OPTIMIZING PERFORMANCE AND SATISFACTION IN VIRTUAL REALITY ENVIRONMENTS WITH INTERVENTIONS USING THE DATA VISUALIZATION LITERACY FRAMEWORK

Andreas Bueckle

Submitted to the faculty of the University Graduate School in partial fulfillment of the requirements for the degree Doctor of Philosophy in the Luddy School of Informatics, Computing, and Engineering, Indiana University

July 2021

Accepted by the Graduate Faculty, Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Doctoral Committee

Katy Börner, Ph.D.

Patrick C. Shih, Ph.D.

John A. Walsh, Ph.D.

Robert D. Montoya, Ph.D.

April 19, 2021

To Mama

To Aga

To Alex

To Chris

ACKNOWLEDGMENTS

First and foremost, I want to express my deep gratitude to my advisor Katy Börner for hundreds, if not thousands of hours of mentorship, direct and indirect, and for her courage to challenge and be challenged together.

Second, I would like to say a heartfelt "Thank You!" to my research committee, consisting of Patrick Shih, John Walsh, and Rob Montoya, for all the guidance, feedback, and support.

Third, I gratefully acknowledge support from the following entities and people: the Department of Information & Library Science (ILS) for the ILS Leadership Award; the Cyberinfrastructure for Network Science Center (CNS) and the Department of Intelligent Systems Engineering (ISE) in the Luddy School of Informatics, Computing, and Engineering (SICE) at Indiana University for my employment as a research assistant and associate instructor as well as the ability to be part of large scientific endeavors; my stellar colleagues at CNS and in SICE; the German Academic Exchange Service (Deutscher Akademischer Austauschdienst, DAAD) for my graduate scholarship as a non-degree student in 2014/2015; and the Advanced Visualization Lab (AVL), especially Chauncey Frend, for tech support and the encouragement to pursue virtual reality as a research topic. Also, I would like to thank Klaus Gasteier at Berlin University of the Arts, who introduced me to interaction design while I was pursuing my Master's degree prior to coming to IU. My thanks also go to Philip Beesley for allowing me to use his 3D model of Luddy Hall in my dissertation, and for the opportunity to collaborate with the world-class artists in his team.

Fourth, I would like to acknowledge the contributions of co-authors, co-creators, and colleagues. Like most interdisciplinary work, the research presented here is the product of collaboration between people with different skills. I gratefully acknowledge my co-authors Katy Börner and Michael Ginda for their work on the Data Visualization Literacy Framework (32), on which I built the theoretical core of this dissertation. Chapters 5 and 6 were co-written with Kilian Buehling, Patrick Shih, and Katy Börner, a great collaboration for which I express my deepest gratitude. Further, JangDong "JD" Seo from the Indiana Statistical Consulting Center (ISCC) provided his expert input on power analysis for determining the required number of subjects for the user studies presented in chapters 5, 6, and 7. Robert L. Goldstone assisted in developing the initial research design for chapter 5. Leonard E. Cross supported the development of the three setups tested in chapter 5 through active play-testing and discussions. Kristen Browne (Bioinformatics and Computational Biosciences Branch, Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases, National Institutes of Health) created the 3D model of the kidney used in chapters 5 and 6. Todd Theriault performed copy-editing and proof-reading for chapter 5. Perla Brown improved the figures in chapter 5. Eight colleagues at CNS served as pilot testers for the virtual reality (VR) setups developed for this dissertation. Lisel Record, Esmé Middaugh, and my brother Alexander Bueckle graciously served as friendly reviewers for my dissertation proposal in 2020. Further, I would like to acknowledge Sanjay Jain, M.D., Ph.D., and Diane Salamon, R.N., from Washington University in St. Louis School of Medicine for providing the pictures of dissected kidneys.

Last but not least, I would like to thank my fellow sailors at the Sailing Club at Indiana University, where I had the honor of serving as Coach, Captain, and Marketing Chair throughout my graduate school experience at IU. Being on the water has always kept me full of energy for the road ahead.

PREFACE

Throughout my educational career, I have admired the transformative power of mediated experiences, or the fact that we can get excited by seeing shadows on the wall of a cave (or, in modern times, colorful pixels on a screen). Challenged and highly excited by “You Are Not A Gadget” (123), a provocative reflection on the internet, modern media, and their influence on society by VR pioneer Jaron Lanier, I found information science and interaction design as a perfect place to set up my new inner intellectual workshop. When I experienced proper VR for the first time at IU, I had a glimmer of an understanding what those people in the late 19th century must have felt when seeing that famous first film record of a train arriving at a station (174). According to an urban legend, when people watched this clip first the first time, many believed the train would dash out of the screen at any moment, simply because they had never experienced this kind of medium before. Being enchanted by the possibilities of a new medium never loses its power, whether it is in 1895 or in 2016. Combining VR with data visualization then, a field with rich theoretical and applied history, seemed like a dream project for a dissertation that is both theoretically sound and practically valuable, and that can help address real-world issues in a world increasingly reliant on our human ability to analyze data, identify patterns, and spot trends. I am grateful that this dissertation worked out the way it did, allowing me to combine data visualization for solving practical problems with the strength of the mediated experience that is VR.

While collecting the data for this dissertation, I had the privilege of watching many dozens of people move and act in VR for the first time. While the VR applications described in this dissertation – owing to the nature of this research – lack the graphical fidelity, action, and larger-than-life environments of video games or similar products of entertainment, I witnessed that VR has an impact almost regardless of content. With immersive media becoming increasingly cheap, immersive content becoming ever more widespread, and producing immersive content getting easier by the day, I cannot wait to see where we can take the emerging field of immersive data visualization.

As a tinkerer, I cannot wait to continue building.

Andreas Bueckle

Friday, July 16, 2021

Andreas Bueckle

OPTIMIZING PERFORMANCE AND SATISFACTION IN VIRTUAL REALITY ENVIRONMENTS WITH INTERVENTIONS USING THE DATA VISUALIZATION LITERACY FRAMEWORK

In the age of big data, interactive data visualizations are becoming increasingly prevalent. The ability to focus on subsets of data is essential for exploring large temporal, geospatial, topical, and network datasets. In order to categorize all parts of data visualizations, various attempts have been made to build frameworks for how to interpret, teach, and construct data visualizations while turning data into insights. This need for interactive data visualizations has sparked interest to classify key interaction types. The recent rise of affordable virtual reality (VR) has introduced yet more ways of interacting with datasets using our visual abilities while enabling user input beyond mouse and keyboard.

In this dissertation, we apply the Data Visualization Literacy Framework (DVL-FW) in a series of three interconnected VR user studies with 152 subjects representing around 123 hours of face-to-face data collection. In the first study, we compare performance and satisfaction across two VR and one desktop implementation of the same 3D manipulation interface. We found that while VR users are about three times as fast and about a third more accurate in terms of rotation than desktop users, there are no significant differences for position accuracy. Building on this experiment, in the second study, we investigate quantitative differences between two user cohorts, where the experiment cohort gets to inspect their own manipulation performance data between trials in a VR setup and with a traditional 2D line graph, depending on their assigned setup ("Reflective phase"). Our findings indicate that while there is no difference in performance between VR users across cohorts, the Reflective phase yields significant differences for desktop users and increases the satisfaction for VR users. Moreover, we identified behavioral metrics for VR users in the Reflective phase that have a favorable effect on performance in subsequent trials. Finally, in the third study, we asked users to travel to various points inside a virtual 3D model of Luddy Hall on the Indiana University campus in Bloomington, IN. We then tested whether the experiment cohort was able to devise faster movement strategies after a Reflective phase where they inspected their own navigation data in a VR visualization with a miniature model of the building. We found that users with a Reflective phase in VR have significantly faster completion times in the second set of trials than those who did not, while also scoring significantly higher on a mid-questionnaire about the topology of the virtual building.

Our methodology combines quantitative and qualitative surveys with a VR software and hardware solution that can be used in future human-subject studies by others. In addition to contributing to the theory of data visualization, specifically the interaction typology of the DVL-FW, this dissertation provides evidence-based design recommendations for matching and movement tasks in VR and on desktop devices by comparing task completion time, accuracy, and user satisfaction across different implementations of the same application. Further, we derive design guidelines for interventions using data visualizations for subjects to reflect on their behavior in VR to improve performance and satisfaction metrics in future trials.

Table of Contents

List of Tables

List of Figures

1. Introduction

As the amount of data displayed to us in everyday life is increasing by several orders of magnitude every year, interaction with data has become an ever more important task for gaining insights from data visualizations. As early as the beginning of this century, German computer scientist Daniel Keim (117) cited a UC Berkeley estimate whereby every year, about 1 million terabyte of data were created. At the time of writing, this was almost two decades ago in 2002. He notes elsewhere (118):

"Effective data mining depends on having a human in the data exploration process while combining this person's flexibility, creativity, and general knowledge with the enormous storage capacity and computational power of today's computers. Visual data exploration seeks to integrate humans in the data exploration process, applying their perceptual abilities to the large data sets now available. The basic idea is to present the data in some visual form, allowing data analysts to gain insight into it and draw conclusions, as well as interact with it." (p. 39)

Data visualization thus allows humans and computers to work together, connected by an empowering user interface (UI). When analyzing data, the visual acuity, ability to recognize patterns and trends, and spatial memory of an analyst make a powerful addition to the computational capabilities and speed of a computer system. Visualizations then allow the user to not only see the data but also interact with it to unearth more insights than are initially visible. This sentiment was mirrored years earlier in Ben Shneiderman's famous report aptly titled "The Eyes Have It" (176) where the author identifies the visual primacy of humans when it comes to parsing information. Many modern data visualization software packages contain native functionality to manipulate graphs, select subsets of data, or annotate them on the spot. While the standard paradigm for data visualization has long been the printed chart and then, with the rise of the personal computer, the 2D monitor plus mouse and keyboard, newer media for deployment have been researched in recent years, including but not limited to, large display walls, mobile phones, tablets, augmented reality (AR), and VR. VR offers new and exciting ways of interacting with virtual worlds as a whole, and elements of graphs, charts, maps, networks, and tree visualizations specifically. While VR as a technology has existed for several decades, it is only over the past few years that we have seen affordable, portable, and comparatively high-spec VR hardware enter the market, along with content as well as the necessary software development kits (SDKs) to allow single users without extensive expertise in computer graphics to produce their own VR content. While the future of VR as a mass medium is all but clear (99), there exist diverse domain applications with data visualizations in which users are immersed in virtual models, look at data as if it were right in front of them, and use their innate visual acuity and fine motor skills to turn data into insights. Indeed, VR has been successfully applied in scientific visualization for a long time (42). As a result, data visualizations developed for and deployed in VR have received significant scholarly attention (19, 23, 55, 60-62, 75, 112, 134, 207, 208).

1.1 Expanding the Data Visualization Literacy Framework

The need for making sense of ever-larger amounts of data has led to a veritable "zoo" of visualizations (102) (p. 60) to describe statistical, temporal, geospatial, topical, and network datasets. In order to categorize and describe the animals in the "zoo", many frameworks have been developed, including the Data Visualization Literacy Framework (DVL-FW), described in depth in section 2.1. It contains seven typologies with terminology to describe a variety of data visualizations from the omnipresent, simple, static scatter graph, to the visually more complex stacked bar graph, to dynamic, real-time, 3D dot density maps and network visualizations. These DVL-FW typologies provide language, retrieved and synthesized from prior literature, to construct, interpret, and teach data visualizations. One of these seven typologies is the interaction typology with nine types as introduced by Börner, Bueckle and Ginda (32) and Börner (31), see Table 1 below. In section 2.1.3, we provide an overview of previous use cases for the DVL-FW. The user studies in this dissertation describe the first application of the DVL-FW to interactive data visualizations in VR.

Table 1. Interaction types currently captured in the DVL-FW.

Interactions
Zoom
Search and locate
Filter
Details on demand
History
Extract
Link and brush
Projection
Distortion

Like the other typologies in the DVL-FW, these interaction types can be used for a variety of purposes: analyzing and critiquing existing visualizations; teaching data visualization literacy; assessing the usability of data analysis algorithms, visual encoding schemes, or interaction techniques; and developing new visualizations. In this dissertation, we propose an extension of the interaction typology from nine to 24 types based on an extended survey of related work from the fields of information visualization, human-computer interaction, informatics, geography, computer science, and statistics (see Table 2). We compiled brief definitions and links to prior work for all 24 interaction types below. Additionally, we reviewed existing interaction typologies in section 2.2. Sources for all types listed here can also be found in Table 4.

Below, we provide descriptions for the interaction types added to the DVL-FW interaction typology in this dissertation, and add explanatory language based on the DVL-FW (32) with quotes from selected publications defining the interaction type. Please note that it is outside of the scope of this dissertation to provide any categorizati1ons or groupings for these types, like the proposed by others (103). Identifying which interaction types have a natural fit with each other in an internal framework for the DVL-FW interaction typology should be the focus of future work.

Aggregate: "concerns methods that change the granularity of visualization elements" (40). Also refers to collecting many units into one. For example, data can be aggregated by ethnicity.

Animate/replay: when data has a temporal dimension, graphic symbols can be displayed based on a time stamp or time value. This allows the user to view data traces of events unfolding in real-time or at a range of playback rates, allowing them to better identify patterns and trends over time. Says Wilkinson (215):
"Graphics can be animated over variables intrinsic or extrinsic to the graph [...]. More importantly, the user should have control over the frames and be able to pause, move forward, or move backward in the animation at will. This can be accomplished with a pause button and a slider that the user can move where each tick on the slider corresponds to a single frame."

Annotate: "allow textual [or graphical] annotation of states within a visual history" (103). Also refers to adding explanatory notes or comments to a visualization.

Arrange/coordinate views: "enable analysts to see their data from different perspectives" (103). Also refers to placing individual views or windows in a visualization either completely manually or assisted by an algorithm.

Derive: "compute new data elements given existing data elements" (40), "summarize the input data, ranging from descriptive statistics […] to model fitting […] and data transformation […]" (103). In our interpretation, this refers to creating variables based on existing ones in the original dataset.

Details on demand: "Select an item or group and get details when needed" (176). This can refer to entire data records or graphical elements representing one or multiple variables.

Distortion: "show portions of the data with a high level of detail and other portions with a lower level of detail" (118).

Extract: "Allow extraction of sub-collections and of the query parameters." (176). This can refer to data records and variables but also parameter and widget settings.

Filter: "focusing on specific information within a graphic. It is usually helpful to see a certain graphic under a set of constraints that are defined either by categories or ranges of continuous values" (215).

Highlight: "focus on certain data points (objects) by giving users the ability to change the appearance of object groups in real-time" (57). This refers to the appearance of graphic symbols and graphic variables.

History: "Keep a history of actions to support undo, replay, and progressive refinement." (176). This refers to the iterative improvement of the entire data visualization.

Link and brush: "coloring or otherwise highlighting a subset of the data [...and] showing information about the highlighted subset in other views" (44).

Manipulate: "manipulate object set parameters through object handles. Direct manipulation refers to operating directly on objects instead of through menus or dialogues." (57). Specifically, we consider adjusting position, rotation, and scale of objects (graphic symbols) to be manipulation by interacting with graphic symbols.

Navigate: "visualizations often function as viewports onto an information space. Analysts need to manipulate these viewports to navigate the space", "One common pattern of navigation adheres to the widely cited visual information-seeking mantra: 'Overview first, zoom and filter, then details-on-demand'" (Heer and Shneiderman (103), citing Shneiderman (176)).

Overview: "Gain an overview of the entire collection" (176). We understand this as the ability to quickly get a bird's-eye view of an entire dataset with the option of having context + focus tools.

Pan: "change the geographic center of the cartographic representation" (163), "shift the start of the value range to be shown" (212).

Projection: "set or change the cartographic projection used for the cartographic representation" (163); another good example are "parallel coordinates" (118). We understand this as the ability to adjust the reference system of the visualization from, e.g., a Cartesian to a polar or parallel coordinate system.

Record: "To support iterative analysis, visual analysis tools can record [i.e., capture] and visualize analysts' interaction histories" (103), "save or capture visualization elements as persistent artefacts" (40).

Relate: "View relationships among items" (176). We understand this as the ability to turn links in a visualization on and off, or manually create new links.

Search and locate: Neither source for this type provides a definition (31, 32). We consider Search and Locate the activity of identifying a data record by using phrases in a search interface, ranging from simple words to complex queries. In the case of Search and locate, the user already knows exactly what item they are looking for.

Select: "demarcation of one or more elements in a visualization, differentiating selected from unselected elements" (40).

Sort: "Ordering [...] is another fundamental operation within a visualization [...]. The most common method of ordering is to sort records according to the value of one or more variables" (103).

Visualize/encode: "show me a different representation" (218), "specify a visualization of data: analysts must indicate which data is to be shown and how it should be depicted" (103). We understand this as the ability to adjust the mapping of data variables to graphic variables of graphic symbols, either for the whole visualization or just a subset of data. It can also mean the change of the visualization type.

Zoom: "change the scale and/or resolution of the cartographic representation" (163), "Through zooming, users can simply change the scale of a representation so that they can see an overview of a larger data set (using zoom-out) or the detailed view of a smaller data set (using zoom-in). A key point here is that the representation is not fundamentally altered during zooming" (218). This can also mean "semantic zoom" where the granularity of the data on display is adjusted rather than the visual representation of that data.

Table 2. Proposed expanded interaction typology of the DVL-FW. Interaction types previously not in the DVL-FW interaction typology are bold. A more detailed alignment of interaction types across frameworks is shown in Table 4.

Interaction types identified in literature	**Highlighting**	**Relate**	**Record**	**Manipulate**	**Aggregate**	Search and locate	**Animate /replay**
Captured by	**(57, 73, 211)**	**(176, 218)**	**(40, 176)**	**(57, 215)**	**(40, 73)**	(31, 32)	**(215)**
Applied in chapter 6							**Animate /replay**
Applied in chapter 7							

Interaction types identified in literature	Details on demand	**Annotate**	**Arrange/ coordinate views**	Projection	History	Extract	**Overview**	**Derive**
Captured by	(31, 32, 73, 176, 201)	**(40, 57, 103, 197, 215)**	**(40, 44, 103, 163, 201)**	(31, 32, 117, 118, 163)	(31, 32, 176, 215)	(31, 32, 163, 176)	**(31, 176)**	**(40, 103, 163)**
Applied in chapter 6								
Applied in chapter 7								

Interaction types identified in literature	Filter	Zoom	Link and brush	Distortion	**Sort**	**Visualize/encode**	**Select**	**Pan**	**Navigate**
Captured by	(31, 32, 40, 57, 73, 103, 117, 118, 163, 176, 197, 211	(31, 32, 44, 73, 117, 118, 163, 176, 201, 211, 212	(31, 32, 44, 57, 73, 117, 118, 201, 211, 215)	(31, 32, 73, 117, 118, 197, 211, 212, 215)	**(73, 103, 163, 197, 211, 215, 218)**	**(40, 73, 103, 163, 201, 211, 218)**	**(40, 73, 103, 201, 211, 212, 215)**	**(44, 73, 163, 211, 212, 215)**	**(40, 103, 197, 211, 212, 215)**
Applied in chapter 6	Filter								**Navigate**
Applied in chapter 7	Filter		Link and brush						**Navigate**

In the user studies in this dissertation, we gave 152 subjects across these studies various cube-matching (*RUI* and *RUI (VR) Reflective* user studies, see chapters 5 and 6) and movement tasks (*Luddy (VR) study*, see chapter 7) while measuring performance for completion time (all studies) and accuracy (chapters 5 and 6 only). The user studies in chapters 5 and 6 constitute one dataset, but the results from chapter 5 are also presented on their own as the solution to an applied research problem. The subjects from the study then served as control cohort for the study in chapter 6. We tasked the experiment cohorts in the RUI VR Reflective and Luddy VR studies to inspect their performance using an interactive VR data visualization during a segment about half-way through the experiment, which we called the "Reflective phase." Specifically interested in expanding the DVL-FW interaction typology, we applied four interactivity types for interventions in VR using data visualizations in the Reflective phases of these two studies:

- **Filter** (already in the typology)

  Refers to the act of "focusing on specific information within a graphic. It is usually helpful to see a certain graphic under a set of constraints that are defined either by categories or ranges of continuous values" (215).
- **Link and brush** (already in the typology)

  Refers to "coloring or otherwise highlighting a subset of the data [...and] showing information about the highlighted subset in other views" (44).
- **Navigate** (newly added)

  Refers to the following: "visualizations often function as viewports onto an information space. Analysts need to manipulate these viewports to navigate the space", "One common pattern of navigation adheres to the widely cited visual information-seeking mantra: 'Overview first, zoom and filter, then details-on-demand'", quote by Heer and Shneiderman (103), citing Shneiderman (176).

- **Animate/replay** (newly added)

  "Graphics can be animated over variables intrinsic or extrinsic to the graph [...]. More importantly, the user should have control over the frames and be able to pause, move forward, or move backward in the animation at will. This can be accomplished with a pause button and a slider that the user can move where each tick on the slider corresponds to a single frame." (215)

1.2 Rationale for choosing these interaction types

Table 2 above indicates which interactivity type was implemented in which user study. We selected "**filter**" and "**link and brush**", because they are interaction types frequently mentioned in the literature reviewed for this dissertation (see section 2.2), and because both are already part of the DVL-FW interaction typology. Further, we consider these two types fundamental for decluttering large, spatial datasets, thus greatly helping making sense of visualizations in VR.

We selected "**navigate**", because we argue that it suits data visualizations in VR naturally, because it describes the act of moving around a dataset to inspect it from different angles, either through different windows (in a screen-based visualization) or, quite literally, through adjusting one's physical position regards to the dataset in VR. The "navigate" interaction type is thus specifically useful for geospatial insight needs. Because VR is unique suited for visualizing data that is already spatial by nature, navigate is a fitting interaction type for a visualization deployed to VR. Contrary to, e.g., viewing a scatter graph printed inside a book, viewing a visualization in VR is never the same from one moment to another. Even if the visualization itself is static, i.e., the graphic symbols that make up the visualization do not change any of their properties (such as position, size, or color), the user constantly “navigates” around it, i.e., changes their perspective and, thus, what is being shown to them at any given moment. Navigation is thus a natural interaction type already built into VR, without the need for designing a specific control scheme to move around a dataset (like one would for a keyboard and mouse interface).

