation of Selection Techniques for 3D Volumetric Displays." In: *Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology*. (UIST). Montreux, Switzerland: Association for Computing Machinery, 2006, pages 3–12. ISBN: 1595933131. DOI: 10.1145/1166253.1166257 (cited on pages 22, 23).

[Hae+16] Steffen Haesler, Karen Obernesser, Tino Raupp, Christoph Jahnke, Jonathan Stapf, Julia Bräker, Paul Lubos, Gerd Bruder, and Frank Steinicke. "Edutainment & Engagement at Exhibitions: A Case Study of Gamification in the Historic Hammaburg Model." In: *Mensch und Computer 2016 - Tagungsband*. Edited by Wolfgang Prinz, Jan Borchers, and Matthias Jarke. Aachen: Gesellschaft für Informatik e.V., 2016. DOI: 10.18420/muc2016-mci-0143 (cited on page 17).

[HSR18] Dustin T. Han, Mohamed Suhail, and Eric D. Ragan. "Evaluating Remapped Physical Reach for Hand Interactions with Passive Haptics in Virtual Reality." In: *IEEE Transactions on Visualization and Computer Graphics* 24.4 (2018), pages 1467–1476. DOI: 10.1109/TVCG.2018.2794659 (cited on page 30).

[Han+03] John Paulin Hansen, Anders Sewerin Johansen, Dan Witzner Hansen, Kenji Ito, and Satoru Mashino. "Command Without a Click: Dwell Time Typing by Mouse and Gaze Selections." In: *Human-Computer Interaction INTERACT '03: IFIP TC13 International Conference on Human-Computer Interaction*. IOS Press, 2003 (cited on pages 11, 24, 135).

[Han+18] John Paulin Hansen, Vijay Rajanna, I. Scott MacKenzie, and Per Bækgaard. "A Fitts' law study of click and dwell interaction by gaze, head and mouse with a head-mounted display." In: *Proceedings of the Workshop on Communication by Gaze Interaction*. (COGAIN). Warsaw, Poland: Association for Computing Machinery, 2018. ISBN: 9781450357906. DOI: 10.1145/3206343.3206344 (cited on pages 11, 24).

[Han92] Andrew J. Hanson. "II.3 - THE ROLLING BALL." In: *Graphics Gems III (IBM Version)*. Edited by David Kirk. San Francisco: Morgan Kaufmann, 1992, pages 51–60. ISBN: 978-0-12-409673-8. DOI: 10.1016/B978-0-08-050755-2.50023-3 (cited on page 20).

[Har06a] 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), pages 904–908. DOI: 10.1177/154193120605000909 (cited on pages 143, 210).

[Har06b] 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 (Oct. 2006), pages 904–908. DOI: 10.1177/154193120605000909 (cited on pages 100, 118).

[Har+21] Judith Hartfill, Jenny Gabel, Lucie Kruse, Susanne Schmidt, Kevin Riebandt, Simone Kühn, and Frank Steinicke. "Analysis of Detection Thresholds for Hand Redirection during Mid-Air Interactions in Virtual Reality." In: *Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology*. (VRST). (Honorable Mention for Best Paper Award). Osaka, Japan: Association for Computing Machinery, 2021, pages 1–10. ISBN: 9781450390927. DOI: 10.1145/3489849.3489866 (cited on pages 14, 29).

[Har+20] Judith Hartfill, Jenny Gabel, Daniel Neves-Coelho, Daniel Vogel, Fabian Räthel, Simon Tiede, Oscar Ariza, and Frank Steinicke. "Word saber: an effective and fun VR vocabulary learning game." In: *Proceedings of Mensch Und Computer 2020*. (MuC). (Best Paper Award). Magdeburg, Germany: Association for Computing Machinery, 2020, pages 145–154. ISBN: 9781450375405. DOI: 10.1145/3404983.3405517 (cited on page 13).

[Her98] Peter Herrmann. *Inschriften von Milet, 2, Inschriften n. 407 - 1019*. De Gruyter, 1998. ISBN: 3110150921 (cited on pages 8, 55).

[Her+22] Julia Hertel, Jenny Gabel, Lucie Kruse, Marcel Wollborn, and Frank Steinicke. "Co-Design of an Augmented Reality Maintenance Tool for Gas Pressure Regulation Stations." In: *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. (ISMARW). Singapore: Institute of Electrical and Electronics Engineers, 2022, pages 720–724. DOI: 10.1109/ISMAR-Adjunct57072.2022.00152 (cited on page 14).

[Höl25] Tonio Hölscher. *Klassische Archäologie: Grundwissen*. Volume 6. Darmstadt: Verlag Herder GmbH, Apr. 2025, pages 150–151. ISBN: 9783534610815 (cited on page 7).

[HLH21] Zhong Hong, Wang Leilei, and Zhang Heqing. "The application of virtual reality technology in the digital preservation of cultural heritage." In: *Computer Science and Information Systems 2021* 18.2 (2021), pages 535–551. DOI: 10.2298/CSIS200208009Z (cited on page 17).

[Hum] Humboldt University Berlin. *Digitales Forum Romanum*. Website. (accessed: August 12, 2024). URL: https://www.projekte.hu-berlin.de/de/digitales-forum-romanum (cited on page 17).

[Inc] Apple Inc. *Design for the iPadOS pointer - WWDC20*. Website. (accessed May 15, 2023). URL: https://developer.apple.com/videos/play/wwdc2020/10640/ (cited on page 30).

[Int] Interaction Design Foundation - IxDF. *What is Extended Reality (XR)?* Website. (accessed: January 10, 2025). URL: https://www.interaction-design.org/literature/topics/extended-reality-xr (cited on page 15).

[Int00] International Organization for Standardization. *ISO 9241-9 Ergonomic requirements for office work with visual display terminals (VDTs) — Part 9: Requirements for non-keyboard input devices*. Website. (accessed August 10, 2023). 2000. URL: https://www.iso.org/standard/30030.html (cited on page 33).

[Int12] International Organization for Standardization. *ISO/TS 9241-411 Ergonomics of human-system interaction — Part 411: Evaluation methods for the design of physical input devices*. Website. (accessed August 10, 2023). 2012. URL: https://www.iso.org/standard/54106.html (cited on page 33).

[Isl18] Hans Peter Isler. *Antike Theaterbauten: Ein Handbuch*. Austrian Academy of Sciences Press, 2018. ISBN: 370017957X (cited on page 61).

[Jer15a] Jason Jerald. "Interaction Patterns and Techniques." In: *The VR Book: Human-Centered Design for Virtual Reality*. Association for Computing Machinery and Morgan & Claypool, 2015. Chapter 28, pages 323–350. DOI: 10.1145/2792790.2792824 (cited on pages 45, 54).

[Jer15b] Jason Jerald. *The VR Book: Human-Centered Design for Virtual Reality*. Association for Computing Machinery and Morgan & Claypool, 2015. ISBN: 9781970001129. DOI: 10.1145/2792790 (cited on pages 1–3, 20, 21, 48, 76).

[JW15] Ricardo Jota and Daniel Wigdor. "Palpebrae superioris: exploring the design space of eyelid gestures." In: *Proceedings of the 41st Graphics Interface Conference*. (GI). Halifax, Nova Scotia, Canada: Canadian Information Processing Society, 2015, pages 273–280. ISBN: 9780994786807. URL: https://dl.acm.org/doi/10.5555/2788890.2788938 (cited on page 28).

[Kar+24] Sukran Karaosmanoglu, Sebastian Cmentowski, Lennart E. Nacke, and Frank Steinicke. "Born to Run, Programmed to Play: Mapping the Extended Reality Exergames Landscape." In: *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. (CHI). Honolulu, HI, USA: Association for Computing Machinery, 2024. ISBN: 9798400703300. DOI: 10.1145/3613904.3642124 (cited on page 15).

[Kei24] Nadja Keilich. "Design, Implementierung und Empirische Evaluierung einer mobilen Webanwendung zur Lokalisierung und Darstellung von Inschriften in Hamburg." Bachelor's Thesis. University of Hamburg, Jan. 2024 (cited on page 65).

[Ken+93] 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), pages 203–220. DOI: 10.1207/s15327108ijap0303_3 (cited on pages 78, 210).

[Kir+23] Johannes Kirchner, Tamara Watson, Jochen Bauer, and Markus Lappe. "Eyeball translations affect saccadic eye movements beyond brainstem control." In: *Journal of Neurophysiology* 130.5 (2023), pages 1334–1343. DOI: 10.1152/jn.00021.2023 (cited on page 142).

[KKS21] Seina Kobayashi, Kei Kanari, and Mie Sato. "An Examination of View-Settings for Long Texts in VR Reading." In: *ACM SIGGRAPH 2021 Posters*. (SIGGRAPH). Virtual Event, USA: Association for Computing Machinery, 2021. ISBN: 9781450383714. DOI: 10.1145/3450618.3469164 (cited on pages 19, 20, 75, 79).

[KWB12] Luv Kohli, Mary C. Whitton, and Frederick P. Brooks. "Redirected touching: The effect of warping space on task performance." In: *2012 IEEE Symposium on 3D User Interfaces (3DUI)*. Costa Mesa, CA, USA: IEEE, 2012, pages 105–112. DOI: 10.1109/3DUI.2012.6184193 (cited on pages 29, 30).

[Koj+22] Tanja Kojic, Maurizio Vergari, Sebastian Möller, and Jan-Niklas Voigt-Antons. "Assessing User Experience of Text Readability with Eye Tracking in Virtual Reality." In: *Virtual, Augmented and Mixed Reality: Design and Development*. Berlin, Heidelberg: Springer-Verlag, 2022, pages 199–211. ISBN: 978-3-031-05938-4. DOI: 10.1007/978-3-031-05939-1_13 (cited on page 19).

[Koj+20] Tanja Kojić, Danish Ali, Robert Greinacher, Sebastian Möller, and Jan-Niklas Voigt-Antons. “User Experience of Reading in Virtual Reality — Finding Values for Text Distance, Size and Contrast.” In: *2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)*. Athlone, Ireland: IEEE, 2020, pages 1–6. DOI: 10.1109/QoMEX48832.2020.9123091 (cited on pages 19, 75, 79).

[Kop+10] Regis Kopper, Doug A. Bowman, Mara G. Silva, and Ryan P. McMahan. “A human motor behavior model for distal pointing tasks.” In: *Int. J. Hum.-Comput. Stud.* 68.10 (Oct. 2010), pages 603–615. ISSN: 1071-5819. DOI: 10.1016/j.ijhcs.2010.05.001 (cited on pages 22, 34).

[KS18] Piotr Kowalczyk and Dariusz Sawicki. “Blink and wink detection as a control tool in multimodal interaction.” In: *Multimedia Tools and Applications* 78.10 (Aug. 2018), pages 13749–13765. DOI: 10.1007/s11042-018-6554-8 (cited on page 11).

[Kra73] Friedrich Krauss. *Milet. Das Theater von Milet / Das hellenistische Theater: der römische Zuschauerbau*. 1st edition. Volume 4. De Gruyter, 1973. ISBN: 311004000X (cited on page 61).