Finally, we chose to implement “**Animate/replay**” to allow the users in our RUI VR Reflective study (see chapter 6) to inspect the temporal dimension of the dataset presented to them. Because animations are based on time and motion, and because the datasets generated in this study focused on completion time as well as position and rotation accuracy, this interaction type appeared to be well-suited to allow insights into the user’s understanding of temporal relationships over time. Note that we decided to give our users a different way of gaining insights into these relationships is the Luddy VR study (see chapter 7), where we visualized completion times in a bar graph instead.

1.3 Interaction types implemented in user studies in this dissertation

While we describe the implementation of these interaction types in detail in the corresponding chapters, with a brief, high-level overview of our user studies presented in chapter 4, we provide a brief summary for each type with examples of user interface (UI) elements that commonly accompany the interaction type below. The implementation of all four interaction types is explained in sections 6.3.3 and 7.5.

Filter:

- Type 1: Checkboxes (binary):

  Simple checkboxes (see Figure 1A) let the user turn parts of the data overlay on and off. Checkboxes correspond to a predetermined subset of the data on display, i.e., the designer has determined what data is affected by the interaction beforehand.

- Type 2: Time slider (range):

  A time slider (see Figure 1B) allows the user to show or hide data based on the time stamp associated with a data record. It offers more control about what kind of data the interaction affects. Furthermore, the time slider can be used for another interaction type: animate/replay, see below.

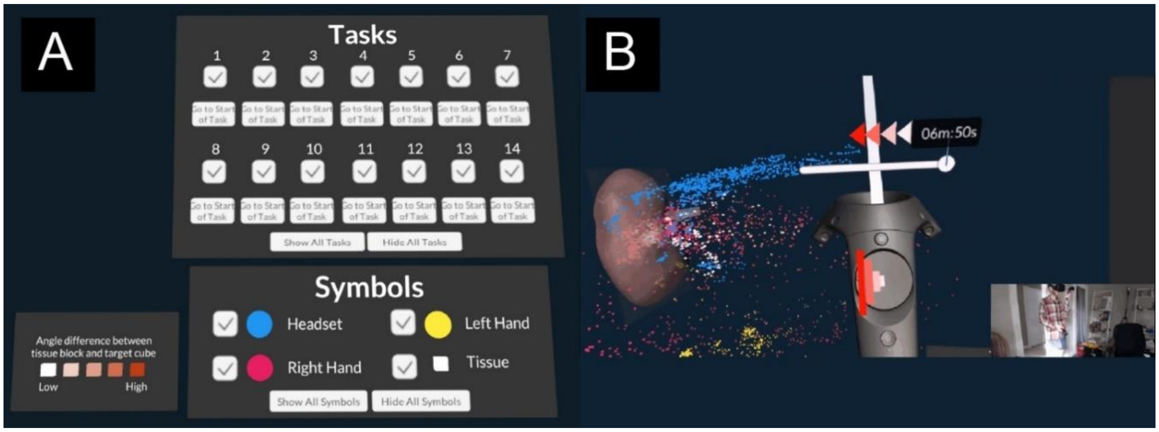

*Figure 1. **A**: checkboxes to turn parts of the data overlay on and off. **B**: time slider, adjusted via the user's input on the thumbpad of the VR controller.*

Animate/replay: Animations are a powerful interactive technique in data visualizations and can help the user see trends, uncover patterns, and understand transitions over time. Animate/replay allows the user to view events unfolding in real-time or at a custom playback speed, allowing them to better identify patterns and trends over time. This is different from filtering in that with a time filter, the user manually moves a time slider, system executes the user input, and the data vis changes one step at a time (similar for a time range filter). With animate/replay, the user controls not a time filter but a play head that allows them to go through the data in real-time or at any desired rate. When using the play head, the user can skip through the dataset at varying speeds, depending on their input (see Figure 1B). This can be observed in everyday life when watching, e.g., a YouTube video: the user can skim through the video to search for a particular time stamp by looking at the video frame displayed to them (time filter), or they can press the play button and watch the video in real-time (or at a faster/slower rate). The difference between these two interactions with a YouTube mirrors the difference between animate/replay and filter. While the animate/replay type is not part of the DVL-FW yet, it is mentioned in relevant literature and should be added to the interaction typology (215). The implementation of the animate/replay type is explained in chapter 6.

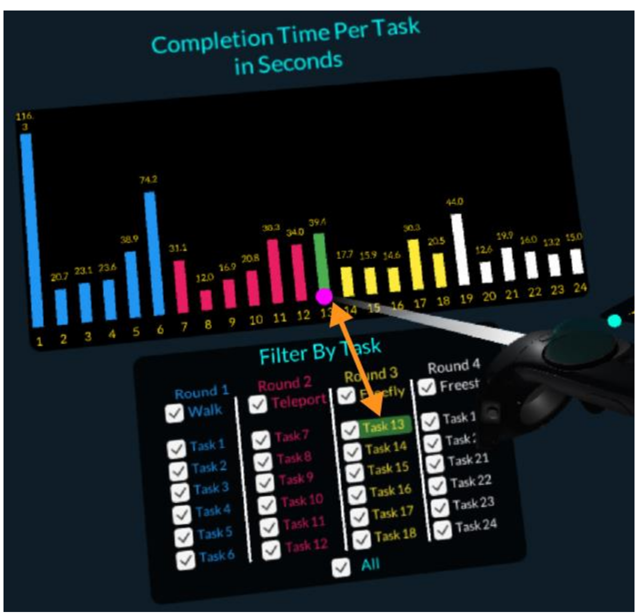

Figure 2. Link and brush example. When the user hovers over a bar in the bar graph of completion times, the corresponding checkbox for a task is highlighted, and vice versa.

Link and brush: This technique (also called coupled windows) highlights a subset of data in one visualization while the user hovers over or otherwise interacts with a data selection in another visualization or a UI element in a menu. It enables the user to detect connections between data subsets through providing an intuitive way of interacting with visualizations by virtue of menus, even in advanced, dynamic, multivariate visualizations such as the VR ones we use in the Reflective phase of our Luddy VR user study on navigation (see chapter 7). The presence of link and brush highlights another feature of our Reflective phase VR visualizations for Luddy VR: By bringing in a traditional, static bar graph of task completion times, we allow the user to explore a complex spatial dataset (their own) in VR while being able to retrieve metadata about their performance with the familiarity and proven effectiveness of the visual encoding of bar graphs, specifically, position and length (59, 101). The implementation of link and brush in this dissertation is explained in chapter 7.

Navigate: This technique can be implemented in different ways, depending on visualization type. It can be a general collection of other interaction techniques to traverse the continuum from overview to detail (103) or, more literally, to move around 2D and 3D virtual space to view a visualization from different angles (215). In the Reflective phase of our user studies, the ability to navigate is an inseparable part of being in VR where the user's body movements are tracked and the virtual camera is adjusted according to the user's head position and rotation. Navigation thus becomes an essential interaction technique for any VR visualization. It bears a somewhat surprising similarity to filtering in that it "moves about the graphic, while filtering moves around the data" (p. 555) (215).

1.4 Goal of this dissertation

While data visualization is a field rich with frameworks and other formalizations for how data visualizations are built, used, and taught, what visual encodings to use, and how user tasks can be defined (31, 87, 89, 146), interactions have not gotten the same attention as remarked elsewhere in the literature (81, 218). We thus propose an extension of the DVL-FW interaction typology and process model (31, 32, 35). The goals of this dissertation are thus:

- To expand the interaction types section of the DVL-FW to be inclusive of more interaction types essential to 3D and dynamic visualizations, specifically **animate/replay** and **navigate** (also **filter**, and **link and brush**).
- To produce **data-based design recommendations** for data visualizations in VR by comparing task completion time, accuracy, and satisfaction **in user studies** for various interaction types in VR compared to screen-deployed (Desktop) implementations of the same data visualization, and for improving performance metrics when completing matching and movement tasks in VR
- To develop **code** and **methodology** to conduct user studies in the game engine Unity 3D, a real-time 3D development platform (200)

The visualizations created for the user studies in this dissertation constitute the first time that data visualizations in VR were designed based on the DVL-FW, which, while able to characterize the data visualization process for static visualizations and a subset of dynamic, interactive visualizations, lacks the ability to describe similar work for VR. Expanding the DVL-FW will be carried out as follows:

Chapter 2 presents related work. We summarize the DVL-FW in its current form and propose an extension of its interaction typology in the form of the Data Visualization Literacy Interaction Component in section 2.1. Subsequently, in section 2.2, we review prior work a on interaction in data visualization in general (section 2.2.1) and present a synthesized view of interaction types (section 2.2.2). Next, we discuss related work for interaction types for VR specifically (section 2.3). Because this dissertation contains human-subject research, we then review VR user study methodologies (section 2.4). Over the course of chapter 2, we synthesize the findings of our literature review with a review of exemplars from the world of VR applications (such as video games) into a theoretical extension of the DVL-FW. This enhanced DVL-FW interaction typology forms the basis for our research questions (see chapter 3) as well as chapters 5, 6, and 7, where we present the design and results from our user studies. We conclude this dissertation with a Discussion containing recommendations, major challenges, and pointers to future in chapter 8.

The following elements provide additional information for the theoretical core of our contribution as well as the user studies presented:

A **Glossary** (p. 283) provides the definitions of domain-specific terms potentially not familiar to all readers.

High-resolution versions of all figures are available on GitHub (https://github.com/andreasbueckle/bueckle-dissertation-supporting-information/tree/main/high_res_figs).

Supporting Information (starting on p. 289) contains study information sheets, data collection instruments, and recruitment materials from the user studies. We added code from our experiments, written in C# for Unity, and videos demonstrating the VR setups and study procedures to the GitHub repository to illustrate how our studies were run and how subjects experienced the tasks and data presented to them (https://github.com/andreasbueckle/bueckle-dissertation-supporting-information).

2. Related Work

In this chapter, we review prior work, including the DVL-FW, interaction types in data and information visualization, both in VR and on traditional interfaces. Further, because this dissertation contains results from user studies, we review VR user study methodology in terms of the equipment used, sample sizes, research questions, and study instruments. We conclude this chapter with an overview of best-practice exemplars for interaction and interface design in commercial products such as video games.

2.1 The Data Visualization Literacy Framework

First, we provide an in-depth review of the DVL-FW, specifically the definitions it contains, the seven typologies, and prior use cases.

2.1.1 Definitions

The DVL-FW (32) has been developed to "define, teach, and assess [data visualization literacy]" (p. 1857). As data becomes increasingly prevalent in our everyday lives, skills relating to the understanding of trends, patterns, and structures of temporal, geospatial, topical, and network data are increasingly important for professional and personal decision-making. Unlike other literacy types such as numeracy (150, 158), textual literacy (151), or visual literacy (16, 18, 93, 100), data visualization literacy has seen formal assessment attempts only very recently. Boy et al. (39) proposed two tests for visualization literacy with line graphs using item response theory, successfully validating their model with a user study involving 40 subjects on Amazon Mechanical Turk (MTurk, https://www.mturk.com/). With similar methodological goals, Lee, Kim and Kwon (127) employ test development in psychology and education to develop the

Visual Literacy Assessment Test (VLAT), designed to measure how non-expert users interpret data visualizations. They validate their test, consisting of 12 data visualizations, 53 multiple-choice items, and eight visualization tasks, using input from five domain experts in data visualization and a user study with 191 MTurk subjects to show a high reliability.

Other work has been performed to characterize the sense-making process when facing data visualizations. Lee et al. (126) presented the NOVIS model to capture how users make sense of unfamiliar data visualizations (in this case, parallel coordinate plots, chord diagrams, and tree maps) using think-aloud sessions. They identify five steps in the sense-making process for unfamiliar visualizations: encountering visualization, constructing a frame, exploring visualization, questioning the frame, floundering on visualization, alongside miscellaneous activities. Similarly, Maltese, Harsh and Svetina (131) assessed data visualization literacy on the novice-expert continuum with 202 participants, finding significant gaps between groups on opposite sides of the spectrum but little between those in the middle. Börner et al. (34) investigated the data visualization literacy of 273 museum visitors in four US science museums, asking subjects to, e.g., name visualization types and describe what kind of data they would be useful for. Their findings reveal the extent to which many users are not able to even name a visualization while highlighting the lack of a common terminology for all but the basic types.

Other researchers have focused on designing interventions to teach visualization literacy in formal and informal educational settings. Alper et al. (9) developed C’est la Vis, a tablet-deployed learning experience for elementary school children. Using pictograms and bar charts, concreteness fading skills are trained when transforming individual elements (pictograms) to more abstract bar charts via the 2D interface. Observing touch interactivity, verbal activity, and class dynamics, they find high levels of interaction with the app, an ample amount of communication between subjects (often for assisting each other), and short but not significant disruptions of the classroom through the app. Beheshti et al. (26), while not strictly focusing on data visualization, present Spark, a science museum experience about electronic circuits in three conditions: single-display (static), single-display (animated), and AR. Users built virtual circuits and were then shown animations of electron flow in the wires (except in the static condition). The authors then tested understanding and collaborative practices in 60 parent-child dyads, and find that children performed better in the non-static conditions and that parents in the AR condition were more likely to adopt the role of a co-learner rather than an educator.

2.1.2 Typology & process model

At the core of the DVL-FW, there are seven typologies to capture building blocks of data visualizations, see Table 3.

Table 3. DVL-FW typology as it appears in Börner, Bueckle and Ginda (32).

Table 1. Typology of the DVL-FW

Insight needs	Data scales	Analyses	Visualizations	Graphic symbols	Graphic variables	Interactions
Categorize/cluster	Nominal	Statistical	Table	Geometric symbols	Spatial	Zoom
Order, rank, sort	Ordinal	Temporal	Chart	Point	Position	Search and locate
Distributions (also outliers)	Interval	Geospatial	Graph	Line	Retinal	Filter
Comparisons	Ratio	Topical	Map	Area	Form	Details on demand
Trends (process and time)		Relational	Tree	Surface	Color	History
Geospatial			Network	Volume	Optics	Extract
Compositions (also of text)				Linguistic symbols	Motion	Link and brush
Correlations/relationships				Text		Projection
				Numerals		Distortion
				Punctuation marks		
				Pictorial symbols		
				Images		
				Icons		
				Statistical glyphs		

These types were previously defined in the Atlas of Knowledge (31), an extensive review of prior literature on data visualization, statistics, and best-practice examples from applied data visualizations spanning a variety of disciplines such as economics, library science, and geo science. Due to its nature as a book, the Atlas is limited to static, paper- or screen-based visualizations, thus demonstrating a lack of exemplars involving interactions.

Insight needs (Table 3, first column), also called *basic task types*, constitute the objectives that users of data visualizations like to achieve. Frameworks for categorizing insight needs have previously been proposed by authors in data visualization (40, 87) and geography (28). Due to the abstract semantics of task types, there is little agreement on how to properly categorize them, and in what granularity they should be described or grouped.

Data scales, on the other hand, are more clearly defined. This typology captures whether data is quantitative (ordinal, interval, ratio) or qualitative (categorical). Since data scales are more low-level and less dependent on semantics, they can be tied quite directly to different types of datasets.

In the DVL-FW, data scales are used to describe a dataset (or a dimension of a dataset) through its numerical (or non-numerical) features. Additionally, data scales limit or enable mathematical operations (such as transformation) on the data or dimension.

The Analyses typology captures the kinds of transformations to be performed on datasets before visualizations. Statistical analyses are often necessary to create anything but the most basic visualizations from an already heavily adapted or simplified dataset. Depending on the nature of the insight needs of a user, these analyses can be purely statistical, or they can be more complex, e.g., temporal trends, geospatial patterns, topical clusters, or relationships. Most datasets must be preprocessed in some way before being visualized.

The visualizations typology captures visualization types. The DVL-FW distinguishes between tables, charts, graphs, maps, trees, and network visualizations. Tables offer a simple yet powerful way of visualizing data via rows and columns, and there has been research on how to make them even more powerful through adding interaction (157). Graphs are another omnipresent type, with essential subtypes being scatter graphs, boxplots, and line graphs. Charts work similar to graphs but lack a clearly defined reference system; a good example is the pie chart. Maps are another type well-known to the general public due to their prevalence in the news media, school books, and the internet, with particularly high exposure during major historical events, such as presidential elections (147, 148) and public health emergencies (54, 76). Finally, network visualizations (and one of its subtypes, the tree) give visual form to relationships among items, using nodes and edges (links). Oftentimes, nodes that share features are arranged close together to reflect local similarity. The terminology used to describe the very basic building blocks for visualization types also stems from the work of cartography: reference system, or base map, and data overlay. Reference systems describe a space into which data is plotted according to mapping rules; the most prominent ones are the Cartesian (x-y-z) and the polar coordinate system (albeit much harder to read). Two-dimensional reference systems provide a horizontal and a vertical dimension (see Figure 3), and they are most commonly used for geographic maps in atlases in navigation systems.

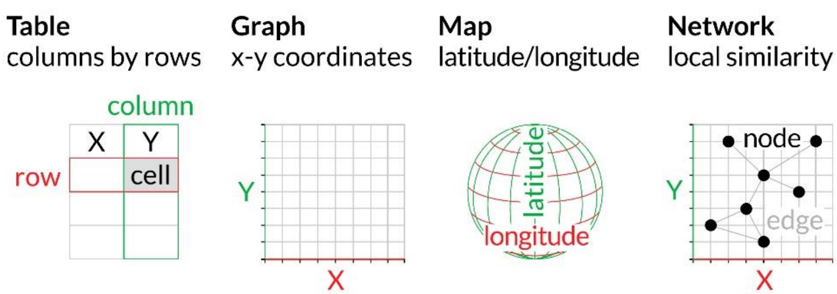

Figure 3. Reference systems across visualization types (32).

There is an extensive body of research on the cognitive processes involved to deconstruct visualizations for extracting insights, ranging from spatial memory (121, 194) to “graph schemas“ (153) and graph comprehension (171). What we know about how humans interpret visualizations in turn affects the way we construct visualizations to communicate insights.

The next two categories in the DVL-FW typology, graphic symbols and graphic variables, see Figure 4, refer to the visual encoding of data variables. Alternative names for graphic symbols and graphic variables are marks and channels (28, 146). Graphic symbols such as points, lines, areas, volumes, and text are paired with graphic variables such as color, position, and size in order to represent data visually as a data overlay within a reference system. The process of mapping data to points, lines, areas, volumes, etc. is called visual encoding. Not all visualizations support all graphic symbol-variable pairings. For example, the graphic variable of motion is not supported on a static visualization printed on paper. Similarly, stereoscopic depth is not supported on monoscopic media. There is a wealth of research on the strength of graphic variables with regards to human cognition. In a landmark user study, Cleveland and McGill (59) identified position and length as the graphic variables best-suited for human use (hence the prevalence of scatter graphs and bar charts), with rotation and area least-suited. Heer and Bostock (101) later ran similar studies using a crowdsourcing setup. Kim and Heer (119) performed user studies linking visual encodings to task types.

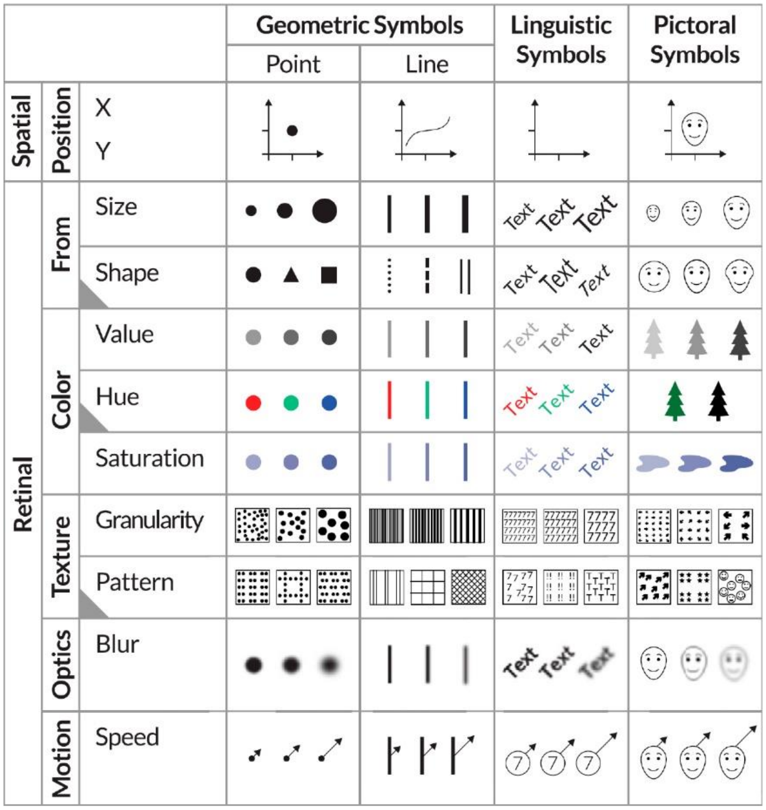

Figure 4. Selection of four graphic symbols and 11 variables from a total of 11 graphic symbols and 24 graphic variables from Börner, Bueckle and Ginda (32). Qualitative variables are marked with a grey triangle.

Lastly, the interactions typology, see Table 3, column 7, contains techniques to allow users to manipulate data views, change visual encodings, and otherwise change the state of a data visualization through pre-defined input parameters. The interaction typology in the DVL-FW consists of nine elements, which were in turn collected from a literature survey.

Categorizations of interactions in data visualizations have previously been proposed. Heer and Shneiderman (103), in a survey of relevant literature and prominent visualization tools, identified 12 interaction types, grouped into three categories: data & view specification, view manipulation, and process & provenance. The types in the data & view specification and view manipulation categories are roughly covered by the existing DVL-FW interaction typology, but process & provenance (record, annotate, share, guide) are not. Regardless of the goal of an interaction, for optimal usability (31), interactions should be "[r]apid, incremental, and reversible" (p. 68).

Interactions are the focus of this dissertation, and we consider the extension of this typology one of the main contributions to data visualization theory of this dissertation. While the DVL-FW interaction typology captures a variety of interaction techniques extracted from an in-depth view of the literature, several important types are missing, many of which are extensively used and studied in scientific visualization. In section 2.1.4, we are investigating potential extensions and improvements of the DVL-FW to better capture data visualizations outside of the Desktop metaphor. Another essential element of the DVL-FW is its process model (DVL-FW PM). It summarizes how data visualizations are constructed through repeated steps. Specifically, it encapsulates the cyclical nature of how data is transformed into visual insights while interlinking the seven typologies shown in Table 3 to steps in the workflow, see Figure 5.

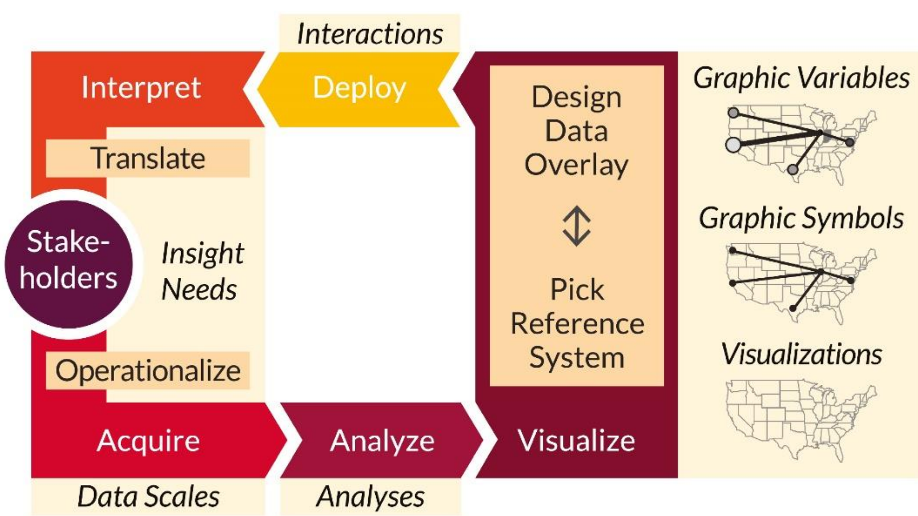

Figure 5. DVL-FW process model by Börner, Bueckle, & Ginda (32)

The PM begins and ends with the stakeholders and their identified insight needs. Oftentimes, broadly formulated questions about data need to be transformed into answerable questions before the process can begin. This process of translating real-world needs into a workflow process is called operationalization; it guides the subsequent step where data needs to be acquired that is likely to satisfy the stakeholders' insight needs. While terminology may vary, this initial process is captured in a variety of other workflow descriptions as well. Elsewhere in the literature (89), the operationalization of a question is described as a division into subtasks, or proxy tasks, to gain a more low-level understanding of what is required of the visualization, aided by proxy values, which are measures in the data set that contain information about a desired value. For example, if one asks: "Which superhero movie was the most popular in 2018?", box office numbers would serve as a proxy values for answering this question.

A crucial improvement of the most recent DVL-FW PM (32) over previous versions (31, 35) is the alignment of the seven typologies into the workflow. Insight needs (Table 3, column 1) need to be investigated when working with the stakeholders at the onset of the data visualization process. Data scales (column 2) and analyses (column 3) are applied when acquiring data upon input from the stakeholders, and then analyzing them, respectively. The visualization process proper consist of two steps: picking a reference system and designing a data overlay, echoing the role of reference system and data overlay as the two primary building blocks of a visualization. The typologies of visualization types (column 4), graphic symbols (column 5), and graphic variables (column 6) are then used to construct a set of visualizations based on prior steps in the workflow. Finally, in the deployment phase, interactions are implemented to match functionality with a target output medium. Not all interactions are possible for all deployment methods; for example, a printed visualization cannot support the interaction type of link and brush, where the user highlights a subset of data in one view and then sees it dynamically highlighted in another view as well. However, some interaction types that are primarily used in interactive media can be supported on static media but with a twist. Consider, for example, the ample use of maps in atlases where the user is presented with one large reference map and several detail views of particular regions of interest. Figure 6. The detail view in Figure 6B is similar to the result produced by interactive zooming, but does not require input via a mouse, as it would in a web interface, but shifting the user's focus or flipping a page. Sea charts for nautical navigation, for example, are published and updated regularly as part of the marine operations of every seafaring nation. These maps come at various scales to empower users with different kinds of insight needs. Some maps cover territories on a very large scale, allowing mariners to plan long-haul trajectories by providing important navigational information for entire regions; smaller-scale maps with more local information then enable seafarers to navigate coasts, waterways, ports, and traffic separation schemes in a local context. This practice predates interactive data visualization.

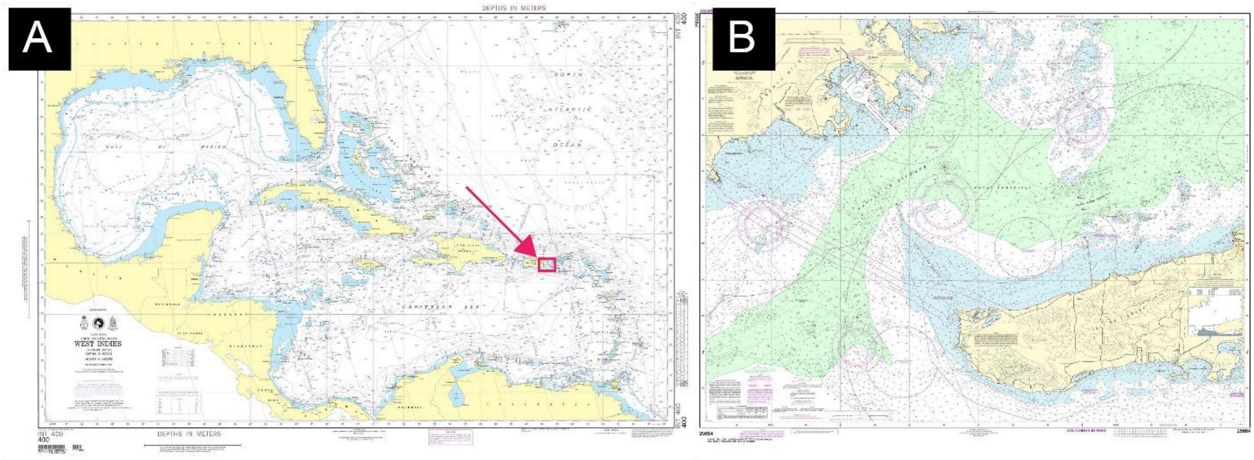

Figure 6. ***A****: NGA Nautical Chart 400, West Indies, showing Florida in the north, the northern tip of South America to the south, and Mexico to the west.* ***B****: Pasaje des Vieques and Radas Roosevelt, which is also shown in* ***A*** *but in much less detail (pink arrow).*

Interpretation and translation then are the last steps in the DVL-FW PM. During this phase, the visualizations created with particular insight needs in mind are brought back to the stakeholders. They are then able to evaluate how well their insight needs were satisfied, often with input from the visualization designer explaining the reasoning behind design decisions.

2.1.3 Validation & usage

The DVL-FW has been used in a variety of settings: research, exhibit design, and education. The theoretical advancement and consolidation of the DVL-FW has been developed in the scholarly literature. The Atlas of Knowledge (31) is a large-format book containing dozens of best-practice examples of (mostly static) visualizations, structured around the typologies of the DVL-FW in high detail alongside relevant domain application-specific terminology. The most recent version of the DVL-FW has been described and published in a major journal (32). Research involving the DVL-FW has been conducted on online education (82), informal science education (34), and the use of base maps for topical and geospatial visualization with user studies (36) and learning sciences research, specifically “Make-a-Vis,” an online learning tool for data visualization based on the DVL-FW where users can drag-and-drop data records to graphic symbols and graphic variables (33, 69).

Furthermore, the DVL-FW has informed the design, optimization, and evaluation of the Information Visualization Massive Open Online Course (IVMOOC), a free, web-based class for teaching the basics of data visualization to anyone with an internet connection (82). An accompanying textbook presents the structure of the course as it covers contents on temporal, geospatial, topical, and network visualization over several weeks (35). More recently, the content of the IVMOOC has been updated for the Visual Analytics Certificate (VAC), a shorter, 6-week online course focusing on professional education for professionals who seek to learn critical data visualization and analytics skills over a variety of industries, including for-profit, non-profit, and government (69). Two pieces of software have been developed around the DVL-FW as teaching tools for these classes: Make-a-Vis and Sci2 (67, 169).

The *Places & Spaces: Mapping Science* exhibit is an annually updated collection of data visualization work from researchers and artists, aiming at presenting visualizations to audiences in informal education environments (68). Currently in its 16th iteration, Places & Spaces contains printed posters implementing a variety of visualization types, including but not limited to maps, bar charts, networks, and combinations thereof. Beginning with the 11th iteration, in 2015, macroscopes have been added to the collection. Macroscopes are interactive, web-deployed visualizations, allowing users to dynamically interact with datasets on migrant flow (2), the 3D layout of galaxies in the universe (7), and the smell of physical locations in cities across the world (155), among many other application domains. Deployed on large touch displays, visualizations in the Places & Spaces exhibit are shared with users all around the world every year (66). Currently, *Places & Spaces* includes 20 of these macroscopes.

2.1.4 Expanding interaction elements

In this dissertation, we address a shortcoming of the DVL-FW: It is mostly focused on visualizations with two-dimensional reference systems. While 3D position is a supported graphic variable in the graphic variable typology and the *Atlas of Knowledge* (31) mentions several interactive visualizations, the best-practice visualizations reviewed in the literature published on the DVL-FW so far are almost exclusively static and 2D. Presumably, there are two reasons for this. First, books and papers naturally favor static visualizations that can be represented well on the two-dimensional, high-resolution medium of printed books. Second, the inclusion of a third spatial dimension has always been met with caution in the field of information visualization. All too often, the third dimension is added to natural two-dimensional reference systems without adding any value beyond aesthetics, the advantages of which are highly debatable. Few (87), in his chapter on general graph design, examines a range of 3D bar charts and line graphs (to be found, for example, in Microsoft Excel), and bluntly advises readers to “[a]void 3-D displays of quantitative data” (p. 203). Similarly, Munzner (146) warns of the “disparity of depth” (p. 118) where the human ability to judge length, shown to be one of the most powerful graphic variables on the x-y plane (59), works nowhere near as well on the z-plane, i.e., depth. Quoting Ware (213), she emphasizes that even though we live in a three-dimensional world, our perception of width and height is far more advanced than our understanding of depth, and thus insinuates that we experience life in a "2.05D" world (p. 119). Occlusion, the phenomenon where objects that are farther away seemingly disappear behind closer objects, is a standard experience for humans with functioning eyesight. In data visualization, it becomes a problem as data items might simply not be immediately visible, necessitating the implementation of **navigation** to view data from different angles. An additional depth cue (and challenge for data visualization) is the phenomenon of perspective distortion or foreshortening. Objects that are farther away seem smaller. Because size (i.e., length) is such an essential graphic variable, this means that if two graphic symbols of the same size are rendered in a visualization where one of them is farther away from the user's viewpoint than the other one, they will (falsely) appear to be of different sizes, thus producing inaccurate results for the viewer. To counteract the confusing effect of unwanted perspective distortion, there exists a variety of tools. A popular one is **orthographic projection**, a method where objects maintain their size no matter the distance from the viewer (see Figure 7).

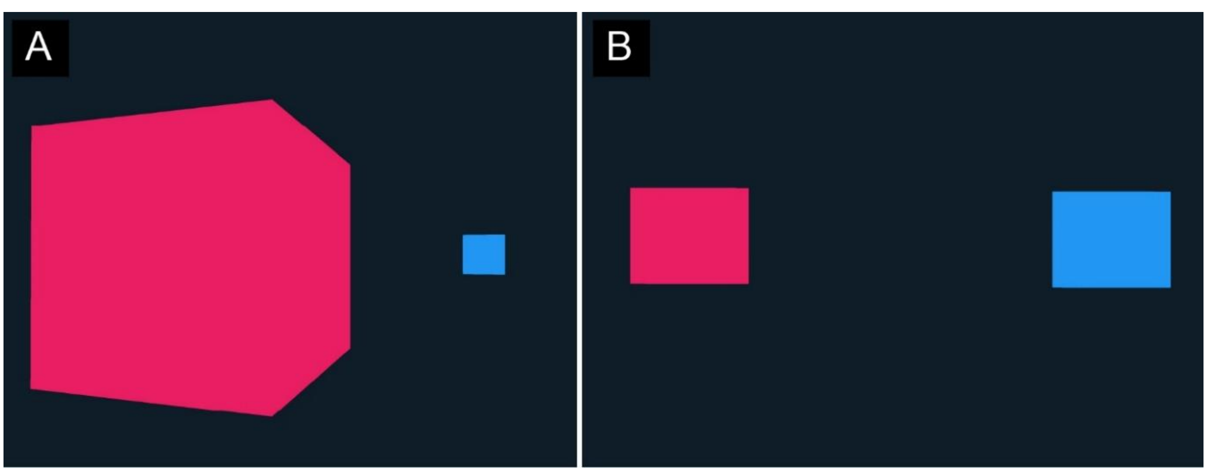

*Figure 7. **A**: Unity (200) scene with two same-size cubes where the blue one is ~10 scene units, i.e., meters in real-world scale, behind the pink one. Distance between pink cube and camera: ~1.7 scene units. **B**: same cubes, same distances but with orthographic projection. Notice how the size is preserved in **B** while in **A**, the pink one seems much larger than the blue one due to foreshortening.*

Of course, applying **navigation** and **orthographic projection** to counteract the negative perceptional effects of these depth cues can be detrimental to the visualization. After all, many 3D visualizations embody the spatial structure of data, thus making use of our immense visual abilities. Flattening a data visualization through orthographic projection, while combating the issue of perspective distortion, also takes away a crucial piece of the visual encoding. Further, advanced interaction techniques such as 3D navigation cost two valuable resources: time and cognitive effort, as Munzner (146) remarks in her chapter on "No unjustified 3D" (p. 117). After all, interaction types often involve complex input and can have a steep learning curve.

In its current version, the DVL-FW offers limited language that allows us to formulate questions surrounding **navigate** and **animate/replay**. Neither of these interaction techniques are mentioned in the current DVL-FW interaction typology and are additions proposed in this dissertation. Thus, the interaction typology of the DVL-FW is limited and will be expanded for non-2D visualizations. This is in contrast to, e.g., the graphic variable typology, where clearly non-2D and/or non-static examples such as stereoscopic depth, velocity, and speed are listed as retinal variables. To understand the full range of interaction techniques in data visualization, a literature review in section 2.2 provides the basis for an integrated overview of interaction taxonomies in Table 4. In three user studies (chapters 5, 6, and 7), we will then investigate the use of two new interaction techniques new to the DVL-FW (**navigate**, **animate/replay**) and two old ones (**filter**, **link and brush**). Further, to avoid pitfalls in

VR research design, we perform a literature review of VR user study methodology in section 2.4. In expanding the DVL-FW interaction typology, we further aim to introduce a **DVL-FW interaction model** according to Beaudouin-Lafon (24), specifically a "framework for guiding designers, developers, and even users […] to create interactive systems" (p. 16), and that can be "evaluated along three dimensions: **descriptive** […], **evaluative** [, and] **generative** power" (p. 17, highlights added by this author). Considering interaction models from these three angles provides us with the ability to formulate research questions that we can then answer with a variety of methods, including user studies. The following questions highlight the goal of each dimension.

Descriptive: How well can our interaction model inform the operationalization process by describing a range of existing interfaces? This dimension is essential for the operationalization of insight needs, as an early exploration of interface options can guide the early stages of a visualization cycle. Best-practice examples can then be incorporated to make informed decisions about what data should be supported for visualization. Additionally, a terminology can be compiled to discuss existing and planned work with a common vocabulary. We add to the descriptive component through an in-depth study of VR interaction exemplars in section 2.5.

Evaluative: How well can our model help us evaluate alternative designs when translating our work back to our stakeholders at the end of a visualization cycle? This component needs to be contextual; specifically, it is paramount to only compare and evaluate against each other interfaces that share a set of common features. We build this component by comparing two essential deployment methods in user study 1 (see chapter 5): Desktop and VR, two different media that are, however, comparable on the basis of user input, task completion time, accuracy, and user satisfaction. Further, the user studies in chapters 6 and 7 contribute to our evaluative component through testing the power of an intervention for improving performance for a control and an experiment cohort of users.

Generative: How well can our model provide guidelines for the visualization process, specifically the various interface design choices that need to be made? The ability of a model to be used by designer and other practitioners when implementing visualizations is of paramount importance. There is little prior research on best-practice rules for implementing data visualizations in VR, and a generative dimension of a VR interaction model could greatly help fill this void. We build this component by translating insights from our user studies into VR design recommendations in chapter 8. The generative component thus follows from a combination of the descriptive and evaluative components.

2.2 Interaction types in data visualization frameworks

The literature review for this dissertation consists of three parts: a brief overview of general frameworks and interaction typologies to outline previous attempts of formalizing interactions for data visualizations; VR user study methodologies; and the challenges of a third spatial dimension, i.e., 3D, in data visualizations in general and for VR in particular. We chose these three rather distinct areas of study in order to address three fundamental challenges when using VR for the visual representation of information:

First, this dissertation focuses on the use of VR for the visual display of data specifically (rather than, e.g., the use of VR in entertainment). We thus aim to investigate how VR can be understood as a medium in the growing arsenal of traditional data visualization, and begin with a brief overview of data visualization frameworks in this chapter, specifically synthesizing previous approaches to categorizing interaction types. We are particularly interested in exploring and summarizing the rationale and the guidelines by which researchers have previously tried to group interactions with data. After investigating existing typologies for various interactions (such as the quasi-standard zoom, filter, and link and brush) in section 2.1.1, we present an integrated overview of interaction taxonomies in tabular form in section 2.2.2 (Table 4). Additionally, previous work for interactions in VR is identified and synthesized in section 2.3.

Second, the rich data coming out of VR user studies and the relative novelty of VR as a subject of user studies pose a methodological challenge that needs to be addressed by building on prior research, see section 2.3.3. Here, this dissertation investigates methodological questions and answers as identified by other researchers with regards to VR user studies. This serves as a preparation for the experimental part of this dissertation, identifying best-practices and pitfalls when collecting data from human subjects.

Third, we postulate the inherent spatial three-dimensionality of VR negates much of the visual simplicity that constitutes one of the key strengths of traditional data visualization by not only providing a multidimensional display, i.e., visual depth, but also requiring multidimensional user input in the form of VR controllers, a device capable of higher-dimensional input than a mouse and a keyboard. This is reflected in a variety of previous work.

2.2.1 Overview

In this section, we investigate the following:

- There is an agreement that interactions are not properly investigated in information visualization literature.
- Existing interaction frameworks categorize interactions in various ways, making cross-framework comparisons difficult.
- There is little guidance on how to develop interactions with data in VR. What exists is situational and often dependent on domain application.

In order to formalize how data visualizations are constructed, interpreted, and taught, a variety of frameworks have been proposed over the past decades. While providing a comprehensive review of the full history of formalizing data visualizations is out of scope for this dissertation, a select few seminal writings of importance for the field are referenced here. Jacques Bertin's *Semiology of Graphics: Diagrams, Networks, Maps* (28) is generally cited as having coined the terms marks and channels to describe the two elements of visual encodings used in data visualization: graphic icons (such as points, lines, and other geometric primitives) alongside properties such as color, position, and size. While other terms are sometimes used, such as graphic symbols and graphic variables (31), marks and channels has proven to be a concept that endured through the past decades of research as evidenced by its continued use in more recently published work (103, 146). Tufte, Goeler and Benson (196) and Tufte

(195) approach data visualization from a design perspective with little to no interest in underlying data structures.

There is also a variety of frameworks that have been implemented in software; specifically, Wilkinson's *The Grammar of Graphics* (215) forms the theoretical basis of ggplot2, a data visualization package for the statistical programming language R (https://ggplot2.tidyverse.org/) as laid out by Wickham (214). Similarly, Polaris (later commercialized as Tableau, https://www.tableau.com/) was built as an extension of the Pivot Table interface commonly known from spreadsheet software with the goal of enabling users to visually explore large quantities of data (191). Similarly, D3.js (37) and its high-level abstractions Vega (167) and Vega-Lite (166) do not constitute data visualization frameworks per se; rather, their creators build their software on top of an already existing and widely used technology (in this case, the Document Object Model of modern websites) and thus leverage their users' preexisting knowledge of web technology to create interactive visualizations in a web browser.

Between the writings of early data visualization pioneers and publications involving more modern hardware and software (such as the web browser), there is a noticeable rift in the treatment of interactions. In her textbook *Visualization Analysis and Design* (146), Tamara Munzner argues:

"[The] hallmarks of the last 20 years of computer-based [visualization] are interactivity rather than simply static presentation and the use of [visualization] for exploration of the unknown […]."

This is certainly due to the fact that many of the seminal works on data visualizations were written for graphs, charts, and maps on paper. Munzner highlights in this brief quote not only the shift towards more user input to adjust what data visualizations show on the spot, but also the role this interactivity has played for the exploration of ever larger datasets. This sentiment is echoed by Börner (31), citing the volume and complexity of modern datasets as a rationale for interactive data visualizations:

"Many data sets are too large to fit on one screen or printout. Interaction permits the user to first gain a global overview of all the data and then to zoom in to that data [...]. The structure and dynamics of data can be explored at multiple orders of magnitude. In principle, any part of the analysis workflow and any layer of the visualization design can be modified via user input." (p. 68)

And Unwin (201) comments on the realm of possibilities opened by including interaction in data visualizations:

"Exploratory graphics have in effect unlimited space [...] and are primarily only for the person who draws them. Presentation graphics are for conveying information, while exploratory graphics are for discovering information." (p. 10)

Dix and Ellis (73) went so far as to "establish that the heart of modern visualization techniques is interaction and propose that interaction can be applied to any representation however simple." (p. 124)

Those quotes make apparent how interactions with data, and thus, the user as an active participant in the data visualization process, became important for the design of a data visualization but also that interactions are essential for the exploratory analysis of large datasets. Of course, interactions in data visualizations are not limited to the narrow scope of visual analysis and statistics, although one could argue the roots of interactive visualization lie there as outlined by Friedman and Stuetzle (94) in their 2002 paper on mathematician John Tukey's work. But where do interactions fit within the workflow underlying data visualization, and how can we conceptualize what constitutes an interaction, exactly? In contrast to tasks in data visualizations where there is ample research (11, 12, 40), interactions are less researched in the field (81).