[KGH85] Myron W. Krueger, Thomas Gionfriddo, and Katrin Hinrichsen. “VIDEOPLACE—an artificial reality.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. (CHI). San Francisco, California, USA: Association for Computing Machinery, 1985, pages 35–40. ISBN: 0897911490. DOI: 10.1145/317456.317463 (cited on page 1).

[Kum+08] Manu Kumar, Jeff Klingner, Rohan Puranik, Terry Winograd, and Andreas Paepcke. “Improving the accuracy of gaze input for interaction.” In: *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications*. (ETRA). Savannah, Georgia: Association for Computing Machinery, 2008, pages 65–68. ISBN: 9781595939821. DOI: 10.1145/1344471.1344488 (cited on pages 26, 131).

[Kyr+20] Christos Kyrlitsias, Maria Christofi, Despina Michael-Grigoriou, Domna Banakou, and Andri Ioannou. “A virtual tour of a hardly accessible archaeological site: The effect of immersive virtual reality on user experience, learning and attitude change.” In: *Frontiers in Computer Science* 2 (Aug. 2020). DOI: 10.3389/fcomp.2020.00023 (cited on page 17).

[Kyt+18] Mikko Kytö, Barrett Ens, Thammathip Piumsomboon, Gun A. Lee, and Mark Billinghurst. “Pinpointing: Precise Head- and Eye-Based Target Selection for Augmented Reality.” In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. (CHI). Montreal QC, Canada: Association for Computing Machinery, 2018, pages 1–14. ISBN: 9781450356206. DOI: 10.1145/3173574.3173655 (cited on page 116).

[LK11] Daniel M Laby and David G Kirschen. “Thoughts on ocular dominance—Is it actually a preference?” In: *Eye & contact lens* 37.3 (2011), pages 140–144. DOI: 10.1097/ICL.0b013e31820e0bdf (cited on page 162).

[Lan+17] Eike Langbehn, Paul Lubos, Gerd Bruder, and Frank Steinicke. “Application of redirected walking in room-scale VR.” In: *2017 IEEE Virtual Reality (VR)*. Los Angeles, CA, USA: IEEE, 2017, pages 449–450. DOI: 10.1109/VR.2017.7892373 (cited on page 29).

[Lan+19] Eike Langbehn, Joel Wittig, Nikolaos Katzakis, and Frank Steinicke. “Turn Your Head Half Round: VR Rotation Techniques for Situations With Physically Limited Turning Angle.” In: *Proceedings of Mensch Und Computer 2019*. (MuC). Hamburg, Germany: Association for Computing Machinery, 2019, pages 235–243. ISBN: 9781450371988. DOI: 10.1145/3340764.3340778 (cited on page 29).

[Lan01] Martin Langner. *Antike Graffitizeichnungen: Motive, Gestaltung und Bedeutung*. Dr Ludwig Reichert, 2001. ISBN: 9783895001888 (cited on pages 8, 55).

[Lan00] Chris Lankford. “Effective eye-gaze input into Windows.” In: *Proceedings of the 2000 Symposium on Eye Tracking Research & Applications*. (ETRA). Palm Beach Gardens, Florida, USA: Association for Computing Machinery, 2000, pages 23–27. ISBN: 1581132808. DOI: 10.1145/355017.355021 (cited on page 11).

[LHS08] Bettina Laugwitz, Theo Held, and Martin Schrepp. “Construction and Evaluation of a User Experience Questionnaire.” In: *HCI and Usability for Education and Work*. Edited by Andreas Holzinger. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pages 63–76. ISBN: 978-3-540-89350-9. DOI: 10.1007/978-3-540-89350-9_6 (cited on pages 78, 215).

[LaV+17] Joseph J. LaViola, Ernst Kruijff, Ryan P. McMahan, Doug A. Bowman, and Ivan Poupyrev. *3D User Interfaces - Theory and Practice*. en. 2nd. Redwood City, CAUnited States: Addison-Wesley, 2017. ISBN: 978-0-134-03432-4 (cited on pages 3, 10, 20–22, 54, 93, 114).

[Lea] Learning Sites, Inc. *The Acropolis, Athens, Greece: Digital Re-creation for Hypothesis Testing, Educational Videos, and Interactive Learning*. Website. (3D Rendering, accessed November 6, 2024). URL: https://www.learningsites.com/Acropolis/Acropolis_home_2016.php (cited on page 17).

[LS09] James R. Lewis and Jeff Sauro. “The Factor Structure of the System Usability Scale.” In: *Human Centered Design*. Edited by Masaaki Kurosu. Springer Berlin Heidelberg, 2009, pages 94–103. ISBN: 978-3-642-02806-9 (cited on page 68).

[LS18] James R. Lewis and Jeff Sauro. “Item benchmarks for the system usability scale.” In: *J. Usability Studies* 13.3 (May 2018), pages 158–167 (cited on pages 148, 157).

[LG94] Jiandong Liang and Mark Green. “JDCAD: A highly interactive 3D modeling system.” In: *Computers &Graphics* 18.4 (1994), pages 499–506. ISSN: 0097-8493. DOI: 10.1016/0097-8493(94)90062-0 (cited on pages 10, 21, 22).

[Liu+09] Lei Liu, Robert van Liere, Catharina Nieuwenhuizen, and Jean-Bernard Martens. “Comparing Aimed Movements in the Real World and in Virtual Reality.” In: *2009 IEEE Virtual Reality Conference*. 2009, pages 219–222. DOI: 10.1109/VR.2009.4811026 (cited on page 34).

[Liu19] Pietro M. Liuzzo. “EAGLE Continued: IDEA. The International Digital Epigraphy Association.” In: *Crossing Experiences in Digital Epigraphy*. Edited by Annamaria De Santis and Irene Rossi. Warsaw, Poland: De Gruyter Open Poland, 2019. Chapter 17, pages 216–230. DOI: 10.1515/9783110607208-018 (cited on page 47).

[LWW23] Xiaolong Lou, Ying Wu, and Yigang Wang. “The Optimal Interactive Space for Hand Controller Interaction in Virtual Reality.” In: *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 2023, pages 839–840. DOI: 10.1109/VRW58643.2023.00262 (cited on page 35).

[Lu+20] Xueshi Lu, Difeng Yu, Hai-Ning Liang, Wenge Xu, Yuzheng Chen, Xiang Li, and Khalad Hasan. “Exploration of Hands-free Text Entry Techniques For Virtual Reality.” In: *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 2020, pages 344–349. DOI: 10.1109/ISMAR50242.2020.00061 (cited on pages 28, 139).

[Luh+20] Thomas Luhmann, Stuart Robson, Stephen Kyle, and Jan Boehm. *Close-Range Photogrammetry and 3D Imaging*. Berlin, Boston: De Gruyter, 2020. ISBN: 9783110607253. DOI: 10.1515/9783110607253 (cited on page 37).

[Lys+24] Mathias N. Lystbæk, Thorbjørn Mikkelsen, Roland Krisztandl, Eric J Gonzalez, Mar Gonzalez-Franco, Hans Gellersen, and Ken Pfeuffer. “Hands-on, Hands-off: Gaze-Assisted Bimanual 3D Interaction.” In: *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*. (UIST). Pittsburgh, PA, USA: Association for Computing Machinery, 2024. ISBN: 9798400706288. DOI: 10.1145/3654777.3676331 (cited on page 25).

[Lys+22] Mathias N. Lystbæk, Peter Rosenberg, Ken Pfeuffer, Jens Emil Grønbæk, and Hans Gellersen. “Gaze-Hand Alignment: Combining Eye Gaze and Mid-Air Pointing for Interacting with Menus in Augmented Reality.” In: *Proc. ACM Hum.-Comput. Interact.* 6.ETRA (May 2022). DOI: 10.1145/3530886 (cited on pages 24–26).

[Ma+21] Ke Ma, Jue Qu, Liping Yang, Wenwen Zhao, and Bernhard Hommel. “Explicit and implicit measures of body ownership and agency: affected by the same manipulations and yet independent.” In: *Experimental Brain Research* 239.7 (May 2021), pages 2159–2170. DOI: 10.1007/s00221-021-06125-5 (cited on pages 100, 119).

[MDC18] Octavian M Machidon, Mihai Duguleana, and Marcello Carrozzino. "Virtual humans in cultural heritage ICT applications: A review." In: *Journal of Cultural Heritage* 33 (2018), pages 249–260. DOI: 10.1016/j.culher.2018.01.007 (cited on page 61).

[Mac92a] I. Scott MacKenzie. "Fitts' law as a research and design tool in human-computer interaction." In: *Human-Computer Interaction* 7.1 (Mar. 1992), pages 91–139. ISSN: 0737-0024. DOI: 10.1207/s15327051hci0701_3 (cited on page 31).

[Mac92b] Ian Scott Mackenzie. "Fitts' law as a performance model in human-computer interaction." (Doctoral dissertation). PhD thesis. Toronto, Ontario, Canada: University of Toronto, 1992. URL: https://www.yorku.ca/mack/phd.html (cited on page 31).

[MWB13] Anne Mangen, Bente R. Walgermo, and Kolbjørn Brønnick. "Reading linear texts on paper versus computer screen: Effects on reading comprehension." In: *International Journal of Educational Research* 58 (2013), pages 61–68. ISSN: 0883-0355. DOI: 10.1016/j.ijer.2012.12.002 (cited on page 19).

[May00] Philipp Mayring. "Qualitative Content Analysis." In: *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* 1.2 (June 2000). DOI: 10.17169/fqs-1.2.1089 (cited on pages 85, 107).

[MG96] Tomasz Mazuryk and Michael Gervautz. *Virtual Reality History, Applications, Technology and Future*. Technical report TR-186-2-96-06. (For human contact, please email: technical-report@cg.tuwien.ac.at). Favoritenstrasse 9-11/E193-02, A-1040 Vienna, Austria: Institute of Computer Graphics and Algorithms, Vienna University of Technology, Feb. 1996. URL: https://www.cg.tuwien.ac.at/research/publications/1996/mazuryk-1996-VRH/ (cited on pages 16, 17).

[Met24] Meta Developers. *User Interface Components*. Website. (accessed August 7, 2024). 2024. URL: https://developers.meta.com/horizon/resources/hands-design-ui/ (cited on page 142).

[Mic24] Microsoft Learn. *App quality criteria overview*. Website. (accessed: 07.08.2024). 2024. URL: https://learn.microsoft.com/en-us/windows/mixed-reality/develop/advanced-concepts/app-quality-criteria-overview (cited on page 142).

[Mica] Microsoft Learn - Mixed Reality. *Interactable object*. Website. (accessed Februrary 5, 2025). URL: https://learn.microsoft.com/en-us/windows/mixed-reality/design/interactable-object (cited on page 20).

[Micb] Microsoft Learn - Mixed Reality. *Interaktives Element [Experimentell] — MRTK2*. Website. (accessed Februrary 5, 2025). URL: https://learn.microsoft.com/de-de/windows/mixed-reality/mrtk-unity/mrtk2/features/experimental/interactive-element?view=mrtkunity-2022-05 (cited on page 20).