Interactions are tightly coupled to tasks as well as the underlying data scale types. After all, a user who interacts with a visualization is likely going to want to complete a task, itself a rather broad term, ranging from simply browsing for new pathways to potential insights to answering a very specific question about properties of subsets of the data, e.g., the change of a variable over time. Despite all the interdependencies between tasks and the underlying data with interactions, and the resulting complex network of relationships, basic buildings blocks of the data visualization workflow are consistent among a variety of frameworks. While the works reviewed in this chapter address a diverse set of issues, three clear themes emerge:

- Interactions can be related to tasks via a taxonomy
- Interacting with visualizations can include interacting with the underlying data in various levels of abstraction.
- Interactions need to be defined in such a way that they are independent of their actual implementation within a medium so that an interaction theory for data visualization can apply to non-traditional media, such as VR.

The goal of this section of the literature review is to construct an integrated overview and alignment of interaction technique taxonomies in the form of a table, see section 2.2.2. The method for identifying literature consisted of investigating two papers that themselves contain expansive literature reviews (40, 218). This part of the literature review thus contains references to papers from a variety of disciplines with the overarching theme of user interaction in data visualization. We will also briefly explore prior work on input devices as this will provide essential background information for our later review of data visualization interaction in VR and VR user study methodology.

Brehmer and Munzner (40) propose a multi-level typology of “domain- and interface-agnostic” visualization tasks, reviewing around 30 extant classifications from a variety of disciplines such as information visualization and human-computer interaction (HCI). The authors propose to describe tasks with three distinct but interconnected components: why, how, and what, where the why relates to the user’s motivation, what defines input and, if applicable, output of the task, and how captures the concrete action being taken to perform the task. While their focus is on creating a taxonomy for tasks in data visualizations (and not on interactions per se), they include 11 interaction types as part of the “how?” of their classification, see also Table 4. While their typology for visualization tasks is not the only one (11, 12), theirs is based on the most through literature review per our knowledge and, as previously mentioned, served as a source for literature for this review. In distinguishing between why, how, and what, Brehmer and Munzner (40) build heavily on Roth (163), who proposes a taxonomy of cartographic interaction primitives. In that framework, there is distinction between goals, objectives, operators (actions), and operands (receivers of action) to describe any user interaction with granularity. While these four terms roughly correspond to the why, how, and what of Munzner & Brehmer’s multi-level typology (40), there are a few essential distinctions: First, Roth’s taxonomy terms are relegated to cartographic interactions. For example, all his operands are of a spatial dimension. While Roth names 30 interaction primitives, only 17 of them contain references to actual interactions, with the rest describing goals and receivers of action. Second, the author does not name input and outputs of these primitives. To validate his framework, the author performed a card sorting experiment with 15 subjects, all of them experts in cartographic interface design and development, with cards having objectives and operators, i.e., interactions, written on them. These had been identified in expert interviews from a previous experiment as well as by reviewing the literature. The author found that there was a much higher variation, i.e., less agreement among experts, for the objective cards than for the operator cards. Unsurprisingly, given that all the subjects were cartographic interaction designers, the most agreed-upon operator was zoom, tightly followed by symbolize and retrieve. While we might expect different results if experts outside of cartography had been interviewed and selected as participants, the empirical method underlying this framework makes it a valid contribution to the investigation of user interactions in data visualization. Roth's proposed interaction types are part of our interaction type overview in Table 4.

Similarly, in prior work, researchers aimed to build "a cognitive model for user-network interaction" by presenting a questionnaire to computer network experts, asking them to rate the importance of each of the proposed tasks (145). This study supported the creation of a mental model for interface usage, and to achieve consistent interface design. The authors considered a low-level list of interaction techniques (such as transform, zoom, and locate), which later informed the what aspect Brehmer & Munzner for their abstract task-based taxonomy (40).

Another example of a taxonomy based on observation contains the data analysis methods of scientists (188). The goal of this study was to develop an overview of elements used in data analysis independent of discipline or tools used. In addition to an unspecified number of pre-interviews, authors performed eight observation sessions with five scientists from three institutions working with their own data. The researchers watched the subjects perform their normal routine while occasionally asking for context when observing specific actions that they felt needed further explanation for someone outside of the subject's domain. Based on this study, the authors propose a framework for scientific data analysis, broken down into a hierarchical structure with two main phases: investigation (interacting with representations, applying math, and maneuvering) and integration of insight (maneuvering and expressing ideas).

Two of their findings seem particularly relevant for the investigation of data visualization interactions: First, the category of maneuvering to denote "movement within and among programs" is part of both the investigation and integration of insight phases. As such, maneuvering can be considered an interaction type that can either be data-related or not, e.g., resizing a window is not connected to the underlying data being analyzed; subsequently, the authors name this part "navigation". They name the second part "data management", which is closer to our current understanding of traversing data, e.g., zooming in and out of maps or network visualization, providing overviews and then details on demand. Second, the presence of an integration of insights, with the categories maneuvering and expressing ideas, is an essential step where the subject describes their conclusions from the work previously performed during the other phases. This is especially of interest since this phase does not necessarily involve any interaction with the data at all but merely with non-data related parts of the application.

This basic distinction between interactions with data and interactions with the application is also mirrored in the taxonomy of interactive dynamics for visual analysis (103). The authors review a wide range of relevant literature as well as commercial and open-source data visualization tools and propose a taxonomy with 12 interaction types across three categories: data & view specification, view manipulation, and process & provenance. While it is not possible to align these two frameworks perfectly with each other, it is viable to view process & provenance (with the interaction types record, annotate, share, and guide) as an elaboration on the expressing ideas phase (188). Similarly, elements from the previously mentioned maneuvering category correspond to view manipulation (with the types select, navigate, coordinate, and organize). Neither view manipulation nor process & provenance, while containing interaction types, necessarily mean interactions with data. The authors do not elaborate on relationships and interdependencies between these 12 interaction types (see also Table 4) but acknowledge that their list can serve to emphasize areas in need of further research.

Ben Shneiderman's much-cited mantra "overview first, zoom and filter, then details-on-demand" (176) is often invoked as a basic rule when implementing data visualizations in general, and interactions with data visualizations in particular. The author identifies seven interaction techniques (overview, zoom, filter, details-on-demand, relate, history, extract), see also Table 4. While we can see a significant overlap with work investigated here previously, Shneiderman does not make any distinctions between interaction types and tasks; instead, he refers to these techniques as tasks to be performed by the user rather than concrete techniques to be employed.

Also, along the line of linking interaction to user tasks and goals, Yi et al. (218) performed a literature review in 2007. While emphasizing the relative lack of research on interactions in data visualizations, they propose seven categories of interactions that "are organized around a user's intent", thus stepping away from focusing on low-level interaction techniques (218), such as previously discussed here and outlined elsewhere (145). The authors review a set of sources with interaction taxonomies, some of which are also addressed in this literature review (11, 44, 56, 73, 117, 176, 197, 212, 215, 220), then follow up with a survey of commercial tools before finally surveying papers in sub-disciplines of information visualization. Similar to what would be presented later (40), the authors found that different interactions can help a user achieve the same goals, and identified user intent as an emerging way of grouping interactions. Subsequently, they presented seven interaction types as shown in Table 4. While their proposed interaction types can be aligned with others, some clearly show a different granularity; for example, arguably, "abstract/elaborate" is not necessarily the (widely agreed upon) "zoom", but given the authors' definition, zoom can clearly be seen as a kind of interaction to abstract (=zoom out) or elaborate (zoom in). A similar focus on user intent (although with a less flexible pairing of intent and interaction) can be found in Buja, Cook and Swayne (44) who distinguish between the rather abstract-sounding finding Gestalt, posing queries, and making comparisons. Their limited number of high-level interaction techniques comprises focusing, linking, and arranging views, with a selection of operands for those actions to be performed on. These types present an extension of their previous work on focusing and linking as abstractions of the "diverse methodology" found in statistical graphics across multiple disciplines (such as, for instance, geographic information systems and time-series analysis) at the time (45). Their assertion of rendering and manipulation as two root areas data visualization would be adopted as representation and interaction in subsequent literature (218). The authors also highlight the added benefit of real-time interaction with instant visual feedback, effectively adding time as a dimension to the data visualization. An interesting aspect of this paper is the focus on 3D visualizations and environments, likely due to their discussion of the XGobi tool and its 3D interaction capabilities in the same paper. While this paper went into greater detail on three specific interactions relevant for their "zoology" (p. 78) of interaction techniques, the authors largely ignore non-multivariate visualizations (thus the noted absence of histograms or line-graph visualizations). In contrast to this rather high-level approach to classifying interactions, Chuah and Roth (56) propose a framework where basic visualization interactions (BVIs) are combined to build potentially complex user inputs. A BVI consists of inputs (attribute, control object, value, formula, focus set), outputs (Graphical state data state, control state), and an operation where the input are set via the UI or by default (through the designer), although this distinction seems poorly worded since the UI, by necessity, was designed by someone before being accessed by the user for the sake of interacting with the visualization. The aforementioned operation consists of condition check and action, allowing for the chaining of multiple BVIs to achieve complex interactions. This idea of interdependencies between interactions is not unlike the how and what modules discussed previously in this review (40), although nothing like the why module is specified. While the interaction types discussed here cannot be fitted into our synthesized typology in Table 4, the authors' focus on the semantic design of interaction (as opposed to the lexical and syntactic parts) first identified earlier in the literature (91) provides a valuable input for this work.

While the semantic design space applies to most of the literature reviewed here so far, the lexical and syntactic space is worth revisiting as well. Chuah and Roth (56) refer to lexical design (methods by which "input and output primitives are derived from basic hardware functions") and syntactic design ("a set of rules by which primitive input or output units can be composed or joined to form ordered sequences of inputs and outputs") to delineate spaces that operate on a hardware-oriented, more physical lower level, and build heavily on prior research on the semantics of input devices by Mackinlay, Card and Robertson (130), see Figure 8, and the classification of computer graphics subtasks (92). In light of the large variety of possible input devices that allow human-computer interaction, Mackinlay, Card and Robertson (130) identify the challenge to unify the "hodgepodge" (p. 147) of such devices into a coherent framework, onto which they then built later on (53). Some of the tools they list in their 1990 paper certainly sound antiquated (headmice, Polhemus cubes) while others are still being pursed (gloves, body suits) for VR after their development was dormant for a few decades. The authors introduce a formalized language with notation reminiscent of object-oriented programming to describe a variety of n-dimensional input devices and mappings from input to output. Employing methodology from one of the authors' prior work on creating application-independent visual encoding for data visualizations (129), the authors propose a framework to analyze expressiveness of user input, i.e., the correct translation of simple or composite user interactions with an input device into output towards an application. They define an input device with a 6-tuple of properties (similar to an object in object-oriented programming): manipulation operator (physical property vector), input domain set (range of values), current state of the device, resolution function to map input set to output set, output domain set, and work function W that allows for resetting the input device if needed, e.g., a spring in a joystick that sets the joystick back to zero if unused. This general terminology allows for various combinations to describe very simple, e.g., the buttons on a radio, to quite complex, e.g., a VR controller with accelerometer, several buttons, and a touchpad. They further differentiate between three composition types: connection composition (where output from one device cascades into another device), layout composition (the device coordinates in physical space), merge composition (where the output from two input devices creates a higher-dimensional set).

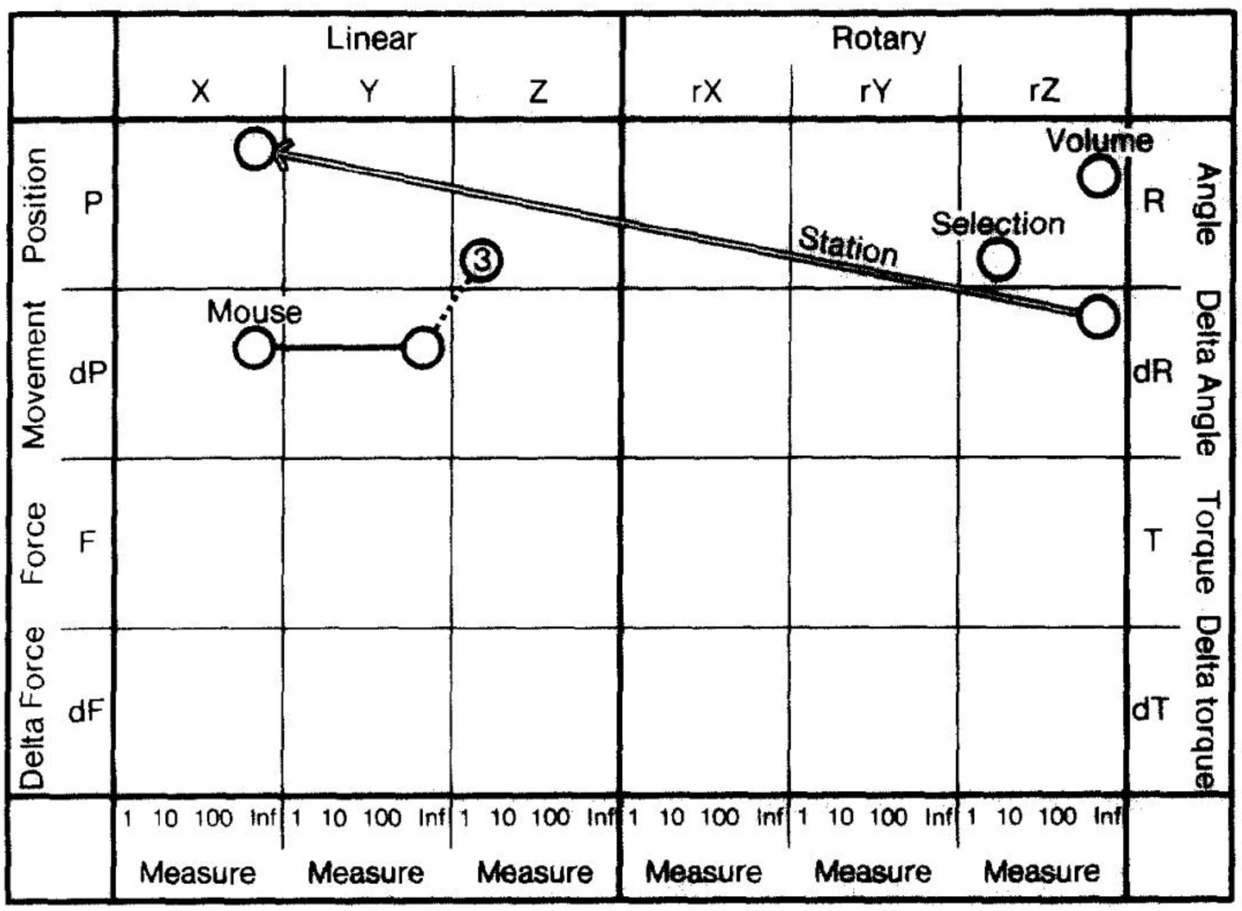

Figure 8. Graphical representation of the design space of input devices (130). Lines with arrowheads indicate cascading input-out mappings. Also note the range of measurements at the bottom, from 1 to infinity.

The authors then apply their framework to design a first-person walking experience in 3D, specifying user input via their formal language presented above. With this having been written in the days prior to 3D gaming, they call it "3D egocentric motion" (p. 173). Their stated goal is to facilitate consistency in interface design, empower users to adjust the interface if needed, and to contribute to the human-centered research in computer science and engineering. In other places, the authors present their framework of formalizing input devices and compositions thereof in a more concise manner and with an added focus on effectiveness, footprint, and bandwidth as other important considerations or "testing points" (p. 5) for successful human input to an application (52), ideas that were first expressed in Mackinlay's earlier work (129). Backtracking further in time, the idea that an input device as "a transducer from the physical properties of the world into logical values of an application" (p. 3) was proposed by Buxton (49) earlier, although as part of a taxonomy for continuous input devices only.

The idea that interaction types are built from utterances between fundamentally different agents (human and machine) is also explored in Robertson, Card and Mackinlay (160) where the authors apply their prior research on data visualization. They define two problems: First, the Multiple Agent Problem arises from the interplay of three elements: the user, the discourse machine, and the application. These need to collaborate in an asynchronous manner with a dialogue consisting of statements from human to the machine (via manipulations of input devices) and vice versa (via graphical, aural, or other displays), mounting a challenge for interface design. Second, the need for fluent transitions and rendering constitute the Animation Problem, where the user needs to be able to follow various objects of interest on screen to understand state changes, which poses another challenge for interface design (the authors suggest a minimum framerate of 10 frames per second (fps), but give 30 fps as a desirable target). As a solution, they present the Cognitive Coprocessor, an abstract architecture where a central animation loop receives information from various other elements such as the Task Queue and the Display Queue to ultimately create advanced visualizations. In these visualizations, the user can enact change through interactive objects such as virtual doors, static and editable text, and even buttons or joysticks (virtual or physical), which ties back to research discussed earlier in this review (52, 130). Interesting and relevant for our discussion of interactions in VR is their dedication to exploring user interaction in simple 3D environments. Nowadays ubiquitous terms such as gaze and body transformations are used throughout these papers. Even though these papers were written long before 3D interactions became a widely used and accessible feature, they contain descriptions of problems that do not cease to be of importance, with specific examples being framerate and delta-time correction, i.e., ensuring framerate-independent execution of the rendering pipeline. While the room metaphor implemented in the Information Visualizer did not carry on well commercially (177), this line of research on organizing virtual workspaces in three dimensions was carried out into the 1990s (104, 161) and continues to be of relevance.

Squarely back in the area of exploring how interactivity can add value to static visualizations, Dix and Ellis (73) investigate a selection of tools available at the time, e.g., Ahlberg and Shneiderman (6), and propose a loosely organized framework of interaction techniques that add value to established static visualizations (their examples included stacked bar graphs, line graphs, and pie charts). Their argument is that almost any visualization can be made interactive, in turn allowing visualization designers to "manage […] trade-offs [i.e., having to make a final decision on what variables to show in what visual encoding as one usually would when designing a static visualization] dynamically" (p. 131). In later work, the authors built on their investigation of interaction techniques by presenting methods to improve computational performance while zooming through statistical sampling (72). The interaction types they suggest can also be found in Table 4.

The idea of interactive zooming (and varying level of detail based on zoom level) is also explored by Keim (117), along with other interaction techniques captured in Table 4.

The idea that different types of visualizations enable users to solve different task, and that therefor interactivity can help designers make visualizations that support different tasks on user demand, can also be found in Tweedie (197). Or, as the author says:

"Whilst this limitation [i.e., how visualizations are tied to supporting specific tasks] applies to static presentations, it is not relevant when interactivity allows different features of the data to be made salient [….]. [The] underlying representation becomes a medium through which different features of the data are made explicit. A single representation can now be used to answer many different questions." (p. 375)

The author examines fourteen visualization and types, isolating five interaction types as shown in Table 4. Furthermore, the author distinguishes between direct and indirect manipulation of the visualization and constructs a continuum from manual to automatic interaction whereby interaction becomes increasingly externalized. For example, clicking on a point in a scatter graph closely resembles the physical action of putting a finger on an object. In contrast, clicking a button to run a selection algorithm is more external and similar to a black box. Concerned with the more direct end of this interaction spectrum, Chuah et al. (57) present work on selective dynamic manipulation (SDM) for a 3D visualization of food supply and demand in a crisis area. They identify five limitations to static data visualizations: maintaining scene context while focusing on individual objects, avoid or mitigate clutter and occlusion, maintain scaling across the scene without “dwarfing” (p. 61) objects representing small values, classify and save annotations to subsets of data, and mitigate height estimation difficulty between objects at varying distance from the user’s point of view (POV). While the domain application presented in this paper is a 3D visualization, all of these limitations apply to visualizations without spatial depth component as well. The authors address these issues by implementing a range of interaction techniques that allow the user to rearrange, rescale, and paint objects to facilitate height and location comparisons. The interaction techniques presented here are also included in Table 4.

Similarly to how SDM allows the user to directly manipulate and select objects, Wilkinson (215) also distinguishes between direct and indirect manipulation tools. He considers direct manipulation tools “operate on the graphic itself” (p. 552), while indirect manipulation tools are outsourced into UI elements. While his assertion that indirect manipulation can offer advantages by using “a game-style GUI for users under the age of 16” (ibid.) may not be an appropriate observation given the vastly successful visualization packages making extensive use of UIs, distinguishing between interaction tools for various levels of expertise certainly rings true. In chapter 17 of his book, the author discusses three modes of control in data visualizations: learning and playing vs. exploring. Learning and playing are for the process of visual data analysis carried out by analysts; exploring comes into play when end users gain insights from visualization software. While he makes no clear distinction between expert users and casual end users for these three modes, it becomes evident that interactions exist on a spectrum from low-level (data manipulation) to high-level (visualization manipulation).

Of further notice for the discussion of direct and indirect manipulation is the framework presented by Ward and Yang (212). Aiming to present an overview of spaces in which interactions happen in data visualizations, the authors identify: operators, i.e., actions, specifically navigation, selection, and distortion; interaction spaces, i.e., screen-space, data value-spaces, data structure-space, attribute-space, object-space, and visualization structure-space; operands, i.e., the portion of the interaction space upon which interaction is imposed; and interaction parameters, i.e., properties of the interaction operator. They name five types: focus, extents, transformation, magnitude, and blender. The authors explicitly mention direct manipulation for focus selection, but do not include a formal definition for (in)direct manipulation or differ between the two in their framework. In later work, Ward, Grinstein and Keim (211) add on to this framework; specifically, they also include the interaction operators of reconfiguring, encoding, connection, and abstraction/elaboration. While these operators function as umbrella terms for more specific techniques, we still included them in Table 4 for alignment with techniques proposed by others.