[Micc] Microsoft Learn, Eye-based interactions. *Gaze and commit*. Website. (accessed August 10, 2023). URL: https://learn.microsoft.com/en-us/windows/mixed-reality/design/gaze-and-commit (cited on page 25).

[MK94] Paul Milgram and Fumio Kishino. "A Taxonomy of Mixed Reality Visual Displays." In: *IEICE Transactions on Information and Systems* 77 (1994), pages 1321–1329. URL: https://api.semanticscholar.org/CorpusID:17783728 (cited on pages 1, 15, 16).

[Mil+95] Paul Milgram, Haruo Takemura, Akira Utsumi, and Fumio Kishino. "Augmented reality: a class of displays on the reality-virtuality continuum." In: *Telemanipulator and Telepresence Technologies*. Edited by Hari Das. Volume 2351. International Society for Optics and Photonics. SPIE, 1995, pages 282–292. DOI: 10.1117/12.197321 (cited on page 16).

[MB10a] Eric Missimer and Margrit Betke. "Blink and wink detection for mouse pointer control." In: *Proceedings of the 3rd International Conference on PErvasive Technologies Related to Assistive Environments*. (PETRA). Samos, Greece: Association for Computing Machinery, 2010. ISBN: 9781450300711. DOI: 10.1145/1839294.1839322 (cited on page 11).

[MB10b] Eric Missimer and Margrit Betke. "Blink and wink detection for mouse pointer control." In: *Proceedings of the 3rd International Conference on Pervasive Technologies Related to Assistive Environments*. (PETRA). Samos, Greece: Association for Computing Machinery, 2010. ISBN: 9781450300711. DOI: 10.1145/1839294.1839322 (cited on page 28).

[Mon+21] Pedro Monteiro, Guilherme Gonçalves, Hugo Coelho, Miguel Melo, and Maximino Bessa. "Hands-free interaction in immersive virtual reality: A systematic review." In: *IEEE Transactions on Visualization and Computer Graphics* 27.5 (2021), pages 2702–2713. DOI: 10.1109/TVCG.2021.3067687 (cited on pages 11, 24).

[Mon+23] Pedro Monteiro, Guilherme Gonçalves, Bruno Peixoto, Miguel Melo, and Maximino Bessa. "Evaluation of Hands-Free VR Interaction Methods During a Fitts' Task: Efficiency and Effectiveness." In: *IEEE Access* 11 (2023), pages 70898–70911. DOI: 10.1109/ACCESS.2023.3293057 (cited on pages 11, 33).

[Mot+17] Martez E. Mott, Shane Williams, Jacob O. Wobbrock, and Meredith Ringel Morris. "Improving Dwell-Based Gaze Typing with Dynamic, Cascading Dwell Times." In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. (CHI). Denver, Colorado, USA: Association for Computing Machinery, 2017, pages 2558–2570. ISBN: 9781450346559. DOI: 10.1145/3025453.3025517 (cited on pages 11, 24, 135).

[MI01] Atsuo Murata and Hirokazu Iwase. "Extending Fitts' law to a three-dimensional pointing task." In: *Human Movement Science* 20.6 (2001), pages 791–805. ISSN: 0167-9457. DOI: 10.1016/S0167-9457(01)00058-6 (cited on page 34).

[Mur62] Bennet B. Murdock. "The serial position effect of free recall." In: *Journal of Experimental Psychology* 64.5 (Nov. 1962), pages 482–488. DOI: 10.1037/h0045106 (cited on page 90).

[Mut+25] Aunnoy K Mutasim, Mohammad Raihanul Bashar, Christof Lutteroth, Anil Ufuk Batmaz, and Wolfgang Stuerzlinger. "There Is More to Dwell Than Meets the Eye: Toward Better Gaze-Based Text Entry Systems With Multi-Threshold Dwell." In: *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. (CHI). Yokohama, Japan: Association for Computing Machinery, 2025. ISBN: 9798400713941. DOI: 10.1145/3706598.3713781 (cited on page 24).

[MBS21] Aunnoy K Mutasim, Anil Ufuk Batmaz, and Wolfgang Stuerzlinger. "Pinch, Click, or Dwell: Comparing Different Selection Techniques for Eye-Gaze-Based Pointing in Virtual Reality." In: *ACM Symposium on Eye Tracking Research and Applications*. (ETRA). Virtual Event, Germany: Association for Computing Machinery, 2021. ISBN: 9781450383455. DOI: 10.1145/3448018.3457998 (cited on pages 24, 25, 135, 138).

[Nil+18a] Niels Christian Nilsson, Tabitha Peck, Gerd Bruder, Eri Hodgson, Stefania Serafin, Mary Whitton, Frank Steinicke, and Evan Suma Rosenberg. "15 Years of Research on Redirected Walking in Immersive Virtual Environments." In: *IEEE Computer Graphics and Applications* 38.2 (2018), pages 44–56. DOI: 10.1109/MCG.2018.111125628 (cited on page 29).

[Nil+18b] Niels Christian Nilsson, Stefania Serafin, Frank Steinicke, and Rolf Nordahl. "Natural Walking in Virtual Reality: A Review." In: *Computers in Entertainment (CIE)* 16.2 (Apr. 2018). DOI: 10.1145/3180658 (cited on pages 29, 30).

[ND86] Donald A. Norman and Stephen W. Draper. *User Centered System Design: New Perspectives on Human-computer Interaction*. New Perspectives on Human-Computer Interaction Series. USA: Lawrence Erlbaum Associates Inc., 1986. ISBN: 9780898597813 (cited on page 48).

[Nuk+16] Tomi Nukarinen, Jari Kangas, Oleg Špakov, Poika Isokoski, Deepak Akkil, Jussi Rantala, and Roope Raisamo. "Evaluation of HeadTurn: An Interaction Technique Using the Gaze and Head Turns." In: *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*. (NordiCHI). Gothenburg, Sweden: Association for Computing Machinery, 2016. ISBN: 9781450347631. DOI: 10.1145/2971485.2971490 (cited on pages 29, 136).

[OBF03] Alex Olwal, Hrvoje Benko, and Steven Feiner. “SenseShapes: using statistical geometry for object selection in a multimodal augmented reality.” In: *The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003. Proceedings.* Tokyo, Japan: IEEE, Oct. 2003, pages 300–301. DOI: 10.1109/ISMAR.2003.1240730 (cited on page 22).

[OF03] Alex Olwal and Steven Feiner. “The flexible pointer: An interaction technique for selection in augmented and virtual reality.” In: *Proceedings of the UIST’2003 ACM Symposium On User Interface Software and Technology*. Vancouver, Canada: Association for Computing Machinery, 2003, Poster. URL: https://uist.hosting.acm.org/uist2005/images/poster_examples/pointer.pdf (cited on page 22).

[Oly] Olympia Back In Time. *Olympia in virtual reality: self guided tour with audio guide*. Website. (Virtual Reality application, accessed November 6, 2024). URL: https://www.olympiabackintime.com/ (cited on page 17).

[Ost25] Ann Lauren Osthof. “City Scripts. Inschriften und die Konstruktion von sozialen Räumen in Milet (Asia Minor).” unpublished PhD thesis. Hamburg, Germany: University of Hamburg, Cluster of Excellence "Understanding Written Artefacts", Centre for the Study of Manuscript Cultures (CSMC), Jan. 2025 (cited on pages 7, 8, 51, 55).

[Pag20] PageRangers. *Flesch Reading Ease (FRE) for the German Language*. Website. (accessed January 12, 2023). 2020. URL: https://pagerangers.com/glossar/flesch-reading-ease/ (cited on page 78).

[Par05] Ross Parry. “Digital heritage and the rise of theory in museum computing.” In: *Museum Management and Curatorship* 20.4 (2005), pages 333–348. DOI: 10.1080/09647770500802004 (cited on page 16).

[Pau+] Erick J. Paul, Vinnie Tieto, Stefano Formicola, and David Coulter. *Comfort - Windows Mixed Reality*. Website. (accessed May 14, 2023). URL: https://learn.microsoft.com/en-us/windows/mixed-reality/design/comfort (cited on pages 99, 118).

[PGG24] Ken Pfeuffer, Hans Gellersen, and Mar Gonzalez-Franco. “Design Principles and Challenges for Gaze + Pinch Interaction in XR.” In: *IEEE Computer Graphics and Applications* 44.3 (2024), pages 74–81. DOI: 10.1109/MCG.2024.3382961 (cited on pages 22, 25, 131).

[Pfe+17] Ken Pfeuffer, Benedikt Mayer, Diako Mardanbegi, and Hans Gellersen. “Gaze + Pinch Interaction in Virtual Reality.” In: *Proceedings of the 5th Symposium on Spatial User Interaction*. (SUI). Brighton, United Kingdom: Association for Computing Machinery, 2017, pages 99–108. ISBN: 9781450354868. DOI: 10.1145/3131277.3132180 (cited on pages vii, ix, 10, 24–26, 113, 116, 136, 137).

[Pie+97] Jeffrey S. Pierce, Andrew S. Forsberg, Matthew J. Conway, Seung Hong, Robert C. Zeleznik, and Mark R. Mine. "Image plane interaction techniques in 3D immersive environments." In: *Proceedings of the 1997 Symposium on Interactive 3D Graphics*. (I3D). Providence, Rhode Island, USA: Association for Computing Machinery, 1997, 39–ff. ISBN: 0897918843. DOI: 10.1145/253284.253303 (cited on page 26).

[Plo+22] Alexander Plopski, Teresa Hirzle, Nahal Norouzi, Long Qian, Gerd Bruder, and Tobias Langlotz. "The Eye in Extended Reality: A Survey on Gaze Interaction and Eye Tracking in Head-worn Extended Reality." In: *ACM Computing Surveys* 55.3 (Mar. 2022). ISSN: 0360-0300. DOI: 10.1145/3491207 (cited on page 24).

[PZ23] Chris Porter and Gary Zammit. "Blink, Pull, Nudge or Tap? The Impact of Secondary Input Modalities on Eye-Typing Performance." In: *HCI International 2023 – Late Breaking Papers: 25th International Conference on Human-Computer Interaction, HCII 2023, Copenhagen, Denmark, July 23–28, 2023, Proceedings, Part I*. Copenhagen, Denmark: Springer-Verlag, 2023, pages 238–258. ISBN: 978-3-031-48037-9. DOI: 10.1007/978-3-031-48038-6_15 (cited on pages 28, 139).

[Pou+96] Ivan Poupyrev, Mark Billinghurst, Suzanne Weghorst, and Tadao Ichikawa. "The Go-Go Interaction Technique: Non-Linear Mapping for Direct Manipulation in VR." In: *Proceedings of the 9th Annual ACM Symposium on User Interface Software and Technology*. (UIST). Seattle, Washington, USA: Association for Computing Machinery, 1996, pages 79–80. ISBN: 0897917987. DOI: 10.1145/237091.237102 (cited on pages 20, 30).

[Pou+98] Ivan Poupyrev, Tadao Ichikawa, Suzanne Weghorst, and Mark Billinghurst. "Egocentric Object Manipulation in Virtual Environments: Empirical Evaluation of Interaction Techniques." In: *Computer Graphics Forum* Volume17.Issue3 (1998), pages 41–52. ISSN: 1467-8659. DOI: 10.1111/1467-8659.00252 (cited on pages 10, 21, 22, 93).

[Pou+20] Florent Poux, Quentin Valembois, Christian Mattes, Leif Kobbelt, and Roland Billen. "Initial User-Centered Design of a Virtual Reality Heritage System: Applications for Digital Tourism." In: *Remote Sensing* 12.16 (2020). ISSN: 2072-4292. DOI: 10.3390/rs12162583 (cited on page 17).

[Ram+21] Argenis Ramirez Ramirez Gomez, Christopher Clarke, Ludwig Sidenmark, and Hans Gellersen. "Gaze+Hold: Eyes-only Direct Manipulation with Continuous Gaze Modulated by Closure of One Eye." In: *ACM Symposium on Eye Tracking Research and Applications*. (ETRA). Virtual Event, Germany: Association for Computing Machinery, 2021. ISBN: 9781450383448. DOI: 10.1145/3448017.3457381 (cited on pages 11, 27, 28).

[Rät+24] Fabian Räthel, Susanne Schmidt, Jenny Gabel, Lukas Posniak, and Frank Steinicke. "An Evaluation of Targeting Methods in Spatial Computing Interfaces with Visual Distractions." In: *Proceedings of the 30th ACM Symposium on Virtual Reality Software* *and Technology*. (VRST). Trier, Germany: Association for Computing Machinery, 2024, pages 1–11. ISBN: 9798400705359. DOI: 10.1145/3641825.3687712 (cited on page 14).

[RZG21] Pei-Luen Patrick Rau, Jian Zheng, and Zhi Guo. "Immersive reading in virtual and augmented reality environment." In: *Information and Learning Sciences* 122.7/8 (2021), pages 464–479. DOI: 10.1108/ils-11-2020-0236 (cited on pages 19, 78).

[Rau+18] Pei-Luen Patrick Rau, Jian Zheng, Zhi Guo, and Jiaqi Li. "Speed reading on virtual reality and augmented reality." In: *Computers & Education* 125 (2018), pages 240–245. ISSN: 0360-1315. DOI: 10.1016/j.compedu.2018.06.016 (cited on pages 19, 78).

[Rea] Realities.io Inc. *realities.io*. Website. (accessed May 10, 2021). URL: https://www.realities.io/ (cited on page 18).

[Reb+23] Mikkel Rosholm Rebsdorf, Theo Khumsan, Jonas Valvik, Niels Christian Nilsson, and Ali Adjorlu. "Blink Don't Wink: Exploring Blinks as Input for VR Games." In: *Proceedings of the 2023 ACM Symposium on Spatial User Interaction*. (SUI). Sydney, NSW, Australia: Association for Computing Machinery, 2023. ISBN: 9798400702815. DOI: 10.1145/3607822.3614527 (cited on page 28).

[RS02] Holder Regenbrecht and Thomas Schubert. "Real and Illusory Interactions Enhance Presence in Virtual Environments." In: *Presence: Teleoper. Virtual Environ.* 11.4 (2002), pages 425–434. ISSN: 1054-7460. DOI: 10.1162/105474602760204318 (cited on pages 68, 212).

[Rei89] Paul Reilly. "Data visualization in archaeology." In: *IBM Systems Journal* 28 (Feb. 1989), pages 569–579. DOI: 10.1147/sj.284.0569 (cited on page 16).

[Rei91] Paul Reilly. "Towards a virtual archaeology." In: *Computer Applications and Quantitative Methods in Archaeology*. (CAA90). 1991, pages 132–139. URL: https://proceedings.caaconference.org/paper/21_reilly_caa_1990/ (cited on page 16).

[Rei21] Jan Reimann. "Virtual History. Das Forum Romanum in Virtual Reality." In: *Potsdamer Lateintage 2018 - 2020*. Das Forum Romanum in Virtual Reality (2021). Edited by Alexandra Forst, pages 117–126. DOI: 10.25932/publishup-53348 (cited on page 17).

[RA] Retro Futuro and Arvizio. *Pompeii. Voices from the past*. Website. (Virtual Reality application, accessed November 6, 2024). URL: https://pompeii.refutur.com/en/ (cited on page 17).

[RT97] William H Ridder 3rd and Alan Tomlinson. "A comparison of saccadic and blink suppression in normal observers." In: *Vision research* 37.22 (1997), pages 3171–3179. DOI: 10.1016/s0042-6989(97)00110-7 (cited on page 162).

[Rou07] Maria Roussou. “The Components of Engagement in Virtual Heritage Environments.” In: *New Heritage – New Media and Cultural Heritage*. 1st. Routledge, 2007, page 17. DOI: 10.4324/9780203937884 (cited on page 16).

[Rza+21] Rufat Rzayev, Polina Ugnivenko, Sarah Graf, Valentin Schwind, and Niels Henze. “Reading in VR: The Effect of Text Presentation Type and Location.” In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. (CHI). Yokohama, Japan: Association for Computing Machinery, 2021. ISBN: 9781450380966. DOI: 10.1145/3411764.3445606 (cited on pages 20, 79).

[San23] Donald H. Sanders. *From Photography to 3D Models and Beyond: Visualizations in Archaeology*. ARCHAEOPRESS, 2023 (cited on pages 16, 17).

[SR19] Annamaria De Santis and Irene Rossi, editors. *Crossing Experiences in Digital Epigraphy*. Warsaw, Poland: De Gruyter Open Poland, 2019. DOI: 10.1515/9783110607208 (cited on pages 18, 47).

[Sat+15] Hironobu Sato, Kiyohiko Abe, Shoichi Ohi, and Minoru Ohyama. “Automatic Classification Between Involuntary and Two Types of Voluntary Blinks Based on an Image Analysis.” In: *Human-Computer Interaction: Interaction Technologies*. Cham: Springer International Publishing, 2015, pages 140–149. ISBN: 978-3-319-20916-6. DOI: 10.1007/978-3-319-20916-6_14 (cited on page 151).

[Sat+17] Hironobu Sato, Kiyohiko Abe, Shoichi Ohi, and Minoru Ohyama. “An automatic classification method for involuntary and two types of voluntary blinks.” In: *Electronics and Communications in Japan* 100.10 (2017), pages 48–58 (cited on page 151).

[SS17] Susanne Schmidt and Frank Steinicke. “A Projection-Based Augmented Reality Setup for Blended Museum Experiences.” In: *Proceedings of the 27th International Conference on Artificial Reality and Telexistence and 22nd Eurographics Symposium on Virtual Environments: Posters and Demos*. (ICAT-EGVE). Adelaide, Australia: Eurographics Association, 2017, pages 5–6. DOI: 10.2312/egve.20171366 (cited on pages 17, 62).

[STH17] Martin Schrepp, Jörg Thomaschewski, and Andreas Hinderks. “Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S).” In: *International Journal of Interactive Multimedia and Artificial Intelligence* 4.6 (Dec. 2017), pages 103–108. ISSN: 1989-1660. DOI: 10.9781/ijimai.2017.09.001 (cited on pages 143, 216).

[Sch03] Thomas W. Schubert. “The sense of presence in virtual environments: A three-component scale measuring spatial presence,involvement, and realness.” In: *Zeitschrift für Medienpsychologie* 15 (2003), pages 69–71. DOI: 10.1026//1617-6383.15.2.69 (cited on pages 68, 212).

[SSM19] Robin Schweigert, Valentin Schwind, and Sven Mayer. "EyePointing: A Gaze-Based Selection Technique." In: *Proceedings of Mensch Und Computer 2019*. (MuC). Hamburg, Germany: Association for Computing Machinery, 2019, pages 719–723. ISBN: 9781450371988. DOI: 10.1145/3340764.3344897 (cited on page 27).

[Sea06] Frank Sear. *Roman Theatres: An Architectural Study*. Oxford University Press, USA, July 2006. ISBN: 9780198144694 (cited on page 61).

[SPR19] Helen Sharp, Jennifer Preece, and Yvonne Rogers. *Interaction Design: Beyond Human-Computer Interaction*. 5th. John Wiley & Sons Inc, 2019. ISBN: 978-1-119-54725-9 (cited on page 48).

[SC18] William R. Sherman and Alan B. Craig. *Understanding Virtual Reality: Interface, Application, and Design*. 2nd edition. San Francisco: Morgan Kaufmann, Nov. 2018. ISBN: 9780128009659. DOI: 10.1016/C2013-0-18583-2 (cited on page 2).

[Sid+20] Ludwig Sidenmark, Christopher Clarke, Xuesong Zhang, Jenny Phu, and Hans Gellersen. "Outline Pursuits: Gaze-assisted Selection of Occluded Objects in Virtual Reality." In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. (CHI). Honolulu, HI, USA: Association for Computing Machinery, 2020, pages 1–13. ISBN: 9781450367080. DOI: 10.1145/3313831.3376438 (cited on pages 27, 135).

[SG19] Ludwig Sidenmark and Hans Gellersen. "Eye&Head: Synergetic Eye and Head Movement for Gaze Pointing and Selection." In: *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*. (UIST). New Orleans, LA, USA: Association for Computing Machinery, 2019, pages 1161–1174. ISBN: 9781450368162. DOI: 10.1145/3332165.3347921 (cited on pages 24, 25, 28, 29).

[Sid+22] Ludwig Sidenmark, Mark Parent, Chi-Hao Wu, Joannes Chan, Michael Glueck, Daniel Wigdor, Tovi Grossman, and Marcello Giordano. "Weighted Pointer: Error-aware Gaze-based Interaction through Fallback Modalities." In: *IEEE Transactions on Visualization and Computer Graphics* 28.11 (2022), pages 3585–3595. DOI: 10.1109/TVCG.2022.3203096 (cited on page 27).

[Ske] Sketchfab Inc. *Sketchfab for cultural heritage*. Website. (accessed May 5, 2021). URL: https://sketchfab.com/museums (cited on page 18).

[Škr+23] Ondřej Škrabal, Leah Mascia, Ann Lauren Osthof, and Malena Ratzke, editors. *Graffiti Scratched, Scrawled, Sprayed: Towards a Cross-Cultural Understanding*. Berlin, Boston: De Gruyter, 2023. ISBN: 9783111326306. DOI: 10.1515/9783111326306 (cited on page 2).

[SR09] Timothy J. Slattery and Keith Rayner. "The influence of text legibility on eye movements during reading." In: *Applied Cognitive Psychology* 24.8 (Oct. 2009), pages 1129–1148. DOI: 10.1002/acp.1623 (cited on page 19).

[SSA20] Johann Anselm Steiger, Michael Schilling, and Stefanie Arend. *Sinnbilder im Sakralraum. Die Kirche in Lucklum – Ein Kompendium der geistlichen Emblematik der Frühen Neuzeit.* Schnell & Steiner, 2020. ISBN: 978-3-7954-3501-1 (cited on pages 5, 6, 41, 51, 52, 54, 58, 75).

[Ste16] Frank Steinicke. *Being Really Virtual - Immersive Natives and the Future of Virtual Reality*. Cham, Switzerland: Springer Cham, 2016. ISBN: 978-3-319-43076-8. DOI: 10.1007/978-3-319-43078-2 (cited on pages 1, 29).