Another source of frameworks for interaction techniques can be found in papers that do not present a taxonomy per se but rather a list of recommendations for implementing tools for exploratory data analysis or visual analysis. Unwin (201) reviews interaction functionality from software such as MANET, Data Desk, JMP, and Visual Insights, none of which are available in their original form anymore (as of the time of this writing, the web links given in the paper are either defunct or relay a visitor to other products). However, the interaction types the author extracts are still valid and can be aligned within our integrated overview of interactions in Table 4 just fine (they identify querying, zooming, variation of displays, multiple views, grouping, rescaling, and linking.). More recently, Heer and Shneiderman (103) set out to compose a taxonomy of interactive dynamics for visual analysis in order "to assist designers, researchers, professional analysts, procurement officers, educators, and students in evaluating and creating visual analysis tools" (p. 1). Their extended definition of the term of "analyst" hints at the wide variety of potential users of interactive data visualizations. Surveying a bigger set of software (the industry giant Tableau among them), they identify 12 distinct interaction techniques, grouped into three categories (data & view specification, view manipulation, and the oddly named but essential process & provenance). The latter category includes interaction types covered only by a small subset of the other works in this literature review, which is perhaps due to the increased focus on collaborative and communicative as well as exploratory data visualization. This trifecta of interactions is echoed in Börner (31) with data transformations, visualization transformations, and visual view transformations (although process & provenance are not covered in that particular framework). Both Börner (31) and Heer and Shneiderman (103) frequently reference existing commercial and non-commercial visualization tools when abstracting interaction types, a technique not without a large amount of precedent in the literature.

For example, Friedman and Stuetzle (94), in their review of mathematician John Tukey's work on interactive graphics, refer to yet an older piece of software: PRIM-9 from the 1970s.

Because papers of this type make frequent references to existing software (and the technical challenges of the time), they date themselves more strongly. Interestingly, some technical challenges of the past, e.g., the comparatively low screen resolution of monitors in the 1990s, have long been resolved, but the base issue persists in other display devices, such as in VR.

A much more recent implementation of interaction techniques, this time for large display walls, can be found in Agarwal, Srinivasan and Stasko (5). In their VisWall software, running on an 84" touch display, users of a spectrum of expertise can create and modify visualizations based on data attributes. While the interactions they define, such as add and merge, are specific to their tool, the navigate alternative visualizations interaction is more closely aligned with more generalized interaction techniques presented in Table 4. The authors attempt to implement (and specifically reference) the guidelines on fluid interaction for data visualization as outlined by Elmqvist et al. (81) who base their work on a review of best-practice implementation of interactions. Referencing literature from information visualization and human-computer interaction (HCI), they investigate the concept of fluid interaction, defined as an interface that promotes flow, supports direct manipulation, and minimizes the gulfs of action, and review six best-practice visualizations. Further, the authors propose design goals (e.g., provide immediate feedback on user input, ensure user has correct mental model of application). These are echoed elsewhere in the literature, explicitly so by Börner (31) who proposes three forms of user guidance in interactive data visualization: manipulation support, coordination support, and self-evaluation support (again ensuring the user's correct mental model of the application).

They finish by calling for visualization criticism as a "skilled practice" (p. 339) based on agreed-upon "visualization design patterns" (ibid.). They explicitly criticize some of the works also discussed in this review (11, 218) for their limited usefulness when it comes to generating and designing actual user experiences (rather than just describing existing ones).

Further, the previously encountered notions of direct and indirect manipulation are discussed, with a call for more direct manipulation to increase interface fluidity (5, 57, 215). The authors also make a case for equal treatment of representation (graphical representation) and manipulation (interaction) as brought forth previously by others (44, 218).

In their best-practice review, the authors specifically mention the movie "Iron Man 2," including a holographic visualization wall from the protagonist's lab into their list of fluid information visualization exemplars. A whole book dedicated to this kind of review of interfaces from popular pieces of entertainment was written by Shedroff and Noessel (172). Although not a piece of peer-reviewed academic literature, their book contains numerous examples of futuristic designs for interaction, split across a variety of categories (e.g., mechanical and visual interfaces, gestures, and augmented reality).

Other works, rather than providing a taxonomy for implementing interactions, provide either explicit support or assign a concrete role to interaction in the workflow of data analysis. The D3 family of visualization packages notes the importance of interactivity for data visualizations across the three papers that introduced academia to their technology. Satyanarayan et al. (166), in their paper on Vega-Lite, note the difference between exploratory and explanatory data visualizations and the differing needs visualization software has to satisfy, depending on what kinds of visualization they want to support: while *exploratory* visualizations are often geared towards analysts aiming to get an initial impression of a dataset to uncover potential paths towards insights, explanatory visualizations are used to convey an insight to more general audiences (although, of course, exploratory visualizations are not exclusively made for experts). Prominent examples of such explanatory visualizations are those found in news media and text books. Satyanarayan et al. (166) argue that high-level languages such as ggplot2 and D3 are mostly geared towards explanatory data visualization, and thus develop a rationale for the development of their high-level Vega-Lite to facilitate more interactivity in data visualizations created with D3. Finding this focus on interactivity specifically in a web-deployed visualization framework is not surprising; after all, modern web standards have a significant amount of native interactivity already built into them.

In *Making Data Visual*, a visualization guide specifically for data analysts, Fisher & Meyer (89) do not address interactivity as an area of its own; rather they identify a number of analysis tasks and how they can be supported by multiple linked views (MLVs) , a concept previously explored by Roberts (159). MLVs enable the user to select a subset of data and then see multiple visualization types concurrently in in order to discover patterns, trends, and other insights by looking at the data from different angles at the same time. While the act of selecting already constitutes an interaction in and of itself, it is not necessarily classified as an interaction type per se.

In summary, various levels of granularity have been applied to categorizing interaction techniques in data visualization, from low-level, i.e., focusing on individual actions (145, 163), to high-level, i.e., focusing on goals and objectives of the user (103, 188, 218). Any interaction model developed hereafter needs to account for this granularity and, subsequently, incorporate some kind of level structure to describe interaction types across different user needs. While interaction has not received the amount of attention in previous research that it needs to qualify as an adequately studied area of data visualization (81), the increasing use of high-resolution displays (plus the recent mass-market rise of VR), increased computing power, and heightened awareness of the importance of data literacy provide a fertile ground for the exploration of interaction types for data visualizations in VR.

After presenting our synthesis of interaction types in the literature in section 2.2.2, we will investigate prior research on interface and interaction design for VR in section 2.3.

2.2.2 An integrated overview of interaction taxonomies

Our goal in this section is to present an interaction typology that is medium-agnostic and can be implemented in VR. Based on the literature review in section 2.2.1, we present a synthesized interaction typology for data visualization, see Table 4. Naturally, the terms across the columns do not necessarily line up 100%, and important semantic differences make the alignment of interaction types from so many papers challenging, but we think it is possible to explore common streams of thoughts in the literature. As evident in section 2.2.1, a large variety of frameworks for interaction techniques in data visualization have been proposed. Some, like the taxonomy of distortion techniques by Leung and Apperley (128) and later revisited by others (56, 117), the widely cited Table Lens paper (157), or the FilmFinder (6), are limited to one specific technique and its implementation(s). Others, like the framework by Ward, Grinstein and Keim (211), introduce umbrella terms, grouping various techniques together. Yet others, like the semantic space of input devices proposed by Mackinlay, Card and Robertson (130), describe a more low-level, almost physical framework. We acknowledge the inherent challenge of aligning these terms but nonetheless believe that a visual representation as a table can provide valuable insights on what interaction types are most commonly considered in the literature. A table, itself a simple but strong visualization type, is a great way of highlighting similarities and differences across this multitude of sources. This summary of previously proposed taxonomies is in no way exhaustive but, to our knowledge, the most comprehensive one to date.

Table 4. A comprehensive overview of interaction types in data visualizations in six parts. New interaction types in the DVL-FW are marked in a darker blue. Column headers have last name(s) of author(s) and reference number.

Column	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Börner (31)	Filter	Zoom	Link and brush	Distortion			
Börner, Bueckle and Ginda (32)	Filter	Zoom	Link and brush	Distortion			
Brehmer and Munzner	Filter					Encode, Change	Select
Buja, Cook and Swayne (44)		[Focus (viewpoint variable,	Linking views				
Chuah et al. (57)	Hiding, dynamic transpare		s[Linking]				
Dix and Ellis (73)	Filter, [Hiding/el ision]	Zoom	Brushing, Linking represent	Fish-eye view	Sort	[Same Represent ation,	Select
Heer and Shneiderma n (103)	Filter				Sort	Visualize	Select
Keim (117)	Filter	Zoom	Link and brush	Distortion			
This dissertation	**Filter**	**Zoom**	**Link and brush**	**Distortio n**	**Sort**	**Visualize /encode**	**Select**

Column	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Keim (118))	Filter	Zoom	Link and brush	Distortion			
Roth (163)	Filter	Zoom			Sequence	[Resymbol ize]	
Shneiderman (176)	Filter	Zoom					
Tweedie (197)	Hiding/filt ering			Algorithmi c transform ation	Reorderin g		
Unwin (201)		Zooming	Selection with linking			Multiple views	Selection with linking
Ward, Grinstein and Keim (211)	Filter, reconfigur e-filter	Elaborate-zoom	[Connect]	Elaborate-distortion	Reconfigur e-sort	Encode	Select
Ward and Yang (212)		Zoom		Distortion			Selection
Wilkinson (215)	Filtering (categorica l,	Zoom	Brushing and linking	Lensing	Categorica l reordering		
Yi et al. (218)	Filter	Abstract/e laborate			Reconfigur e	Encode	Select
This dissertation	**Filter**	**Zoom**	**Link and brush**	**Distortio n**	**Sort**	**Visualize/ encode**	**Select**

Row	8	9	10	11	12	13	14	15	16	17
Börner (31)			Details on demand			Projection	History	Extract	Overview	
Börner, Bueckle and Ginda (32)			Details on demand		s	Projection	History	Extract		
Brehmer and Munzner (40)		Navigate		Annotate	[Arrange]					Derive
Buja, Cook and Swayne (44)	[Focus (pan)]				Arrange views					
Chuah et al. (57)				Change object (set) property/paint						
Dix and Ellis (73)	Pan		]Accessing Extra Informa							
Heer and Shneiderman (103)		Navigate		Annotate	Coordinate views					Derive
Keim (117)						Projection				
This dissertation	**Pan**	**Navigate**	**Details on demand**	**Annotate**	**Arrange/ coordinate**	**Projection**	**History**	**Extract**	**Overview**	**Derive**

Row	8	9	10	11	12	13	14	15	16	17
Keim (118))						Projection				
Roth (163)	Pan				Arrange	Reproject		Retrieve		Calculate
Shneiderman (176)			Details on demand				History	Extract	Overview	
Tweedie (197)		Animated navigation		Labeling/Boolean encodin						
Unwin (201)			Querying		Variation of					
Ward, Grinstein and Keim (211)	Pan	Navigation, rotate								
Ward and Yang (212)	Elaborate-pan	Navigate								
Wilkinson (215)	Pan	Navigate		Annotation			Reproducible log			
Yi et al. (218)										
This dissertation	**Pan**	**Navigate**	**Details on deman**	**Annotate**	**Arrange/ coordin**	**Projection**	**History**	**Extract**	**Overview**	**Derive**

9

Row	18	19	20	21	22	23	24
Börner (31)						Search and locate	
Börner, Bueckle and Ginda (32)						Search and locate	
Brehmer and Munzner (40)			Record		Aggregate		
Buja, Cook and Swayne (44)							
Chuah et al. (57)	[Focus], highlightin g			Manipulate position, (rotation), size			
Dix and Ellis (73)	Interactive highlightin g, temporal fusion,				Aggregate		
Heer and Shneiderman (103)			Record				
Keim (117)							
This dissertation	**Highlighti ng**	**Rela te**	**Record**	**Manipulat e**	**Aggregate**	**Search and locate**	**Anima te/ replay**

Row	18	19	20	21	22	23	24
Keim (118))							
Roth (163)							
Shneiderman (176)		Relate					
Tweedie (197)							
Unwin (201)							
Ward, Grinstein and Keim (211)	Highlighting						
Ward and Yang (212)							
Wilkinson (215)				Manipulate (node dragging), rotating			Frame Animation
Yi et al. (218)		[Connect]					
This dissertation	**Highlighting**	**Relate**	**Record**	**Manipulate**	**Aggregate**	**Search and loca**	**Animate/ replay**

2.3 Data visualizations in VR

Modern VR hardware poses challenges: limited screen space and resolution (with a field-of-view around 110 degrees on the most expensive devices and only 2880 x 1600 pixels); immature ergonomics of the head-mounted device (HMD); proneness to induce motion sickness at low frame rates. Additionally, some of its shortcomings have thus far prevented it from becoming a mass-market medium: the lack of choice for powerful standalone HMDs that do not need to be tethered to a powerful, expensive computer; the absence of a killer app that would prompt many customers to buy a VR system; the monetary investment. However, VR has been a subject of research for data visualization for a while. In 2016, two essential commercial VR setups were released: the Oculus Rift (by Oculus) and the HTC Vive (by phone company HTC and the software company Valve). Valve also owns the video game platform Steam, which offers access to a majority of commercial video games. The recent announcement of the Valve title “Half Life: Alyx” (203) has been met with great excitement and might become the killer app that prompts a new wave of VR purchases by end users (99). A comprehensive history of VR and how its various technological components were combined to eventually become the VR as we know it today, starting from the specification of linear rendering in the 1500s to the release of the HTC Vive in 2016, has been written by Sherman and Craig (173).

In the previous section, we found that interactions in data visualization have not received the same amount of attention in the literature as visual encodings, user task types, etc. Now we investigate previous work on interactions in VR visualizations.

2.3.1 Interaction frameworks & embodiment

Many definitions of VR exist; some focus on a reality-virtuality continuum (139), some on internal factors such as telepresence and experience (190), and yet others on the environments VR users get exposed to (38), and yet others on everyday life usage (132) and the whether there is "self-representation" of a physical body as a virtual one (181) (p. 131). We use the following operational definition of VR:

VR is a medium where a user wears a head-mounted device (HMD) with one display per eye, controls a virtual camera through head movement. Optionally, two controllers with a variety of buttons, triggers, and a touchpad can be used to capture user input. VR enables the user to interact with virtual space and virtual object through natural input by virtue of tracking the position, rotation, and acceleration of their head and hands.

This definition captures the immersive nature of VR, where the experience requires more physical user input than just the keyboard and mouse paradigm, and where the users "perform tasks relative to the body (egocentric)" (113), p. 202. While the technical specs and cost of VR hardware has changed substantially over the past decades, the operational definition given above holds true whether we discuss setups from the 1990s or the late 2010s.

VR is never static; animation and movement are built-in. Even the most basic input in VR constitutes an interaction (73):

"VR without interaction is simply computer graphics! Where visualisations do not support full 3D navigation […], the objects within that space must be interactively spun, moved or otherwise manipulated to reveal their nature." (p. 126)

Even the most basic and affordable VR gear offers head-tracking and, based on this, the use of the user's gaze to select items and perform basic navigation. This allows VR developers to create experiences even when no controllers are available and if the user's position in space cannot be tracked. Likewise one can argue that no visualization is ever static, even if the interaction is not built in. Yi et al. (218), for example, observed that even static data visualizations can be interacted with by physically moving closer and looking at different parts of the visualization. VR thus requires interactivity by default. Even in the absence of any of the typical interaction techniques such as filtering or searching, the simplest possible VR experience needs a movable camera. VR is a medium where a virtual camera not only simulates a first-person view but also allows for a more intuitive interactions through tracking of body parts. While there are hardware and software solutions to provide a more immersive experience with a 2D screen (such as eye tracking support in some video games), the experience of six degrees of freedom (6-DoF, see also Glossary) means not only more perceived liberty of movement but also creates more dimensions for user interaction. VR thus presents a departure from what is commonly called the “window, icon, menu, pointing device" (WIMP) interaction style popularized in the 1970s and could be part of a larger paradigm shift towards a new generation of interfaces. A well-written, very brief history of the WIMP interface metaphor was written by Van Dam (206). To capture a variety of interfaces substantially different from the common desktop metaphor, Jacob et al. (113) propose the framework of reality-based interaction (RBI):

“We believe that all of these new interaction styles draw strength by building on users’ pre-existing knowledge of the everyday [world]. They employ themes of reality such as users’ understanding of naïve physics, their own bodies, the surrounding environment, and other people.” (p. 201)

The authors describe how RBI themes such as naïve physics, body awareness and skills, environmental awareness and skills, as well as social awareness and skills can be leveraged by interaction designers, and present four case studies highlighting the concept of RBI as a tool to evaluate existing designs, e.g., two experiments on walking in VR (182, 202), also discussed in section 2.3.3 of this dissertation.

Essential is the awareness of tradeoffs between these RBI themes and the desired qualities of the implemented systems, e.g., expressive power, efficiency, and practicality. RBIs can be analyzed and evaluated using this notion of tradeoff, although the authors admit that their framework does not include a methodology to do so. We will review previous work on VR user study methodology in section 2.4. While Jacob et al. (113) propose an overarching framework to understand post-WIMP user interaction as a whole with a focus on diverse physical and cognitive abilities of humans, others have focused on individual interaction styles. For example, Coutrix and Nigay (63) present their Mixed Interaction model, aiming to unify existing research on mixed reality systems while focusing on input modalities consisting of a device d and a language l. Subsequently, they investigate the model's descriptive, generative, and comparative power, and implement it in a mixed reality video game. The Mixed Interaction model extends the Instrumental Interaction model presented several years earlier by Beaudouin-Lafon (25), where a set of rules to guide the development of post-WIMP interfaces is discussed and implemented in a text search engine.

Aside from the affordances of (post-)WIMP interfaces, another important concept in understanding interaction for VR is embodiment, and embodiment-inspired interaction frameworks (90). Klemmer, Hartmann and Takayama (120) present five themes on interaction design and how we can leverage a user's body for better interface design: thinking through doing (emphasizing the possibility to offload cognition to physical action), performance (the high throughput achievable with physical input, e.g., our hands), as well as visibility, risk, and thickness of practice. The authors note that GUIs unify practices across multiple activities, e.g., practices that are very different in real life become unified when using a WIMP interface, such as making music, writing text, and editing photos. Further, they emphasize the versatility of the human hand as an input device with its expressive (gesture), sensing (touch), and operational (grasping, swiping, etc.) capabilities. For example, Buxton and Myers (51) ran two experiments on two-handed input and parallel task completion and, among other things, found that using two hands for positioning, scaling, and selection tasks can significantly decrease completion time. Buxton (50) further elaborates that interfaces (at the time) were designed with a visual primacy, largely ignoring other sensory systems such as the ear, and not making adequate use of human dexterity. Abstracting from the findings in his experiments, he later generalized his views about how visible and tangible interfaces make use of physical human properties (such as motor skills) to how these interfaces reflect cognitive and social properties (48). In a literature review published more recently, Lee et al. (124) propose four design dimensions (individual, technology, social interactions, interspace between person and technology) to categorize non-WIMP and natural user interfaces (NUIs). They suggest improving interactions in data visualization by going “beyond mouse and keyboard”, providing “high freedom of expression”, paying attention to “social aspects” of technology use, reducing “the gap between person and technology”, and strive to understand “people’s behavior” better (p. 2697). Witmer and Singer (217) also note that natural interaction with a virtual environment can increase immersion in that environment.

In this section, we investigated a select number of works on interaction frameworks and embodied interaction beyond the WIMP interface paradigm. Since one of the key properties of VR is input via head and hand rotation (and movement, depending on the setup), we consider this line of research essential in discovering and understanding previous attempts at formalizing and testing the usefulness of new interfaces and the input devices they need as well as affordances they bring.

2.3.2 Immersive Analytics

In this section, we focus on a more recent line of research that heavily involves VR. Immersive Analytics (19, 55, 60, 79, 134, 179) is an evolving practice where reality-extending technologies, i.e., VR, augmented reality (AR), or mixed reality (MR), collectively known as XR, are used to visualize and investigate abstract data using 6-DoF VR controllers in a virtual environment, or what Bowman et al. (38) called "information-rich environments" where the virtual objects are representations of non-physical entities. This is in contrast to many scientific visualization applications where the data is inherently spatial, and thus presents a natural use case for human depth perception and 3D user input. Use cases include assessing risk in mining with seismic data (116), protein-docking (13), astronomy (74, 75), computer-aided design (183), geographical information science (106), and health care (110). Bryson (42) made the case for using VR for scientific visualizations by pointing at the affordances VR offers for interaction with complex phenomena and their representations in data: "We want to create the effect of interacting with things, not with pictures of things" (p. 63).

Chandler et al. (55) propose Immersive Analytics as “a new facet of data analytics research that seeks to unify these efforts [i.e., leveraging emerging interface and display techniques such as VR and AR for data analysis] to identify the most enabling aspects of these emerging natural user interface and augmented reality technologies for real-world analysis of data” (p. 2). A comprehensive coverage of this emerging research field was written by Marriott et al. (134), with articles on multisensory input (136), interaction (47), collaboration (29) alongside a wide selection of case studies for the application of Immersive Analytics to solve real-world problems (70).

In terms of implementations of Immersive Analytics, Simpson et al. (179) present a prototype to analyze the output of a Dynamic Integrated Climate-Economy (DICE) model for environmental decision-making. Hurter et al. (107) present FiberClay, a system to visualize 3D trajectories in VR. Batch et al. (23) present a field study using economic data with multiple phases of refinement (design stage, in-the-wild deployment, and summative) with data-domain expert users. While there are many hypotheses and findings in this study, two noteworthy ones are: when presenting their own visualizations to an external observer, the participants rearranged their visualizations into chronological order, and would also build more complex visualizations then. Additionally, users reported high engagement while voicing concerns about the usefulness of a heavy VR setup for everyday data analysis as an economist (or any data analyst for that matter). We investigate VR user study methodology more deeply in section 2.4.