[Ste+08] Frank Steinicke, Gerd Bruder, Luv Kohli, Jason Jerald, and Klaus Hinrichs. "Taxonomy and Implementation of Redirection Techniques for Ubiquitous Passive Haptic Feedback." In: *2008 International Conference on Cyberworlds*. 2008, pages 217–223. DOI: 10.1109/CW.2008.53 (cited on page 29).

[SRH06] Frank Steinicke, Timo Ropinski, and Klaus Hinrichs. "Object Selection in Virtual Environments Using an Improved Virtual Pointer Metaphor." In: *Computer Vision and Graphics: International Conference, ICCVG 2004, Warsaw, Poland, September 2004, Proceedings*. Dordrecht: Springer Netherlands, 2006, pages 320–326. ISBN: 978-1-4020-4179-2. DOI: 10.1007/1-4020-4179-9_46 (cited on pages 22, 23, 94).

[Ste+13] Frank Steinicke, Yon Visell, Jennifer Campos, and Anatole Lécuyer, editors. *Human Walking in Virtual Environments - Perception, Technology, and Applications*. New York, NY: Springer, 2013. ISBN: 978-1-4419-8431-9. DOI: 10.1007/978-1-4419-8432-6 (cited on page 45).

[SCP95] Richard Stoakley, Matthew J. Conway, and Randy Pausch. "Virtual reality on a WIM: interactive worlds in miniature." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. (CHI). Denver, Colorado, USA: ACM Press/Addison-Wesley Publishing Co., 1995, pages 265–272. ISBN: 0201847051. DOI: 10.1145/223904.223938 (cited on page 20).

[Sul16] Ann-Marie T Sullivan. "Cultural Heritage & New Media: A Future for the Past, 15 J. Marshall Rev. Intell. Prop. L. 604 (2016)." In: *John Marshall Review of Intellectual Property Law* 15 (2016), page 11. URL: https://api.semanticscholar.org/CorpusID:156600160 (cited on page 16).

[Sut68] Ivan E. Sutherland. "A head-mounted three dimensional display." In: *Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I*. (AFIPS). San Francisco, California: Association for Computing Machinery, 1968, pages 757–764. ISBN: 9781450378994. DOI: 10.1145/1476589.1476686 (cited on page 1).

[TS11] Robert J. Teather and Wolfgang Stuerzlinger. "Pointing at 3D targets in a stereo head-tracked virtual environment." In: *2011 IEEE Symposium on 3D User Interfaces (3DUI)*. Singapore: IEEE, Mar. 2011, pages 87–94. DOI: 10.1109/3DUI.2011.5759222 (cited on pages 33, 34, 102, 121).

[TBV20] Jennifer Tidwell, Charles Brewer, and Aynne Valencia. *Designing Interfaces*. Edited by Jennifer Pollock, Angela Rufino, Christopher Faucher, and Octal Publishing, LLC. 3rd edition. O'Reilly Media, Inc., 2020. ISBN: 9781492051961 (cited on pages 54, 76, 91).

[Tim+25] Tim Rolff and Jenny Gabel, Lauren Zerbin, Niklas Hypki, Susanne Schmidt, Markus Lappe, and Frank Steinicke. *A Hands-free Spatial Selection and Interaction Technique using Gaze and Blink Input with Blink Prediction for Extended Reality*. (Author's Original Manuscript (AOM) preprint). 2025. DOI: 10.48550/arXiv.2501.11540. arXiv: 2501.11540 [cs.HC] (cited on pages 14, 135, 150–152, 162).

[Tox21] Anders Toxboe. *User Interface Design patterns*. Website. (accessed January 11, 2023). 2021. URL: https://ui-patterns.com/ (cited on page 76).

[TL21] Eleftherios Triantafyllidis and Zhibin Li. "The Challenges in Modeling Human Performance in 3D Space with Fitts' Law." In: *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. (CHI EA). Yokohama, Japan: Association for Computing Machinery, 2021. ISBN: 9781450380959. DOI: 10.1145/3411763.3443442 (cited on pages 31, 34).

[UNE09] UNESCO Institute for Statistics. *Cultural heritage - Definition*. Website. (accessed January 10, 2025). 2009. URL: https://uis.unesco.org/en/glossary-term/cultural-heritage (cited on page 16).

[Unia] Unity XR Interaction Toolkit. *Near-Far Interactor*. Website. (accessed Februrary 5, 2025). URL: https://docs.unity3d.com/Packages/com.unity.xr.interaction.toolkit@3.1/manual/near-far-interactor.html (cited on page 20).

[Unib] Universität Tübingen (MUT). *Tempelsteuer und Taubenhändler: Geld im Jerusalemer Tempel zur Zeit Jesu in Virtual Reality*. Website. (Research project 1391 'Andere Ästhetik', accessed November 6, 2024). URL: https://www.unimuseum.uni-tuebingen.de/de/ausstellungen/sonderausstellungen/tempelsteuer-und-taubenhaendler-1 (cited on page 18).

[Uso+99] Martin Usoh, Kevin Arthur, Mary C. Whitton, Rui Bastos, Anthony Steed, Mel Slater, and Frederick P. Brooks. "Walking > Walking-in-Place > Flying, in Virtual Environments." In: *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques*. (SIGGRAPH). 1999, pages 359–364. DOI: 10.1145/311535.311589 (cited on page 45).

[VGC07] Lode Vanacken, Tovi Grossman, and Karin Coninx. "Exploring the Effects of Environment Density and Target Visibility on Object Selection in 3D Virtual Environments." In: *2007 IEEE Symposium on 3D User Interfaces*. Charlotte, NC, USA: IEEE, Mar. 2007. DOI: 10.1109/3DUI.2007.340783 (cited on pages 22, 23).

[VC16] Eduardo Velloso and Marcus Carter. "The Emergence of EyePlay: A Survey of Eye Interaction in Games." In: *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play*. (CHI PLAY). Austin, Texas, USA: Association for Computing Machinery, 2016, pages 171–185. ISBN: 9781450344562. DOI: 10.1145/2967934.2968084 (cited on page 28).

[Vic+14] Rodrigo Vicencio-Moreira, Regan L. Mandryk, Carl Gutwin, and Scott Bateman. "The effectiveness (or lack thereof) of aim-assist techniques in first-person shooter games." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. (CHI). Toronto, Ontario, Canada: Association for Computing Machinery, 2014, pages 937–946. ISBN: 9781450324731. DOI: 10.1145/2556288.2557308 (cited on page 30).

[Wag+23] Uta Wagner, Mathias N. Lystbæk, Pavel Manakhov, Jens Emil Sloth Grønbæk, Ken Pfeuffer, and Hans Gellersen. "A Fitts' Law Study of Gaze-Hand Alignment for Selection in 3D User Interfaces." In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. (CHI). Hamburg, Germany: Association for Computing Machinery, 2023. ISBN: 9781450394215. DOI: 10.1145/3544548.3581423 (cited on pages 24–26, 33, 34, 116).

[Wäs+05] Erik Wästlund, Henrik Reinikka, Torsten Norlander, and Trevor Archer. "Effects of VDT and paper presentation on consumption and production of information: Psychological and physiological factors." In: *Computers in Human Behavior* 21.2 (2005), pages 377–394. ISSN: 0747-5632. DOI: 10.1016/j.chb.2004.02.007 (cited on page 19).

[Wel68] Alan T. Welford. "Movement." In: *Fundamentals of Skill*. London: Methuen & Co Ltd, 1968, pages 137–160 (cited on pages 31, 32).

[WBR02] Chadwick A. Wingrave, Doug A. Bowman, and Naren Ramakrishnan. "Towards Preferences in Virtual Environment Interfaces." In: *Eurographics Workshop on Virtual Environments*. Edited by S. Mueller and W. Stuerzlinger. The Eurographics Association, 2002. ISBN: 1-58113-535-1. DOI: 10.2312/EGVE/EGVE02/063-072 (cited on page 22).

[Wol+20] 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 2020 CHI Conference on Human Factors in Computing Systems*. (CHI). Honolulu, HI, USA: Association for Computing Machinery, 2020, pages 1–10. ISBN: 9781450367080. DOI: 10.1145/3313831.3376876 (cited on page 22).

[WWC19] Zebulun M. Wood, Albert William, and Andrea Copeland. "Virtual Reality for Preservation: Production of Virtual Reality Heritage Spaces in the Classrooom." In: *3D/VR in the Academic Library: Emerging Practices and Trends*. Council on Library and

Information Resources, 2019. Chapter 3, pages 39–53. ISBN: 978-1-932326-60-4. URL: http://hdl.handle.net/1805/23685 (cited on page 17).

[Woo+18] Zebulun M. Wood, Albert William, Ayoung Yoon, and Andrea Copeland. “Virtual Bethel: Preservation of Indianapolis’s Oldest Black Church.” In: *Research Methods for the Digital Humanities*. Edited by lewis levenberg lewis, Tai Neilson, and David Rheams. Cham: Springer International Publishing, 2018, pages 195–210. ISBN: 978-3-319-96713-4. DOI: 10.1007/978-3-319-96713-4_11 (cited on page 17).

[Woo99] Robert S. Woodworth. “Accuracy of voluntary movement.” In: *The Psychological Review: Monograph Supplements* 3.3 (1899), pages i–114. DOI: 10.1037/h0092992 (cited on page 34).

[XQL19] Jing Xiao, Jun Qu, and Yuanqing Li. “An Electrooculogram-Based Interaction Method and Its Music-on-Demand Application in a Virtual Reality Environment.” In: *IEEE Access* 7 (2019), pages 22059–22070. DOI: 10.1109/ACCESS.2019.2898324 (cited on page 28).

[Yu+18] Difeng Yu, Hai-Ning Liang, Feiyu Lu, Vijayakumar Nanjappan, Konstantinos Papangelis, and Wei Wang. “Target Selection in Head-Mounted Display Virtual Reality Environments.” In: *Journal of Universal Computer Science* 24 (2018), pages 1217–1243. DOI: 10.3217/jucs-024-09-1217 (cited on page 23).

[ZP15] René Zeelenberg and Diane Pecher. “A method for simultaneously counterbalancing condition order and assignment of stimulus materials to conditions.” In: *Behavior Research Methods* 47 (2015), pages 127–133. DOI: 10.3758/s13428-014-0476-9 (cited on page 78).

[Zen+24] André Zenner, Chiara Karr, Martin Feick, Oscar Ariza, and Antonio Krüger. “Beyond the Blink: Investigating Combined Saccadic & Blink-Suppressed Hand Redirection in Virtual Reality.” In: *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. (CHI). Honolulu, HI, USA: Association for Computing Machinery, 2024. ISBN: 979-8-4007-0330-0. DOI: 10.1145/3613904.3642073 (cited on page 30).

[ZKK21] André Zenner, Hannah Maria Kriegler, and Antonio Krüger. “HaRT - The Virtual Reality Hand Redirection Toolkit.” In: *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, 2021. ISBN: 9781450380959. DOI: 10.1145/3411763.3451814 (cited on pages 29, 30).

[ZK19] André Zenner and Antonio Krüger. “Estimating Detection Thresholds for Desktop-Scale Hand Redirection in Virtual Reality.” In: *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. Osaka, Japan: IEEE, Mar. 2019, pages 47–55. DOI: 10.1109/VR.2019.8798143 (cited on pages 29, 30).

[ZRK21] André Zenner, Kora Persephone Regitz, and Antonio Krüger. "Blink-Suppressed Hand Redirection." In: *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. Lisbon, Portugal, 2021, pages 75–84. DOI: 10.1109/VR50410.2021.00028 (cited on page 30).

[ZMI99] Shumin Zhai, Carlos Morimoto, and Steven Ihde. "Manual and gaze input cascaded (MAGIC) pointing." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. (CHI). Pittsburgh, Pennsylvania, USA: Association for Computing Machinery, 1999, pages 246–253. ISBN: 0201485591. DOI: 10.1145/302979.303053 (cited on pages 27, 116).

APPENDIX

QUESTIONNAIRES

A1.1 Simulator Sickness Questionnaire

Simulator Sickness Questionnaire (SSQ) by Kennedy et al. [Ken+93].

Please rate how much each symptom is affecting you **right now**:

	None	Slight	Moderate	Severe
General discomfort	◯	◯	◯	◯
Fatigue	◯	◯	◯	◯
Headache	◯	◯	◯	◯
Eyestrain	◯	◯	◯	◯
Difficulty focusing	◯	◯	◯	◯
Increased salivation	◯	◯	◯	◯
Sweating	◯	◯	◯	◯
Nausea	◯	◯	◯	◯
Difficulty concentrating	◯	◯	◯	◯
Fullness of head	◯	◯	◯	◯
Blurred vision	◯	◯	◯	◯
Dizzy (eyes open)	◯	◯	◯	◯
Dizzy (eyes closed)	◯	◯	◯	◯
Vertigo	◯	◯	◯	◯
Stomach awareness	◯	◯	◯	◯
Burping	◯	◯	◯	◯

A1.2 NASA Task Load Index

The NASA Task Load Index (NASA-TLX) by Hart [Har06a] assesses the perceived workload based on 6 subscales.

Mental Demand – How mentally demanding was the task?

◯ ◯ ◯ ◯ ◯ ◯ ◯ ◯ ◯ ◯ ◯ ◯ ◯ ◯ ◯ ◯ ◯ ◯ ◯ ◯

Very Low Very High

Physical Demand – How physically demanding was the task?

◯ ◯ ◯ ◯ ◯ ◯ ◯ ◯ ◯ ◯ ◯ ◯ ◯ ◯ ◯ ◯ ◯ ◯ ◯ ◯

Very Low Very High

Temporal Demand – How hurried or rushed was the pace of the task?

○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○

Very Low | Very High

Performance – How successful were you in accomplishing what you were asked to do?

○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○

Perfect | Failure

Effort – How hard did you have to work to accomplish your level of performance?

○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○

Very Low | Very High

Frustration – How insecure, discouraged, irritated, stressed, and annoyed were you?

○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○

Very Low | Very High

A1.3 System Usability Scale

System Usability Scale (SUS) according to Brooke et al. [Bro+96].

I think that I would like to use this system frequently.

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I found the system unnecessarily complex.

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I thought the system was easy to use.

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I think that I would need the support of a technical person to be able to use this system.

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I found the various functions in this system were well integrated.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

I thought there was too much inconsistency in this system.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

I would imagine that most people would learn to use this system very quickly.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

I found the system very cumbersome to use.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

I felt very confident using the system.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

I needed to learn a lot of things before I could get going with this system.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

A1.4 Igroup Presence Questionnaire

Igroup Presence Questionnaire (IPQ) according to Regenbrecht and Schubert [RS02] and Schubert [Sch03].

In the computer generated world I had a sense of “being there”.

	-3	-2	-1	0	1	2	3	
not at all	◯	◯	◯	◯	◯	◯	◯	very much

Somehow I felt that the virtual world surrounded me.

	-3	-2	-1	0	1	2	3	
fully disagree	○	○	○	○	○	○	○	fully agree

I felt like I was just perceiving pictures.

	-3	-2	-1	0	1	2	3	
fully disagree	○	○	○	○	○	○	○	fully agree

I did not feel present in the virtual space.

	-3	-2	-1	0	1	2	3	
did not feel	○	○	○	○	○	○	○	felt present

I had a sense of acting in the virtual space, rather than operating something from outside.

	-3	-2	-1	0	1	2	3	
fully disagree	○	○	○	○	○	○	○	fully agree

I felt present in the virtual space.

	-3	-2	-1	0	1	2	3	
fully disagree	○	○	○	○	○	○	○	fully agree

How aware were you of the real world surrounding while navigating in the virtual world? (i.e. sounds, room temperature, other people, etc.)?

	-3	-2	-1	0	1	2	3	
extremely aware	○	○	○	○	○	○	○	not aware at all

I was not aware of my real environment.

	-3	-2	-1	0	1	2	3	
fully disagree	○	○	○	○	○	○	○	fully agree

I still paid attention to the real environment.

	-3	-2	-1	0	1	2	3	
fully disagree	◯	◯	◯	◯	◯	◯	◯	fully agree

I was completely captivated by the virtual world.

	-3	-2	-1	0	1	2	3	
fully disagree	◯	◯	◯	◯	◯	◯	◯	fully agree

How real did the virtual world seem to you?

	-3	-2	-1	0	1	2	3	
completely real	◯	◯	◯	◯	◯	◯	◯	not real at all

How much did your experience in the virtual environment seem consistent with your real world experience?

	-3	-2	-1	0	1	2	3	
not consistent	◯	◯	◯	◯	◯	◯	◯	very consistent

How real did the virtual world seem to you?

	-3	-2	-1	0	1	2	3	
about as real as an imagined world	◯	◯	◯	◯	◯	◯	◯	indistinguishable from the real world

The virtual world seemed more realistic than the real world.

	-3	-2	-1	0	1	2	3	
fully disagree	◯	◯	◯	◯	◯	◯	◯	fully agree

A1.5 User Experience Questionnaire

User Experience Questionnaire (UEQ) with 26 items representing the 6 scales *Attractiveness*, *Perspicuity*, *Efficiency*, *Dependability*, *Stimulation*, and *Novelty* according to Laugwitz et al. [LHS08].

	1	2	3	4	5	6	7	
annoying	○	○	○	○	○	○	○	enjoyable
not understandable	○	○	○	○	○	○	○	understandable
creative	○	○	○	○	○	○	○	dull
easy to learn	○	○	○	○	○	○	○	difficult to learn
valuable	○	○	○	○	○	○	○	inferior
boring	○	○	○	○	○	○	○	exciting
not interesting	○	○	○	○	○	○	○	interesting
unpredictable	○	○	○	○	○	○	○	predictable
fast	○	○	○	○	○	○	○	slow
inventive	○	○	○	○	○	○	○	conventional
obstructive	○	○	○	○	○	○	○	supportive
good	○	○	○	○	○	○	○	bad
complicated	○	○	○	○	○	○	○	easy
unlikable	○	○	○	○	○	○	○	pleasing
usual	○	○	○	○	○	○	○	leading edge
unpleasant	○	○	○	○	○	○	○	pleasant
secure	○	○	○	○	○	○	○	not secure
motivating	○	○	○	○	○	○	○	demotivating
meets expectations	○	○	○	○	○	○	○	does not meet expectations
inefficient	○	○	○	○	○	○	○	efficient
clear	○	○	○	○	○	○	○	confusing
impractical	○	○	○	○	○	○	○	practical
organized	○	○	○	○	○	○	○	cluttered
attractive	○	○	○	○	○	○	○	unattractive
friendly	○	○	○	○	○	○	○	unfriendly
conservative	○	○	○	○	○	○	○	innovative

A1.5.1 UEQ-S

The short version of the User Experience Questionnaire by Schrepp et al. [STH17] consisting of 8 items, with 4 items representing the scale *Pragmatic Quality* and 4 the scale *Hedonic Quality*.

	1	2	3	4	5	6	7	
obstructive	◯	◯	◯	◯	◯	◯	◯	supportive
complicated	◯	◯	◯	◯	◯	◯	◯	easy
inefficient	◯	◯	◯	◯	◯	◯	◯	efficient
confusing	◯	◯	◯	◯	◯	◯	◯	clear
boring	◯	◯	◯	◯	◯	◯	◯	exciting
not interesting	◯	◯	◯	◯	◯	◯	◯	interesting
conventional	◯	◯	◯	◯	◯	◯	◯	inventive
usual	◯	◯	◯	◯	◯	◯	◯	leading edge

SPATIAL USER INTERFACE FOR READING LONG TEXTS

A2.1 Reading Task in VR – German Texts

The four long texts used in the user study for reading in VR were taken from the University of Hamburg's magazine *Neunzehn* from 2021 [Ann21]. The texts are in German and were partially shortened and slightly modified to fit the FRE and row number requirements for our study (cf. Section 5.3.2). All four texts are about researchers who present their international research projects.

Text 1

67 Tage gemeinsam auf einem Schiff und kein einziger Landgang. Die Expedition auf dem Forschungsschiff SONNE im Frühjahr 21 war auch für die erfahrensten der 13 Teilnehmenden eine Reise voller Extreme. Ich war Fahrtleiter auf dem Schiff, das von der Leitstelle Deutsche Forschungsschiffe an der Universität Hamburg koordiniert wird. Schon vorab mussten wir in Quarantäne und haben danach zweieinhalb Monate durchgehend auf engstem Raum verbracht. Das war eine sehr intensive Zeit mit wenig Privatsphäre. Hinzu kam die Ungewissheit, ob es uns gelingen würde, unser Ziel zu erreichen. Können wir die unter Wasser liegenden Instrumente mit Forschungsdaten der vergangenen zwei Jahre retten? Für unsere wissenschaftliche Arbeit, unter anderem zum Klimawandel, haben wir verschiedene Messgeräte und viele Sensoren ausgesetzt. Mit diesen sammeln wir in langen Beobachtungsreihen Proben und Daten, etwa zu Temperatur, Salzgehalt, Meeresströmung oder Partikelfluss. Doch wichtige Teile der Systeme, etwa um sie wieder an die Oberfläche zu holen, werden mit Batterien betrieben, die regelmäßig gewartet werden müssen. Da wir zunächst aufgrund der Corona-Pandemie nicht ausfahren konnten, bestand die Gefahr, dass wir die Geräte verlieren. Und damit alle kostbaren Messdaten und Proben. Als wir schließlich an der Südspitze Afrikas ankamen, standen wir vor der bangen Frage: Laufen die Batterien überhaupt noch? Um die Verankerungen der Instrumente zu lösen, schicken wir ein spezielles Schallsignal durch eine Wassersäule. Das setzt einen Mechanismus in Gang und löst so die Systeme vom Ankergewicht am Meeresboden. Dann kommt das Messgerät hoch, wir können es an der Meeresoberfläche einsammeln und an Bord holen. Doch was, wenn der Auslöser nicht mehr funktioniert? Dann die Erlösung: Ja, es ist gelungen! Wir waren mit der Expedition sehr erfolgreich und haben alle Messgeräte wiederbekommen. Zurück in Hamburg werden die vielen Proben und Daten aus dem Südatlantik nun ausgewertet und analysiert. Sie helfen zum Beispiel, die Frage zu beantworten, welchen Einfluss die Meere auf den Klimawandel haben. Die Meere haben bislang 30 bis 50 Prozent des gesamten Kohlendioxids aufgenommen, das der Mensch in den vergangenen 150 Jahren emittiert hat. Wenn wir diesen Prozess richtig verstehen, dann erweitern wir nicht nur unser Wissen über das Erdsystem, sondern können auch Vorhersagen über den Klimawandel verbessern.