Batch et al. (23) built their VR analysis environment using ImAxes (61), a framework for authoring visualizations for abstract data in VR, allowing the user to build 2D and

3D scatterplots, multiple linked visualizations, and parallel coordinate plots. ImAxes was built using Unity. Implementing the direct manipulation paradigm (175) in VR, ImAxes is centered around data axes as "embodied affordances" (p. 71) where the location and orientation of these 3D data axes can be used to create a variety of 2D and 3D visualizations in virtual space. This makes ImAxes fluidly interactive (81), using a formal grammar to allow the user to construct visualizations by manipulating data axes alone, without menus. Similarly, the Immersive Analytics Toolkit (IATK) developed by Cordeil et al. (60), likewise built with fluid interaction (81) in VR in mind, implements four design goals: expressiveness, simple authoring, scalability, and integration of MR (support of augmenting a user's desktop with virtual space). It comes with a high-level GUI as well as a low-level API for finer control over the created visualizations, and allows for a variety of interaction techniques such as filtering, brushing and linking, and details-on-demand, with animated transitions and animations. While ImAxes (61) was developed for end users in VR, IATK is a package for designers using Unity, and is thus built to make use of Unity's GUI and programming environment.

Immersive Analytics is at the forefront of data visualizations for VR, and the implementations outlined in this section use very similar setups to the ones developed later in this dissertation (chapters 6 and 7). Investigating this string of previous work highlights the affordances and limitations to VR for data visualization, and gives a firm understanding of the advantages and disadvantages of modern hardware and software solutions available to us (namely, Unity and commercially available VR gaming headsets).

2.3.3 User interface design in VR

In section 2.3.1, we investigated the WIMP interaction paradigm where windows, icons, menus, and pointers are the primary tools for a user when interacting with an application. WIMP is optimized for 2D screens. In 1993, Robertson, Card and Mackinlay (161) argued that technological advancement (especially speed vs cost) and application demand drive UI innovation. They voice their disappointment with the prevalence of the WIMP interface paradigm, established for decades at this point and famously captured in essence, for the first time, in Douglas Engelbart's famous Mother of All Demos almost 30 years before (78), see Figure 9.

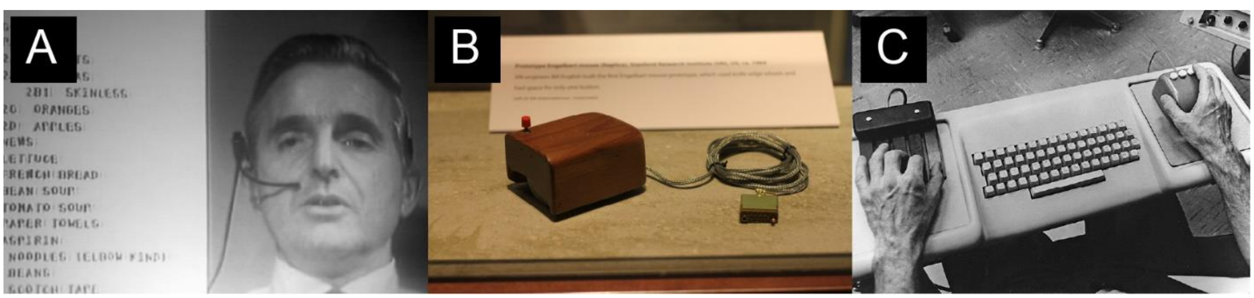

*Figure 9. **A**: Douglas Engelbart during the Mother of All Demos. **B**: Prototype of the first computer mouse. **C**: The keyboard and mouse setup used by Engelbart during the Mother of All Demos.*

The authors present four methods to leverage recent advances in computer graphics in order to cheapen the cost of information retrieval: larger workspace, agents, real-time interaction, and visual abstractions. Going back to their previously discussed work on the Information and the 3D Exploratory (104, 160), they formulate solutions to ongoing challenges of optimizing information management and retrieval through the use of advanced computer graphics, and argue "that information access will be a primary force in shaping the successor to the desktop metaphor" (p. 57). The authors' call for the research and development of new post-WIMP UIs has since come to fruition. There is a variety of research on tangible (20, 105, 108), non-command (149), and the previously discussed reality-based interfaces (113). Some VR applications do not use WIMP elements at all and implement direct manipulation instead (61). Direct manipulation has been described as providing direct engagement between user and system as well as low "semantic distance" between input and output vocabularies (109). The authors then famously introduce the concept of golf of execution and the gulf of evaluation, where both can be "bridged" (p. 318) via good usability to "minimize cognitive effort" (ibid.).

More recently, authors have been investigating post-WIMP interfaces in areas of entertainment such as video games, TV shows, and movies. Shneiderman (177) cautions against merely copying the "richness of 3D reality" (p. 12) and warns of the challenges this could introduce into the interface, i.e., occlusion, complex user action, and confusing navigation. He suggests restraining a user's action instead and calls for making 3D interfaces "better than reality" (ibid.) by proposing a set of guidelines for designers of 3D interfaces. Investigating a range of 3D applications relevant at the time (such as There.com and ActiveWorlds), he gathers 22 suggestions for developing user-friendly 3D interfaces. In their brilliant survey of interface design in science-fiction movies and TV shows, Shedroff and Noessel (172) dedicate a whole chapter to visual interfaces. After investigating the choice of typeface, color, display shape, and visual effects like glow and transparency, in several pieces of entertainment, they discuss 3D file system and the advantages and disadvantages associated with them. Pointing at humans' good spatial memory, they caution again over-challenging less “spatially adept” users who “frequently lose their keys” (p. 62). The intention behind this work is to provide guidance to interface designers by helping them understand how interfaces are made to look futuristic yet realistic. Similarly, to bridge the gap between VR designers and those who implement their designs, Tanriverdi and Jacob (193) propose the Virtual Reality Interface Design (VRID) model and methodology comprised of five components (graphics, behavior, interaction, mediator, communication). In a high-level design phase the graphics, behaviors, etc. of objects are determined, which is then followed by a low-level design phase for implementation details whose output then guides the development of a VR interface in software. Sherman and Craig (173) devote four full chapters to interfaces and interactions in VR. We further review VR UIs in video games and a selection of commercial applications in section 2.5.

Two challenges when implementing traditional WIMP interfaces in VR are the low resolution and the added depth of a 3D pointer. Even state-of-the-art VR headsets can only display a comparatively small number of pixels, making it hard to, e.g., make small text legible and display icons in a satisfying quality. In the Supporting Information (section Modern VR hardware), we provide an overview of a select few HMDs with pertinent technical specs. In terms of added depth, the challenge of adapting a 2D button for a 3D environment is often solved with the usage of visualizing a ray into the scene coming out from the VR controller, which can then be used to hover over and click a button or use a slider, see Figure 11. However, this added depth when using a WIMP interface in VR can make it awkward to use the interface at all, especially when parts of the UI become occluded or if the user has trouble aiming correctly. Also, since 2D buttons in VR have a flat shape, there is a considerable angle from where the effective size of the buttons is very small for the user to hit; on a 2D interface, buttons always face the user at a rotational difference of 0 degrees. In VR, if the orientation of the button is fixed, then, when the user moves through the virtual space, the button can become almost invisible from certain angles. Many 3D video games solve this problem by making the UI part of the virtual environment as object-fixed UI. Many VR applications have buttons as part of a controller menu, allowing the user to determine the orientation of the button with their own hands, see Figure 10. While this approach allows developers and designers to leverage our learned abilities to use classic WIMP interfaces, it can still pose challenges, especially to novice users.

Figure 10. An object-fixed, contextual UI panel, linked to a VR controller (left).

Additionally, while implementing a mouse click on a button in 2D environments is a standard functionality available in most SDKs and platforms, such as browser-deployed visualizations (27, 37), building a VR pointer, on the other hand, as of the time of this writing, comes with a significant amount of coding and is thus not easily accessible. It has therefore been explicitly recommended to avoid using 2D UI elements in VR but to enable the user to use VR controller input instead (107). We will review interactions and UIs in exemplars from video games and commercial applications in section 2.5.

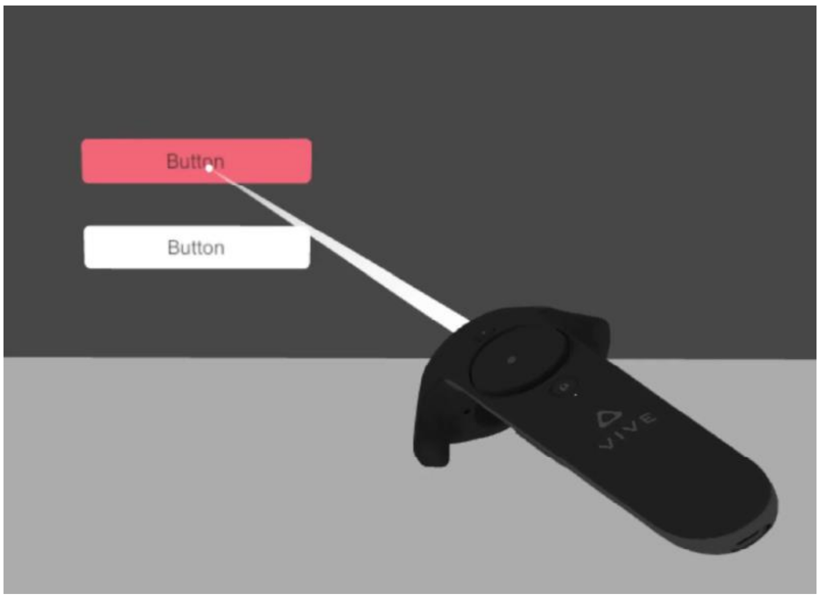

Figure 11. Basic VR pointer in Unity.

2.4 VR user study methodology

In this section, we investigate prior work on designing and performing user studies for interaction design and for VR. We are also interested in reviewing individual data collection tools. Table 5 provides a summary of user study setups by several dimensions. The goal of this section is to understand sample sizes, conditions, hardware and software used, what telemetry data was captured (if applicable), and research designs employed to lay the groundwork for chapters 5 to 7. While a full review of VR user study methodology is out of scope for this dissertation, we investigate a few select implementations to set the stage for the experiments run in this dissertation.

Table 5. Selected user study setups for VR.

Source	Witmer and Singer (217)	Slater, Usoh and Steed (181)
Study instruments	Experiment + questionnaire	Experiment + pre-, post-questionnaire
What is measured	Presence	Presence as sense of "'being there'" (p. 130)
How it is measured	Questionnaire with 7-point scale items	Questionnaire
Conditions	Psychomotor tasks, further elaborated in Bailey and Witmer (21), complex navigation, further elaborated in Witmer et al. (216)	Automatic transportation vs. going through virtual doors. Other factors: gravity, precipice, presence of virtual actor, numbers of levels
Hardware/software	**Controls**: stereo viewer (psychomotor tasks), Virtual Reality Flight Helmet + joystick (complex navigation). **Visual fidelity**: visually simple for psychomotor tasks, complex	DIVISION ProVision200, DIVISION 3D mouse, Virtual Research Flight Helmet
#subjects, type & demographics	152 across four experiments, college students (91 males, 61 females)	24 (across a variety of occupations)

Source	Seay et al. (170)	Slater et al. (180)	Cordeil et al. (62)
Study instruments	Experiment + pre, post-questionnaire	Experiment + pre-questionnaire, spatial awareness test, post-questionnaire	Experiment + pre-questionnaire (demographics, prior VR experience) and post-questionnaire
What is measured	Simulator sickness	Influence of immersion on performance	HMD vs. CAVE2 for collaborative analysis of abstract data (networks); task completion time,
How it is measured	Questionnaire	Questionnaire + task score from experiment	Telemetry + questionnaires, audio recording
Conditions	FOV (one screen vs. three screens), display fidelity (stereo- vs. monoscopic), user role (driver vs.	Level of immersion (exocentric with screen vs. egocentric with HMD), level of environment	HMD condition (nine teams (9 teams), CAVE2 condition (8 teams), in teams of two)
Hardware/software	NAVE (Non-expensive Automatic Virtual Environment) with three 8' x 6' screens at 120-degree angles, joystick	DIVISION ProVision100, Virtual Research Flight Helmet, , DIVISION 3D Mouse	Oculus Rift DK, CAVE2
#subjects, type & demographics	156 undergraduates (133 males, 23 females)	24 (16 males, 8 females)	34 (21 males, 13 females)

Source	Kwon et al. (122)	Prabhat et al. (154)	Millais, Jones and Kelly (140)
Study instruments	Experiment + post-questionnaire	Experiment + pre- and post-questionnaire	Experiment
What is measured	Impact pf layout, rendering, and interactions for graph exploration; task completion time, accuracy, number of interactions	Performance and subjective evaluation of three display conditions while analyzing biological	Task-workload and user satisfaction in 2D vs. VR
How it is measured	Telemetry + post-questionnaire	Correct answers in experiment (via video recording), qualitative evaluation in post-questionnaire	Audio recordings from think-aloud session in experiment, NASA TLX questionnaire administered post-
Conditions	2D layout, spherical layout without depth routing, spherical layout with depth routing	Desktop, Fishtank, CAVE (in all possible orders)	2D vs. VR
Hardware/software	Oculus Rift DK2, NVIDIA GTX 980 graphics card	Linux computers with NVidia 4500G graphics card, displays: IBM P260 (16” x 12”) for Desktop and Fishtank, Marquee 9500 (8’ x 8’) for CAVE, 3D joystick (all	2D: laptop with 15” screen at 2560x1600 resolution + mouse VR: Google Daydream VR HMD + controllers
#subjects, type &	21 (12 males, 9 females)	12 (9 male, 3 female)	16 (9 female, 7 male, 1 prefer not to say)

Source	Batch et al. (23)	Steed et al. (189)	Pfeuffer et al. (152)
Study instruments	Experiment + demographics survey, exit interview	Experiment + pre- and post-questionnaire	Experiment + pre-questionnaire
What is measured	Qualitative feedback/think-aloud utterances	Presence and embodiment in VR "in the wild"	Identifying characteristic body movements in VR (pointing, grabbing, walking, typing)
How it is measured	Video recording, screen-recording	In-app questionnaires, sent back via the internet, limited telemetry	Telemetry + pre-questionnaire
Conditions	All in VR, expert users visualize their own data	All in VR. Modifiers: avatar yes/no, being asked to tap along yes/no, singer looks at user yes/no,	No conditions
Hardware/software	HTC Vive + controllers	Samsung Gear VR or Google Cardboard (Android-based)	HTC Vive with eye tracker, NVidia GTX 1080 on Windows 10 PC
#subjects, type & demographics	Pilot, Formative, Summative, Summative 2: All expert users	59	22 (18 males, 4 females)

Source	Cikajlo and Potisk (58)	Mottelson and Hornbæk (144)
Study instruments	Experiment + post-questionnaire, pre-screening survey	Experiment
What is measured	Performance in patients with Parkinson's disease when placing cubes	Feasibility, validity, and scalability of out-of-lab VR user studies
How it is measured	Telemetry (e.g., time from first touch to placement of cube) + questionnaire on effort	Telemetry via the internet
Conditions	Desktop vs. VR	In-lab (31 subjects) vs. out-of-lab (57 subjects)
Hardware/software	Desktop: LCD screen + Leap Motion controller VR: Oculus Rift CV1 + Leap Motion controller, NVidia GeForce graphics card	In-lab: HTC Vive Out-of-lab: Google Cardboard VR glasses
#subjects, type & demographics	20	88 subjects (47 males, 41 females)

2.4.1 Measuring presence

An essential element of experiencing VR is presence. Steuer (190) defines presence as "the sense of being in an environment" (p. 77), further proposing the term telepresence as "the experience of presence in an environment by means of a communication medium" (ibid.). This is opposed to the concept of VR as an ensemble of hardware devices such as HMDs and bespoke input devices. A continuous stream of sensory input needs to be supplied to the user, either as "patterned sensory impressions" (181) (p. 131) or "[coherent] stimulus flow" (217) (p. 226).

Measuring presence has been an essential task in collecting data from user studies and is usually measured via self-report. Witmer and Singer (217), for example, propose two questionnaires on presence (PQ) and immersive tendencies (ITQ) to quantify experience of presence and the experience of involvement (high level of attention and focus) and immersion (exclusion of the physical world as a result of a constant flow of stimuli from a virtual environment), respectively, using a 7-point scale. The authors discuss involvement and immersion as prerequisites for presence, where immersion is based on the experience of a person rather than a descriptor of a specific VR technology as proposed by Slater et al. (180), discussed later in this section. They further synthesize four factors from previous work that influence involvement and/or immersion and, thus, presence: control, sensory, distraction, and realism factors. They corroborate these factors over four user studies with the PQ and the ITQ, finding significant correlations between a majority of factors with the total PQ and ITQ scores. They argue that while no conclusive evidence has been produced yet, increasing presence can lead to increased learning and performance.

The concept that presence follows immersion is also shared by Slater, Usoh and Steed (181). The authors present a user study with 24 participants across four conditions to test how the experience of presence, or "'being there'" (p. 131), is related to "stacking environments" (p. 133). Interested in exploring the importance of transitions between environments, they develop a model to quantify presence as a result of being n steps removed from physical reality. In an experiment with 24 subjects, gravity, the presence of virtual physical danger, virtual actors, and the number of stacked environments (or levels) were treated as independent variables to observe presence. To minimize contact between the subject and the researcher, a fairy tale-style plotline was presented where participants had to perform certain tasks in the virtual environment. Half of the subjects went from level to level by putting on a virtual HMD while in VR while the other half did so by going through a virtual door. A pre-questionnaire asked about visual, aural, or kinesthetic primacy and perceptual position of the user to determine their proneness to certain senses. A post-questionnaire assessed the participant's experience of presence. The study finds that presence correlates positively with visual and kinesthetic primacy, and negatively with aural. Furthermore, a positive correlation was observed between presence and transitions via virtual HMD, and a negative one between presence and going through doors. The authors emphasize the shortcomings of assessing presence via self-report.

Interplay between presence and other factors for VR have also been studied. Seay et al. (170) use the previously discussed ITQ (217) as well as a Simulator Sickness Questionnaire (SSQ), both administered pre-treatment, together with the aforementioned PQ (217) and the second part of the SSQ post-treatment in their study of simulator sickness (measured as nausea, oculomotor stress, and disorientation) with 156 undergraduate student subjects. There were four experimental groups (field of view, or FOV at one screen vs. three screens, stereo- vs. monoscopic display) with pairs of participants being assigned either a driver or passenger role in a flight-simulator. The authors find no correlation between the scores on the ITQ, PQ, and SSQ, no difference between the cohorts. However, they do find that a higher FOV or being the driver increased the PQ score. Users with high FOV reported higher levels of nausea. The authors argue that motion without interaction from passenger added to the feeling of nausea, along with the rapidness of the content in the peripheral vision when using FOV displays. The authors thus consider a large FOV a "double-edged sword" (p. 300), and that the feeling of presence and experiencing sickness are not mutually exclusive as one might expect.

2.4.2 Measuring task performance

In later work, and building on research by Mizell et al. (141) on evaluating VR vs. desktop workstations, Slater et al. (180) studied how immersion influences performance in a user study with 24 participants in four cohorts where the subjects have to recreate the moves seen on a virtual 3D chess board on a physical one. Specifically, they wanted to test differences between VR and 2D screens for understanding geometric structure, knowledge transfer, and how immersion influences performance. Subjects with prior experience in VR were evenly distributed among cohorts. After a pre-questionnaire about demographics and prior exposure to VR and video games, subjects listened to an introduction to the 3D chess board, a VR practice session in a virtual kitchen introduced the subjects to navigating and interaction in VR. In the following reproduction task, the participants first observed a virtual chess game, then had to recreate those moves before filling out a post-questionnaire about confidence about performance, nausea, and presence. The authors find that the level of immersion (in this case, the egocentric condition) and environment (in this case, realistic) had a significant influence on performance. Immersion (in this case, egocentric) was also significant for the sense of self-reported presence, but environment was not. The authors close by recommending further research into the spectrum of immersion that can achieved from combining different display media with different rendering settings to produce experiences across the spectrum between the “two extremes” (p. 171) they chose for immersive settings.

A useful way to measure task performance is telemetry, i.e., extracting position, rotation, and gaze data from an experimental VR setup. More recently, Cordeil et al. (62) compared the first generation of mass-market HMDs and CAVE2 systems (86) with regards to functionality, collaboration, and user experience for analyzing abstract data (connectivity of 3D networks). In a user study involving 34 participants in groups of two, divided into two cohorts (HMD and CAVE2), users had to perform two tasks in tandem (finding the shortest path between nodes and counting triangles), reporting their answers as a team. The authors measured task completion time, accuracy, and user experience (via a post-questionnaire), and find that users in the HMD condition needed significantly less time for their tasks without differences in accuracy. Also, they did not find shared focus strategies to have an impact on task accuracy but on completion time for the pathfinding tasks in both conditions. Further, using the HMD was reported to be more impersonal than face-to-face communication compared to the CAVE2. Furthermore, subjects in the CAVE2 condition whose head movements did not control the virtual camera were found to perform significantly less head movements than their partners. Assigning subtasks was more unclear in the HMD condition. In terms of usability, the study did not find any significant differences between the two conditions.

Kwon et al. (122) performed a user study comparing layout, rendering, and interaction with graphs in 2D and VR. They implemented a cursor paradigm where the cursor is controlled via mouse in the spherical graph layout and can be reset via a hotkey in three conditions: 2D layout, spherical layout without depth routing (i.e., edge bundles that are farther away appear darker), spherical layout with depth routing. In a user study with 21 participants, they measure completion time, accuracy, and the number of interactions (pointing, highlighting, selecting) for the four tasks (find common neighbors, highest degree, path, recall node locations) on four graphs (training, small, medium, large) in the study. The authors find that users in the spherical layout condition with depth routing performed significantly faster than those in the other conditions while using a significantly smaller number of interactions. While the choice to display a 2D graph in an HMD rather than on a 2D screen might be considered unfair handicapping of the users in that condition, the study highlights and quantifies the effect of different layouts and rendering on completion time, accuracy, and interactions needed to perform tasks in VR. In a post-questionnaire, the users expressed their preference and the perceived ease-of-use of the spherical layout with depth rendering of the other two conditions. Worth noting is the fact that this experiment was not implemented in Unity but in Unreal Engine, a software with advanced graphics capabilities with less work needed on the side of the developers that has traditionally been very popular for video game development.