Text 2

In der durch einen Ausbruch des Vesuvs 79 n. Chr. verschütteten Stadt Pompeji wird bereits seit dem 18. Jahrhundert gegraben. Entsprechend gibt es bereits einen großen Wissensschatz. Trotzdem ergeben sich immer wieder spannende neue Fragestellungen. Ich untersuche die Inschriften, die auf den Häusern zu finden sind und schaue mir dabei an, wie die Inschriften im Kontext des öffentlichen Raums gewirkt haben. Dafür grabe ich nicht selber vor Ort, sondern arbeite mit dem bereits freigelegten Material. Insgesamt werden alternativ zu Grabungen immer mehr nicht-invasive Verfahren genutzt, etwa Begehungen der Oberfläche. In Pompeji sind viele Inschriften nach ihrer Ausgrabung leider nicht erhalten geblieben, da der Putz, auf dem sie gemalt waren, zerfallen ist, weshalb ich manchmal auf alte Fotos zurückgreifen muss. Doch viele Informationen kann ich nur vor Ort sehen, etwa den Ort der Inschrift und wie sie in der Stadt sichtbar war. Es gibt eine ganze Vielzahl an alltäglichen, gemalten Spuren: von der praktischen Anweisung, an dieser Stelle keinen Unrat abzuladen, über Wahlaufrufe bis zu Ankündigungen von Veranstaltungen. Entsprechend unterschiedlich sind die Inschriften. Besonders interessant sind für mich zum Beispiel die Inschriften an den Begräbnisstätten vor der Stadt, den sogenannten Nekropolen. Auch hier finden sich Hinweise auf Veranstaltungen und die damaligen Wahlplakate. Damals waren Friedhöfe ein Ort der Kommunikation, was heute in dieser Form gar nicht vorstellbar ist. Für die Feldforschung bin ich meist zwei bis drei Wochen vor Ort und sammle Daten. Das erfordert immer eine gezielte Vorbereitung, da ich den Zugang – insbesondere zu den nicht öffentlichen Bereichen der Stadt – beantragen muss. Zudem arbeite ich oft im Archäologischen Nationalmuseum in Neapel, wo viele Putzstücke mit Wandmalereien und Inschriften im Magazin eingelagert sind. Während der Pandemie habe ich nun vor allem mit den gesammelten Daten gearbeitet, aber tat sachlich musste ich dringend wieder hinfahren – auch um geplante Folgeprojekte vorzubereiten. Wenn es wieder möglich ist, werden wir auch wieder mit Studierenden auf Exkursionen rund um das Mittelmeer fahren, denn solche Studienreisen sind ein wichtiger Teil der Ausbildung. Man merkt immer wieder, dass es einen unglaublichen Mehrwert bringt, vor Ort zu sein und mit dem Material konfrontiert zu werden. Das kann man zu Hause, in der Bibliothek oder im Seminar nicht simulieren.

Text 3

Wie werden in Krisensituationen Hilfsmaßnahmen koordiniert, bei denen es in einer unübersichtlichen Lage auf schnelles Handeln einer Vielzahl von Akteurinnen und Akteuren ankommt? Zu dieser Frage forsche ich seit mehr als zehn Jahren mit meinem Team. 2017 flüchteten innerhalb eines Jahres 1,2 Millionen Menschen aus dem nordostafrikanischen Südsudan vor dem Bürgerkrieg ins benachbarte Uganda. Dort mussten sie schnellstmöglich mit dem Nötigsten versorgt werden. Eine der zentralen Kräfte im Hilfseinsatz war das Ugandische Rote Kreuz, mit dem wir bis heute eng zusammenarbeiteten. Ich hatte damals einen privaten Kontakt, sodass wir relativ spontan nach Uganda geflogen sind, um die Hilfskräfte bei diesem Notfalleinsatz zu begleiten. Entscheidend für unsere Forschung ist, dass wir zwar dabei sind, aber durch unsere Teilnahme möglichst wenig Einfluss auf das Geschehen nehmen. Natürlich interessieren uns auch die offiziellen Verfahrensanweisungen und wir fuhren begleitende Interviews. Vor allem aber geht es uns darum, zu verstehen, wie in der Praxis konkret gehandelt wird. Das heißt, man muss möglichst viel Zeit mit den Leuten verbringen und Vertrauen aufbauen. Daher sind wir immer mindestens vier bis sechs Wochen am Stuck vor Ort – seit 2017 insgesamt schon sieben Mal, auch während der Ebola-Epidemie in den Jahren seit 2018. Wir sind meist in verschiedenen Gegenden des Landes unterwegs, wohnen so wie die Einsatzkräfte und fahren mit zu den Einsätzen. Uns ist es sehr wichtig, die Arbeit nicht zu behindern, indem wir zum Beispiel wertvolle Ressourcen wie einen eigenen Jeep beanspruchen. Ich werde manchmal gefragt, ob es nicht gefährlich ist, in Uganda und dann auch noch während einer Krise Feldforschung zu betreiben. Aber wir sind dort mit Profis unterwegs, die die Lage sehr gut einschätzen können und wir gehen keinerlei Risiken ein. Im März 2020 waren wir auch schon auf dem Weg nach Uganda, mussten dann aber wegen kurzfristig geschlossener Grenzen direkt wieder zurückfliegen. Seitdem sind wir mit den Mitarbeitenden vor Ort regelmäßig über Messengerdienste in Kontakt und haben im vergangenen Jahr auch drei digitale Workshops durchgeführt. Aber natürlich ist unsere Forschung darauf angewiesen, dass wir die Leute im Einsatz begleiten. Wenn es wieder möglich ist, werden wir hoffentlich schnell nach Uganda reisen können. Denn das machen aus meiner Sicht viel zu wenig Forschende: sich anschauen, was wirklich getan wird.

Text 4

Für meine Masterarbeit habe ich vier Monate lang in einem Stadtviertel von Neu-Delhi gelebt, das von tibetischen Geflüchteten gegründet wurde. Dort habe ich zu der Frage gearbeitet, was einen Ort ausmachen muss, um ihn als ein Zuhause zu empfinden. Auch für meine Promotion wollte ich deshalb nun wieder nach Indien. In einem Teilprojekt des Exzellenzclusters CLICCS untersuchen wir, wie Menschen in verschiedenen Regionen der Welt über die Zukunft des Klimas denken, welche Entwicklungen sie für denkbar halten und wie sie entsprechend dieser Vorstellungen handeln. Ich wollte vergangenes Jahr nach Darjeeling reisen und dort zwölf Monate lang Teebäuerinnen und Teebauern auf den Plantagen begleiten. Am Tee lassen sich Klimaveränderungen gut ablesen, denn der Geschmack verändert sich und die Ernte ist durch die zunehmenden Wetterextreme rückläufig. Das wäre ein sehr interessantes Feld gewesen, doch die Pandemie hat alles gestoppt. Für ethnologische Forschung muss man vor Ort sein, Beziehungen knüpfen und Vertrauen aufbauen – und das ist in Indien auf absehbare Zeit nicht möglich. Daher habe ich im Herbst 2020 mein Projekt umgestellt – von Indien in den Harz. Diese Region ist besonders stark vom Waldsterben betroffen und die Entscheidung, was man heute pflanzt, wird die kommenden Generationen stark beeinflussen. Im ländlichen Sachsen-Anhalt nutze ich dabei die gleichen Methoden wie in Indien, etwa die teilnehmende Beobachtung. Dabei begleite ich meine Forschungspartnerinnen und -partner in ihrem Alltag und beteilige mich an ihren Aufgaben. In diesem Frühjahr war ich bereits für einen Monat in der Nähe von Wernigerode, um das Projekt vorzubereiten, und im Juni konnte ich zu einem längeren Forschungsaufenthalt starten. Ich bleibe voraussichtlich zehn Monate und arbeite unter anderem eng mit Försterinnen und Förstern sowie Mitarbeitenden des Nationalparks zusammen. Ich treffe sie vor allem im Wald, doch das hatte ich auch ohne Pandemie so machen wollen, da das ja ihr Arbeitsort ist. So möchte ich herausfinden, welche Rolle Wald- und Klimawandel für die Menschen in der Region spielen. Dass es mit Indien nicht geklappt hat, finde ich sehr schade. Aber an der ethnologischen Forschung mag ich am liebsten, dass man so eng mit Menschen zusammenarbeitet – und das kann ich auch im Harz tun. Wenn die Alternative wäre, auf unabsehbare Zeit digital arbeiten zu müssen, ist das für mich so in Ordnung.

A2.2 Reading Task in VR – Text Questions

For measuring the information *Recall* of the content in the text that participants had to read during the study, three content questions were asked for each text. For measuring the *Spatial Recall*, the participants were also asked to mark the text row on a paper printout where the answer to the content question was located.

Content Questions Text 1

1. Aus welchem Ozean konnten die Forschenden die Messinstrumente bergen?
2. Wie lautet der Name des Forschungsschiffs?
3. Durch welche Art von Signal wurden die Messinstrumente vom Meeresboden gelöst?

Spatial Recall Questions Text 1

1. An welcher Stelle des Textes wurde erwähnt, aus welchem Ozean die Messinstrumente geborgen wurden?
2. An welcher Stelle des Textes wurde erwähnt, wie der Name des Forschungsschiffs lautet?
3. An welcher Stelle des Textes wurde erwähnt, durch welche Art von Signal die Messinstrumente vom Meeresboden gelöst wurden?

Content Questions Text 2

1. Welcher spezielle Ort war im damaligen Pompeji ein Ort der Kommunikation?
2. Wodurch wurde die Stadt Pompeji 79 n. Chr. verschüttet?
3. In welcher Stadt liegt das archäologische Nationalmuseum, in dem die Forscherin ebenfalls arbeitet?

Spatial Recall Questions Text 1

1. An welcher Stelle des Textes wurde erwähnt, was im damaligen Pompeji ein Ort der Kommunikation war?
2. An welcher Stelle des Textes wurde erwähnt, wodurch die Stadt Pompeji verschüttet wurde?
3. An welcher Stelle des Textes wurde erwähnt, in welcher Stadt das archäologische Nationalmuseum liegt, in dem die Forscherin arbeitet?

Content Questions Text 3

1. Wie bleiben die Forschenden mit den Mitarbeitenden vor Ort in Kontakt, solange sie selbst nicht reisen können?
2. Wie oft waren die Forschenden seit 2017 in Krisengebieten vor Ort?
3. Aus welchem Land flüchteten 2017 1,2 Millionen Menschen nach Uganda?