Prabhat et al. (154) evaluate subjective preference and quantitative performance for biological data analysis for three display metaphors: Desktop, Fishtank, and CAVE. In a user study with 12 participants where every participant performed the tasks on all three platforms (with variations of the order in which the platforms were used), the authors find that the CAVE system significantly outperformed the other platforms with regards to performance score. Methodologically, two factors make this user study distinct from the others reviewed in this section: the extended length of this experiment, which is marked as two to three hours as per the study information sheet (probably due to the fact that each subject used all three platforms), and the high involvement of the experimenter, prompting subjects to perform tasks in as much detail as possible through repeated questioning. In a similar setup but with only two conditions (2D and VR), Millais, Jones and Kelly (140) compared user performance for visual data exploration in 2D and VR while measuring user satisfaction via questionnaires. They implemented a scatter graph as well as a parallel coordinate plot in 2D (using d3.js) and VR (using Unity), deploying them to a laptop and a Google Daydream VR HMD, respectively. The study finds that participants in VR reported feeling more successful as well as satisfied with their work than 2D users while reporting – somewhat expectedly – higher physical demand. Additionally, VR participants submitted fewer insights coded as incorrect.

2.4.3 Telemetry & remote data collection

While presence (2.4.1) and task performance (2.4.2) can be measured via qualitative methods such as observations and questionnaires, metrics like movement are easier to track with telemetry, i.e., the automated logging of quantitative data such as position, rotation, and interaction. Modern SDKs such as SteamVR and Oculus SDK, together with the scripting capabilities of Unity and Unreal Engine, allow research designers to create sophisticated logging functionality for their user studies. Work discussed here previously (23, 62, 122) has made extensive use of this feature. The previously discussed work by Batch et al. (23) also uses telemetry in combination with observation and interviews to measure task performance for VR setups. The authors of this study used the HTC Vive in conjunction with Unity, the same setup in this dissertation. Comparing the experimental setups of Cordeil et al. (62) and Batch et al. (23), it becomes clear that the authors of the more latter paper were able to construct more natural user interactions by virtue of the Vive controllers shipped with the HTC Vive system, while the setup in the former was a self-built one (using a Leap Motion hand tracker strapped to the user's head). As more advanced input devices become available at a better price, the ability of researchers to collect more data also changes. Similarly, Pfeuffer et al. (152) conducted a user study to identify users from body movements from four basic tasks in VR (pointing, grabbing, walking, typing). The authors collected biometrics data from 22 participants logging position, rotation, velocity, and angular velocity of both controllers and the HMD as well as collision information about the rays emitted from these devices. They find that head motions as well as the distance between the HMD and the VR controllers are the most reliable feature to identify a user. Similarly occupied with assessing body movements, Cikajlo and Potisk (58) performed a user study with 22 patients with Parkinson's disease a cube-placing task, they compare two cohorts (Desktop vs. VR), where subjects in the Desktop condition placed cubes while looking at a 2D screen while VR users did so while wearing an Oculus HMD. The hand motions for all users were captured with a Leap Motion controller. The study finds that participants in the VR group showed significantly faster performance, placed more cubes, and exhibited healthier tremor indicators while also registering higher motivation and interest from the VR cohort.

Experiments previously discussed in this review utilized telemetry setups in Unity or Unreal Engine (23, 62, 122), but all of these studies were (for the most part) conducted in lab environments. While there is an ample amount of user studies with 1000+ subjects for data visualization using remote data collection via the World Wide Web (101, 119) or in longitudinal studies (178), similar setups for VR are barely researched, presumably due to the lack of standards and the prohibitive cost of advanced VR equipment as well as the low capabilities of cheap VR. Similarly, "in the wild" user studies (162) for VR are few and far between. As an example, the previously discussed Immersive Analytics user study by Batch et al. (23) deployed a VR setup running their application in a federal agency for a short duration of time without a researcher present before continuing towards a more controlled data collection environment. Steed et al. (189) performed a user study with 59 participants via an Android app for Samsung Gear VR and Google Cardboard, two VR devices at the cheaper end of the cost spectrum of modern VR. Because of the low-spec hardware, and for reasons of physical safety when collecting data "in the wild", this study was limited to comparatively simple tasks (watching a virtual singer perform in a virtual bar while sitting). Another issue raised by the authors is the problem of ethics and consent when collecting data from participants using their application. The authors estimate that the app was installed a total of 400 times, and after filtering, there were 59 datasets usable for data analysis. The study finds that being asked to tap along in the virtual experience lowered the user's self-reported presence score (presumably because the avatar then started tapping their foot without user input). Further, seeing one's own avatar increased a sense of fear of being hurt from a falling object inside the experience. Whether the virtual singer looked at the user made no significant impact on presence. Telemetry data (head rotation only) was recorded but not used for data analysis. The authors close with remarking that developing an “in the wild” research design requires more effort than a lab research design due to the need to optimize for cheap user hardware and the infrastructure needed to collect telemetry data remotely. In a similar study but with an additional in-lab condition, Mottelson and Hornbæk (144) compared effect sizes between in-lab (n = 31, using an HTC Vive) and out-of-lab (n = 57, using Google Cardboards on smartphones) cohorts. Specifically, they designed three VR tasks and experiences (pointing, 3D tracing, body ownership illusion) and measured differences in effect sizes between the cohorts. The study finds that while completion times, accuracy, and throughput for the pointing task were significantly different for the in-lab subjects, users in both cohorts were able to perform the tasks without an experimenter, and that effect sizes between sub-conditions in the 3D tracing and body ownership illusion experiences were similar for both conditions. The authors see this as proof that out-of-lab, crowdsourced VR user studies are feasible.

2.5 VR interaction and UI exemplars from video games

2.5.1 Interactions

All VR visualizations (and most printed maps) are, by default, interactive. At the very least, a VR visualization allows the user to move their head, controlling a virtual camera. What Yi et al. (218) as well as Spence (186) call “passive interaction” for static visualizations (such as looking at it from different distances) becomes rather active interaction in VR. Even a simple change of gaze leads to adjustments on the backend of the visualization. A rich ground for innovation in user interaction is the world of entertainment. Since the early 1990s, 3D video games have provided a financial and creative incentive for companies to develop user interactions and UIs that seamlessly support gaming experiences. Many VR user studies reviewed in section 2.4 asked questions about the gaming experience of their subjects as gaming is likely to be the most prevalent activity involving 3D most users have encountered. The previously discussed prevalence of the WIMP paradigm since the 1980s led to a certain uniformity in UIs in video games as video game developers copied what worked well so as to avoid the high cost of changing interaction paradigms (24), and as UI frameworks in video game engines became standardized as well. To that effect, most UI elements have not changed significantly since the early days of 3D gaming. Of course, VR games present a significant challenge to this WIMP paradigm. The limited screen resolution and complete enclosure of the visual field prohibits the player from engaging in user interfaces traditionally designed for keyboard and mouse or controller. Similarly, within non-VR games, there is a breadth of implementations for user interactions with the two most prevalent forms of user input being mouse/keyboard, and controllers. Video game genres like first-person shooter (FPS) can be played with both types of input devices, while strategy games usually require a mouse and keyboard to be played in a satisfying manner.

2.5.2 UI design dimensions

UIs in video games can be categorized along several dimensions: fixture, association, level of aggregation, manipulation directness, persistence, and visual encoding. UI can be screen-fixed, world-fixed, object-fixed, or user-fixed (degree of fixture), spatially or visually explicit (association), minimal, distributed, or complex (level of aggregation), susceptible to direct manipulation or indirect manipulation (degree of directness), displayed continuously or just contextually (degree of persistence), and can include linguistic, pictorial, and geometrical symbols (degree of visual encoding). This distinction builds on the design space for information display in Information-Rich Virtual Environments by Bowman et al. (38), who considers display location, (fixture in Table 6), association (spatially explicit, visually implicit/explicit), and level of aggregation (how much information to visualize in one display). We extend this triple of design dimensions by three more: manipulation directness (how the player interacts with the display, if at all), persistence (how long it is visible), and visual encoding (what graphic symbols are used to construct the display). With these six dimensions in mind, we can (de-)construct a majority of UI elements in 3D applications.

Table 6. Design dimensions for UI in 3D environments.

	Fixture	**Association**	**Level of aggregation**	**Manipulation directness**	**Persistence**	**Visual encoding**
Elements are	Screen-fixed	Spatially explicit	Minimal number of variables per display	None (static)	Rarely visible	Geometric, linguistic, pictorial
	World-fixed	Visually implicit	Various combinations on various displays	Indirect (player interacts with system)	Sometimes visible	
	Object-fixed		One complex display with all variables	Direct (player interacts with UI element)	Always visible	
	User-fixed					

Degree of fixture

Most commonly in 3D applications, important UI elements are attached to a location on the x-y plane of the 2D screen, which is referred to as screen-fixed UI. This type of UI has traditionally been used for the persistent display of essential stats about the player's health, supplies, and other pertinent data. Figure 12 shows screenshots from Doom (111) and the Outer Worlds (83), released 26 years apart, but both using this same UI modality. Of course, screen-fixed UI is the standard for almost all WIMP interfaces, and thus for most desktop software employed in today's world, whether in Microsoft Word, Adobe Photoshop, or Tableau.

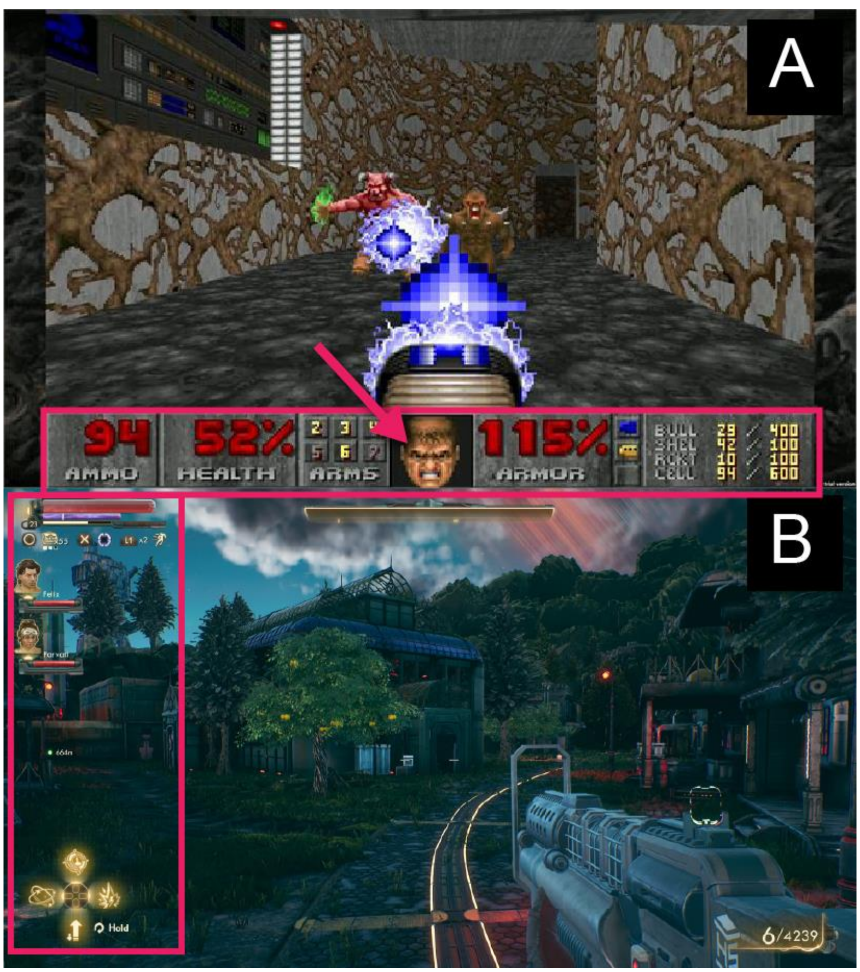

*Figure 12. **A**: Doom, a 1993 game, with its screen-fixed UI at the bottom, displaying data about the player's remaining supplies and health. Note the use of a Chernov-style face to convey additional information about the current health status. **B**: The Outer Worlds, a 2019 game, still using screen-fixed UI for persistent display of health, supplies, other players, etc.*

Screen-fixed UI is not limited to being 2D. Outer Wilds (142) presents a variety of supply and navigation-related data for three dimensions, see Figure 13.

Figure 13. Supply and navigational information in Outer Wilds, a 2019 game.

In VR, however, there is usually no fixed screen space to attach those UI elements to; as a consequence, such content needs to be either *world-fixed*, *object-fixed*, or *user-fixed*. *World-fixed UI* elements are tethered to a set point within the 3D world, and need to be traveled to in order to be used. A great example are location markers such as the blue navigational aids in Far Cry 5 (199), see Figure 14. When the player enters a destination into the in-game map system, this overlay is created in order to guide the player to said location by car. This leverages player's real-world experience with car navigation systems.

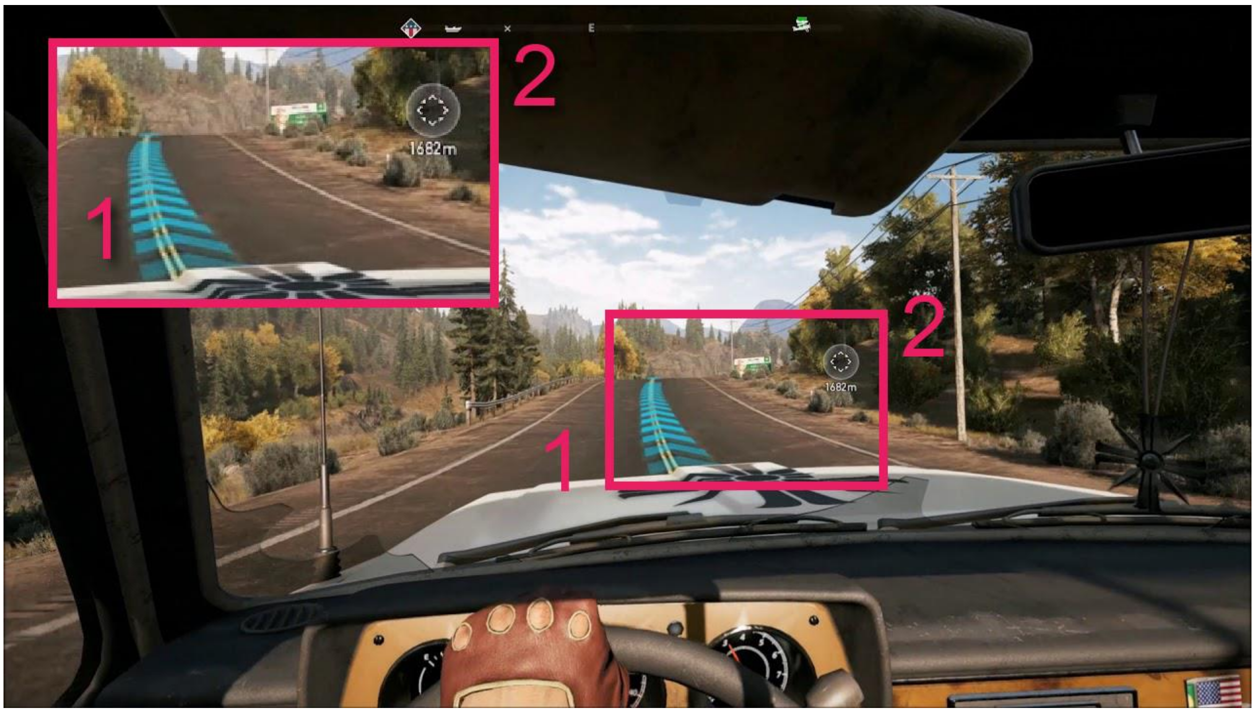

Figure 14. An example of world-fixed UI in Far Cry 5. Notice the blue arrows laid over the road in front of the car (#1), acting as waypoints for the user driving the car to the destination as indicated by another piece of world-fixed UI (#2), indicating the distance to the destination.

World-fixed UI is extremely prevalent in all forms of 3D experiences, especially when it comes to navigation. The ability to see immovable objects in a virtual space marked without occlusion allows the user to determine the validity of their overall course. World-fixed UI is a standard practice not only in video games but also in more real-world scenarios. Google Maps (10), for example, has a novel Live View feature where the user can use the video stream of their smartphone camera to get an AR experience. While the app does not add virtual waypoints to the street (opting for a separate view instead, see Figure 15), the parallels to the navigation approach in Far Cry 5 become immediately clear.

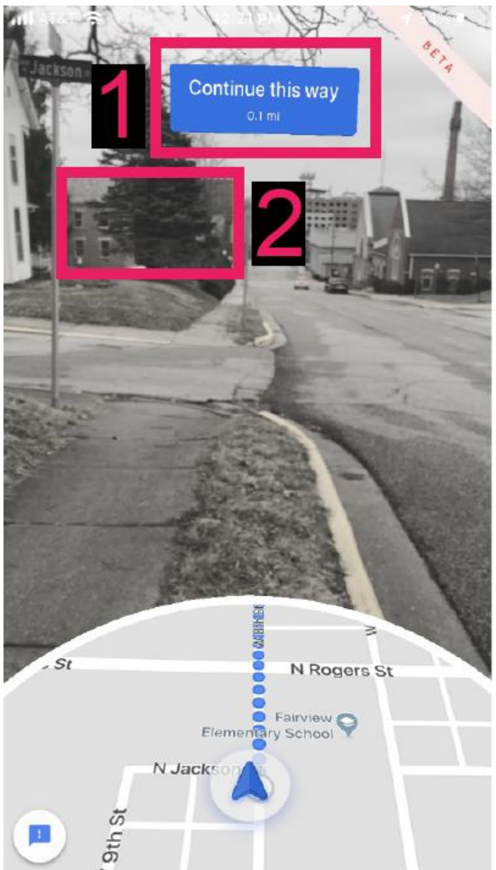

Figure 15. Google Maps Live View feature. #1: world-fixed text panel with distance indicator, attached to destination (tall building behind tree, #2). A separate panel at the bottom shows a map with the user's current location and next waypoints.

While we define world-fixed UI as interface elements that are attached to immovable objects (and thus as having a fixed position), o*bject-fixed UI* is attached to an object in 3D space and can travel with it. In video games, this is often used to mark non-playable characters (NPCs), such as allies and enemies in combat situations. In Rise of the Tomb Raider (64), the player can perform stealth actions to remain undetected when sneaking up on enemies. A skull icon on top of an NPC indicates the availability of such an action, see Figure 16, top. Another example of object-fixed UI comes from Life is Strange (77), where the main gameplay consist of the player moving through a high school and interacting with a variety of NPC and objects, see Figure 16, bottom. Unlike the NPCs in Rise of the Tomb Raider, these objects do not move.

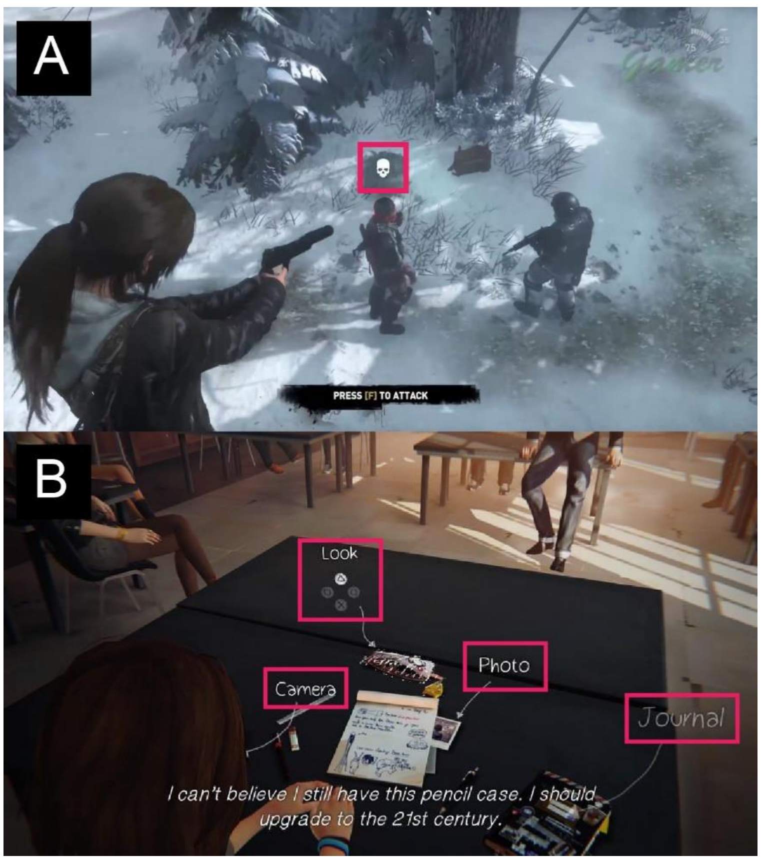

*Figure 16. **A**: a skull icon on top of an NPC indicates a potential action for the player to take. When the NPC moves, the icon moves with it. **B**: object-fixed UI in the video game* Life is Strange (77)*. Note how UI highlights potential interactions with objects.*

Object-fixed UI is contextual and needs to be updated according to the position and gaze direction of the user as well as the current game state to determine possible interactions at that particular time. The visibility and interactivity of object-fixed UI thus depends on a variety of factors, and anticipating the user's need for non-persistent UI it is a challenging task.

In VR, object-fixed UI becomes especially important when using VR controllers. Commonly, developers and designers can let the user compose complex interactions via controller-attached menus. In Tilt Brush (97), see Figure 17, users draw and paint in 3D, using static or animated materials and a variety of colors, stencils, and brushes. In order to allow the user to combine ingredients, the developers designed a multidimensional menu system where one VR controller allows the user to open and navigate between menus while the other controller lets the user make selections. These menus are attached to the 3D representation of the VR controllers and thus qualifies as object-fixed UI.

Figure 17. Another example of object-fixed UI: a menu rendered on top of a VR controller representation in Tilt Brush.