Spatial Recall Questions Text 3

1. An welcher Stelle des Textes wurde erwähnt, wie die Forschenden mit den Mitarbeitenden vor Ort in Kontakt bleiben?
2. An welcher Stelle des Textes wurde erwähnt, wie oft die Forschenden seit 2017 in Krisengebieten vor Ort waren?
3. An welcher Stelle des Textes wurde erwähnt, aus welchem Land 2017 1,2 Millionen Menschen nach Uganda flüchteten?

Content Questions Text 4

1. Welche Personen wollte die Forscherin für ihre Promotion im vergangenen Jahr in Indien bei der Arbeit begleiten?
2. Was mag die Forscherin an der ethnologischen Forschung am liebsten?
3. In welches Bundesland hat die Forscherin ihr Projekt verlegt?

Spatial Recall Questions Text 4

1. An welcher Stelle des Textes wurde erwähnt, wen die Forscherin fur ihre Promotion im vergangenen Jahr in Indien begleiten wollte?
2. An welcher Stelle des Textes wurde erwähnt, was die Forscherin an der ethnologischen Forschung am liebsten mag?
3. An welcher Stelle des Textes wurde erwähnt, in welches Bundesland die Forscherin ihr Projekt verlegt hat?

ASSISTIVE CONTROLLER-BASED RAYCAST REDIRECTION

A3.1 User Study Results – NASA RAW TLX

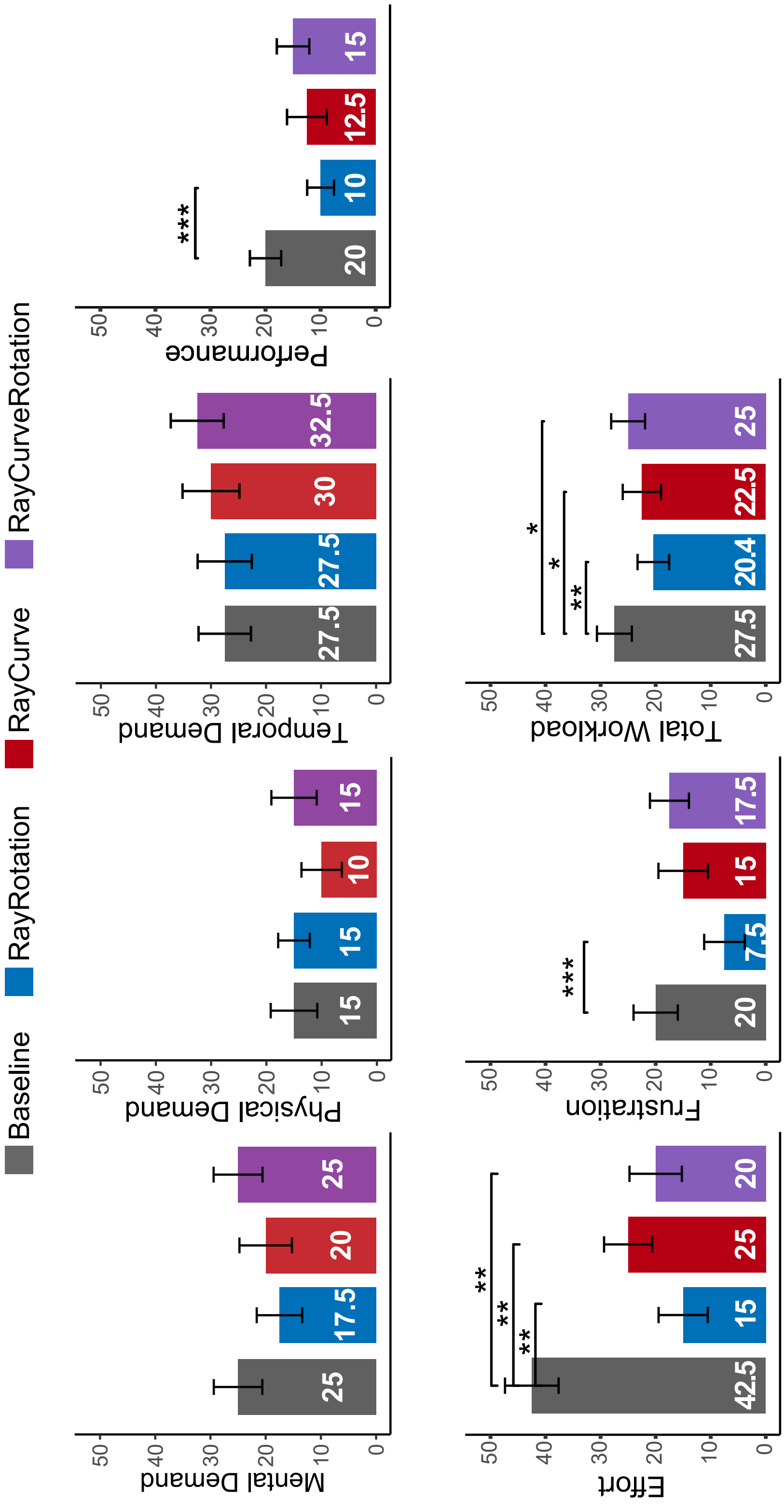


Figure A3.7: Nasa RAW TLX with median and standard error bars – all subscales and Total Workload (significance levels: * < .05 , ** <.01, *** < .001)

A4.1 User Study Results – NASA RAW TLX

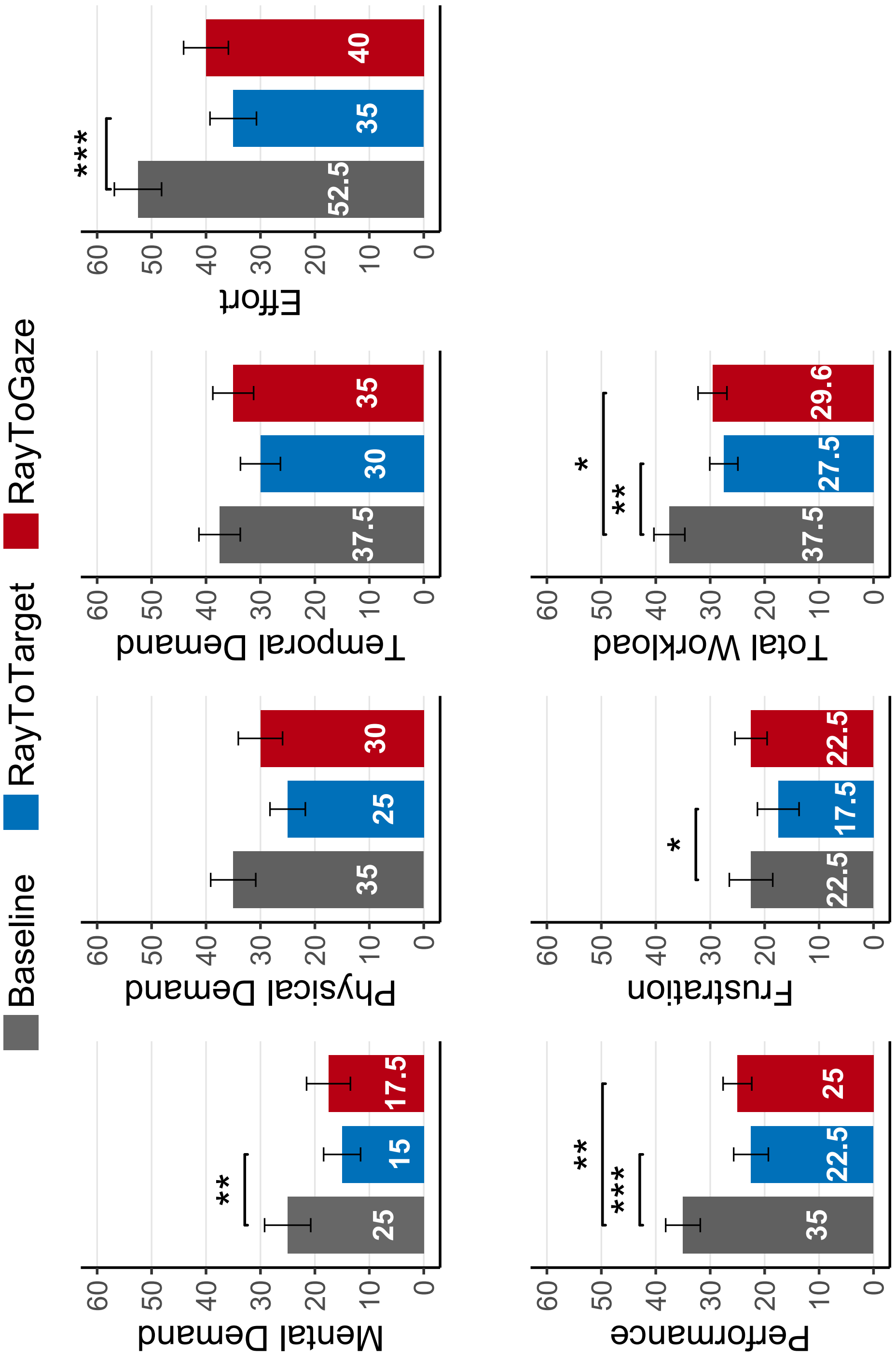


Figure A4.8: Nasa RAW TLX with median and standard error bars – all subscales and Total Workload (significance levels: * < .05 , ** <.01, *** < .001)

DECLARATION

UHHGPT [4] is a web tool operated by the Center for Sustainable Research Data Management and the Regional Computing Center of the University of Hamburg, which provides institutional access to the OpenAI API of ChatGPT (versions GPT4 omni and GPT4 omni mini). Following the ACM guidelines [5], I used UHHGPT to rephrase sentences for improving the writing style, clarity and readability of my texts. UHHGPT helped improve the overall writing quality and coherence of the dissertation, while all original research and analysis remain my own.

In addition, UHHGPT and DeepL Translator Pro [6] were used for translations from German into English. LanguageTool Premium [7] was used check and improve the grammar, spelling, punctuation, and clarity of the texts.

Sample prompts for UHHGPT are listed below.

- Please help me improve the writing style and clarity for this sentence.
- Please rewrite and shorten the following text to improve the clarity and readability.
- Please rephrase the following text to improve the clarity and coherence.
- Please translate this German text into English.

[4] https://uhhgpt.uni-hamburg.de
[5] https://www.acm.org/publications/policies/frequently-asked-questions
[6] https://www.deepl.com/en/translator
[7] https://languagetool.org/

EIDESSTATTLICHE VERSICHERUNG – AFFIDAVIT

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Sofern im Zuge der Erstellung der vorliegenden Dissertationsschrift generative Künstliche Intelligenz (gKI) basierte elektronische Hilfsmittel verwendet wurden, versichere ich, dass meine eigene Leistung im Vordergrund stand und dass eine vollständige Dokumentation aller verwendeten Hilfsmittel gemäß der Guten wissenschaftlichen Praxis vorliegt. Ich trage die Verantwortung für eventuell durch die gKI generierte fehlerhafte oder verzerrte Inhalte, fehlerhafte Referenzen, Verstöße gegen das Datenschutz- und Urheberrecht oder Plagiate.

Hamburg, May 20, 2025

(Jenny Gabel)