Finally, user-fixed UI describes elements that move with the player. Depending on the display system, user-fixed UI may be equivalent to screen-fixed UI; for instance, this is the case with first-person games and experiences as can be seen in Figure 12. In the presence of a 3D avatar, such as in Dead Space (80), see Figure 18, user-fixed UI can be attached to the player's representation of themselves. The UI is embodied into the player such that health and stamina are displayed not as abstract entities fixed to the screen but as part of the avatar's space suit. Similarly, the remaining supplies and target reticle, relevant for defense against attackers, is presented as natural part of the device the avatar is using, not as an overlay visible only to the player.

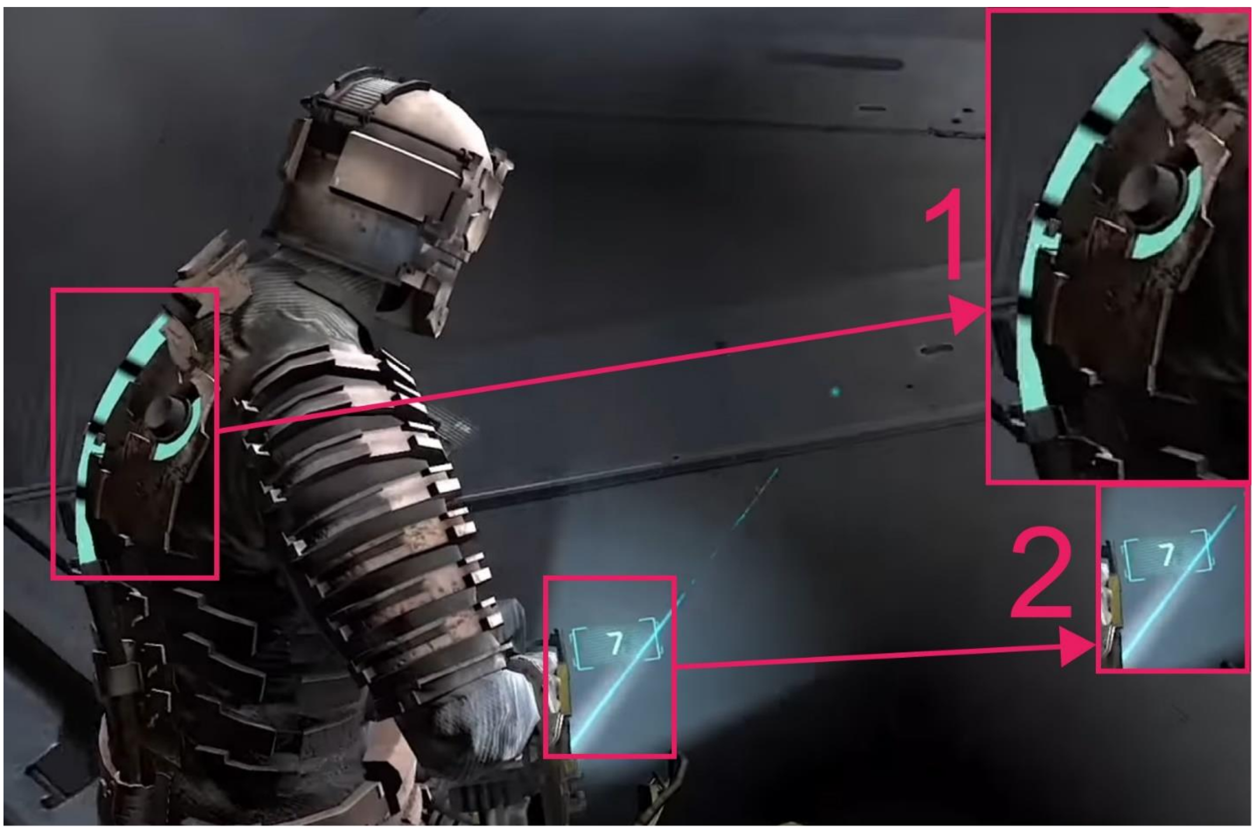

Figure 18. Dead Space with its minimalistic UI. #1: health and stamina. #2: remaining ammunition and targeting reticle

Association

Relationships between UI elements and virtual content, which Bowman et al. (38) refer to as "perceptual information" (p. 83), can be spatially explicit, meaning that the UI element and the 3D object it relates to are in physical proximity (within the context of the virtual environment). The various action-focused overlays in Figure 16 are a good example, as are most object-fixed UI elements. Visually implicit associations, on the other hand, comprise cases where the UI element is removed from the object it refers to, but when selecting and interacting with one element (the object or the UI element), the other is highlighted as well. The bar graph in Figure 21 below is a good example; when the player places charges on a virtual asteroid (not pictured).the bar graph is updated to aid the player in finding the optimum charge level, keeping this vital visualization in one place.

Level of aggregation

Virtual environments can feature individual variables across multiple displays, combination of variables across displays, or all variables contained in one complex visualization. While we could not find many examples to illustrate the differences between these, the settlement information display from Civilization VI (88) in Figure 19 exemplifies how UI elements with dense information displays can be used by the player to manage complex actions in a strategy game by condensing bigger amounts of data.

Figure 19. Settlement display in Civilization VI.

Degree of manipulation directness

The degree of manipulation directness indicates whether a UI element is susceptible to the user's direct input or whether it is updated through the system based on the player's actions. Some 3D UI elements are static in that they cannot be changed through player input at all (static). This comprises all interface elements that do not reflect the user's current state within the 3D experience. Common examples from video games are splash screens when the game is started, version numbers, and company logos. Completely static UI is quite rare and normally used to provide persistent information that is unlikely to change over the course of the experience. Most UI elements allow for at least some interactivity. Object-fixed UI, for example, can be moved through player interaction. Screen-fixed UI can be updated. Another common form of interactive UI are health and status bars, see Figure 12, label #1. While these cannot be clicked, dragged, or otherwise interacted with directly, they are still affected by user actions (indirect manipulation).

The most common fully interactive UI elements are buttons, lists, and sliders, as can commonly be found in the options section of video games, see Figure 20.

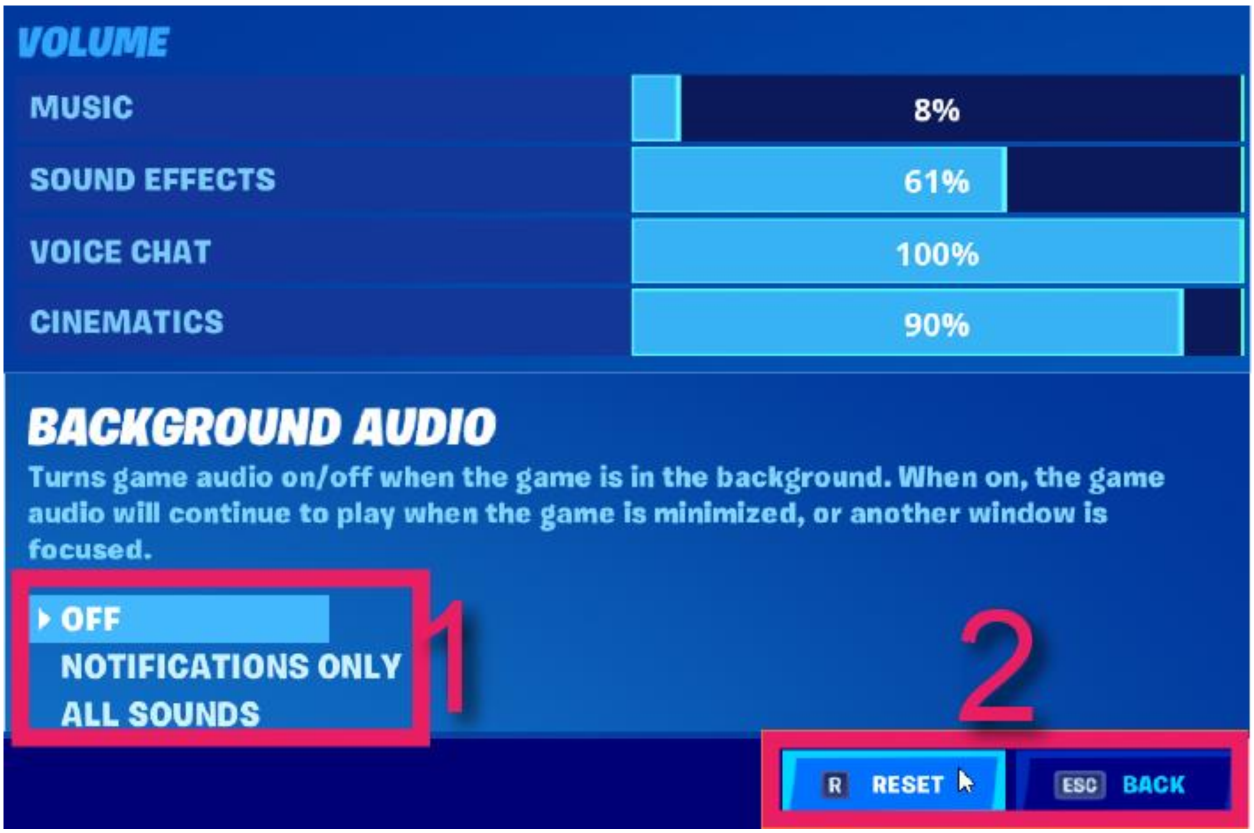

Figure 20. Options menu in Fortnite (84), such as sliders for volume control and a list of options for a setting (#1) with two buttons (#2), active (left) and inactive (right).

Degree of persistence

UI elements are displayed either very briefly, for extended periods of time, or constantly. An example of the latter are the health and progress bars as shown in Figure 12, or the embodied UI in Figure 18. Elements with limited persistence are usually displayed on-demand according to program logic, and are thus contextual. The object-fixed UI in Figure 16 is a great example; it only makes sense for these texts and icons to be shown when action can be taken by the player. The increasing prevalence of non-persistent UI elements, especially in 3D environments, allows for the development of the minimalistic UI encountered in that same figure.

Degree of visual encoding

Lastly, UI in 3D games is composed out of a combination of graphic symbols: geometric, linguistic, and pictorial. Since UI elements often encode game data in a way that is intuitively understandable for players, data visualization principles have a great influence on the design decisions for UI elements. The volume sliders in Figure 20 are based on 100% stacked bar charts, as are the health bars in Figure 12 and Figure 18. Figure 21 shows an example from the space exploration game Elite Dangerous (96): In order to mine valuable materials, the player can put seismic charges into asteroids. However, the player needs to carefully balance the amount of low, medium, and heavy charges to achieve an optimum yield balance. To communicate this, the developers implemented a bar chart with color-coding that visualizes state changes between charges (labels #1 and #2). Additionally, linguistic (#3) and pictorial (#4) symbols are used to communicate the current state of the player's actions.

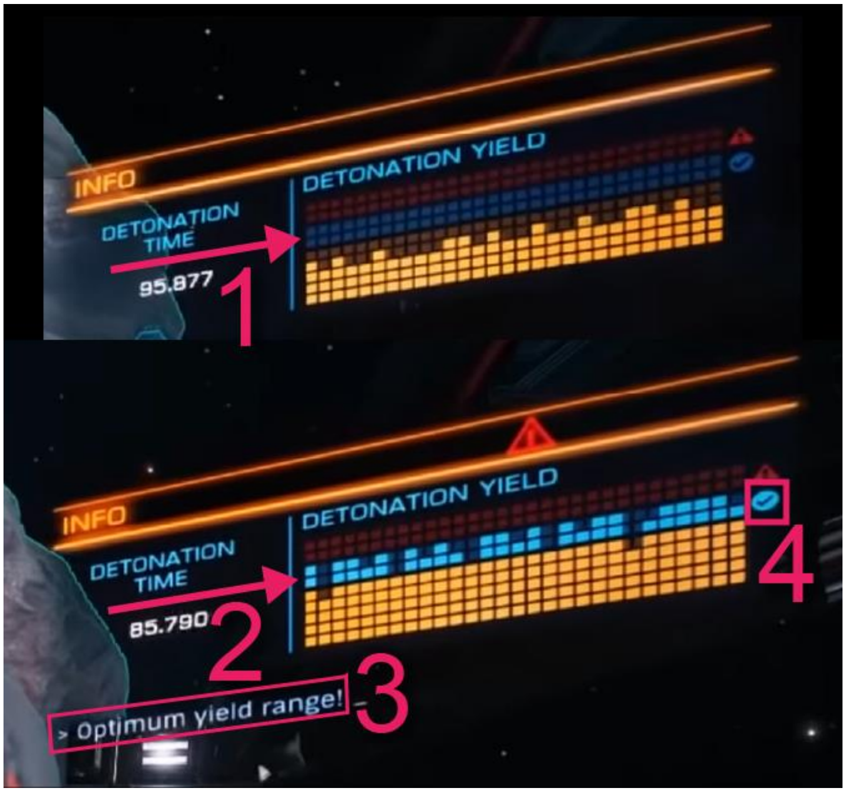

Figure 21. A bar graph to inform the player about the current charge level.

3. Research Questions

Each user study in this dissertation has its own set of research questions. For convenience, they are listed together below. Hypotheses can be found in the corresponding sections for each study. We compiled an overview of hypotheses and confirmed/rejected status in Table 14, Table 15, and Table 16.

User study 1 (Comparing Time, Accuracy, Satisfaction in VR vs. Desktop Registration User Interface (RUI), chapter 5):

- **RQ1**: What position and rotation accuracy and completion time can be achieved with the three different RUI setups?
- **RQ2**: What are the error and bias, i.e. the deviations for each axis as well as the cumulative deviation (see section Bias and error), for position accuracy in all three dimensions?
- **RQ3**: How does task complexity (e.g., smaller tissue block size or more rotation, larger distance between tissue block and target) impact accuracy and completion time?
- **RQ4**: What is the maximum performance level that a user can reasonably achieve, and how many practice tasks are required before performance levels out? That is, after how many tasks do users reach a plateau when accuracy or completion time do not significantly change anymore.
- **RQ5**: What is the tradeoff between accuracy and completion time? For example, if users are asked to register fast, does accuracy decrease? If users are asked to register accurately, do they need a longer time to complete the task?

- **RQ6**: How satisfied are users with the results achieved in the three different setups?

User study 2 (Using VR Data Visualizations to Improve Time, Accuracy, Satisfaction in VR vs. Desktop RUI, chapter 6):

- **RQ1**: When presented with a VR visualization of their own movement data from the **Ramp-Up** phase during their **Reflective phase**, do users in the **experiment** cohort have a better performance in the following **Plateau phase**, measured in accuracy and completion time, compared to the **control** cohort who does the Plateau phase **without** a **Reflective phase**?
- **RQ2**: Between-subject and within-condition, across cohorts, are there significantly different usage patterns between the experiment and control cohort during the Plateau phase?
- **RQ3**: In the Reflective phase, which interactive tool do users most often apply (identified by logging all user inputs and state of the interactivity tools)? How many times do they change the time slider (see Figure 43 below)? How often do they turn the kidney (base map) on and off (see Figure 40 below)?
- **RQ4**: Within the experiment group, in the Reflective phase, are metrics on user actions and interactive tool usage, measured through telemetry, correlated with higher performance in the Plateau phase?
- **RQ5:** In the mid-questionnaire between the intro and main part of the Reflective phase, what is the relationship between the task score for this questionnaire and the performance in the Plateau phase for 2D Desktop users?

User study 3(Improving Completion Time, Memory, and Satisfaction for Traversing Virtual Buildings Using VR Data Visualizations, chapter 7)

- **RQ1**: Is there a difference in completion time between the control and experiment cohorts during trial 2?
- **RQ2**: Is there a difference in the rate of change in completion time from trial 1 to trial 2 between the two cohorts? That is, when computing the differences in completion time per trial and per subject, and then compare these values between the cohorts, is there a significant difference?
- **RQ3**: When asked questions about the tasks and the virtual building after taking a break (control) and completing their Reflective phase (experiment), is there a difference in score between the two cohorts?
- **RQ4**: What are the preferred choices of navigation methods during the last round of tasks?
- **RQ5**: Is there a difference in self-reported satisfaction between the two cohorts at the end of the experiment?

4. Methods

In Table 7, we provide a summary of the three user studies in this dissertation by number of subjects, cohorts, setups, and interaction types tested.

Table 7. Overview of studies.

	Study 1	**Study 2**	**Study 3**
Title	Comparing Time, Accuracy, Satisfaction in VR vs. Desktop RUI	Using VR Data Visualizations to Improve Time, Accuracy, Satisfaction in VR vs. Desktop RUI	Improving Completion Time, Memory, and Satisfaction for Traversing Virtual Buildings Using VR Data Visualizations
Number of subjects	42	84 (42 + 42 from Study 1)	68
Cohorts (control and experiment)	1	2	2
Setups	3 (2D Desktop, VR Tabletop, VR Standup)	3 (2D Desktop, VR Tabletop, VR Standup)	1 (VR)
Interaction types		Filter, navigate, animate/replay	Filter, navigate, link and brush

User study 1, while not containing data visualizations of interaction types, served as control for study 2. Further, this study was motivated by and provides research evidence on how to best solve a practical problem by applying virtual reality to gross-anatomical tissue registration. Finally, it represents a major research and development effort tied to an international scientific endeavor and thus constitutes a vital part of our dissertation research.

Studies 2 and 3 had a control and an experiment cohort. The control cohort performed their tasks without any interventions. As mentioned above, the subjects in user study 1 (see chapter 5) functioned as control cohort for the subjects in user study 2 (see chapter 6). The procedures for each user study are illustrated individually in Figure 25, Figure 31, and Figure 56.

5. User Study 1: Comparing Time, Accuracy, Satisfaction in 2D Desktop vs. VR RUI

5.1 Introduction

The human body consists of trillions of cells. Understanding what cells exist in which anatomical structures and spatial contexts is essential for developing novel approaches to curing diseases. HuBMAP is a research effort carried out by hundreds of researchers in several dozen institutions in the U.S. and abroad (184). The goal of the multi-year project is to create a reference atlas of the healthy human body at single-cell resolution, capturing spatial information about cells and tissues in unprecedented detail. In order to facilitate HuBMAP's ambitious mission, different tools are being developed. This paper presents a novel 3D object manipulation user interface, called the Registration User Interface or (RUI), developed to support tissue registration performed by tissue mapping centers (TMCs), as well as Transformative Technology Development (TTD) and Rapid Technology Implementation (RTI) teams. In the remainder of this section, we review typical approaches to registering tissue data together with registration accuracy typically achieved. We then derive a list of requirements for a qualitatively novel approach to tissue registration and discuss research questions and hypotheses.

5.1.1 Tissue registration procedure and prior work

Developing a human reference atlas at single-cell resolution requires recording the size, position, and rotation of tissue extracted from living or post-mortem patients—before the tissue is processed for spatially explicit analysis. Figure 22A shows a photo of a typical setup: a kidney was butterflied and placed on a dissecting board to capture its size and shape, as well as the size, position, and rotation of a tissue block (outlined in pink) extracted from it. Commonly, a computer is close to the dissection work area so data can be entered and uploaded.

The documentation of extraction sites is non-trivial as different donors might have organs of different sizes and the number and shape of anatomical structures (e.g., the number of renal pyramids per kidney) might differ across individuals. It is common to use exemplary organs derived from an individual donor's data as a reference. An example is the male left kidney derived from the Visible Human (VH) dataset (3, 187) published by the National Library of Medicine (NLM). This 3D model is about 100 mm high (see green y-axis), 60 mm wide (red x-axis), and 40 mm deep (blue z-axis)—see Figure 22, B and C.

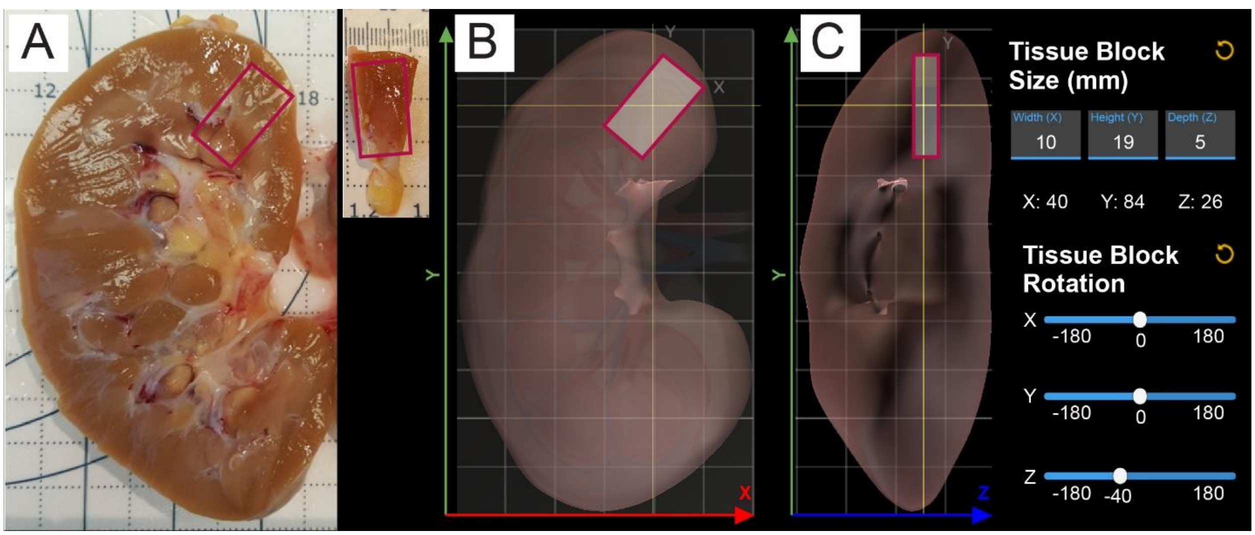

*Figure 22. Physical vs. virtual tissue registration. **A**: Bisected kidney on a dissecting board. Pink outlines indicate where the tissue block highlighted pink (shown in top right) will be extracted. **B**: RUI with reference kidney of about the same size in x-y view. **C**: RUI in z-y view with user interface that supports entry of tissue block size in mm, review of x, y, z position values, and change of tissue block rotation in 3D.*

Shown in Figure 22C is the interface used for entering 'Tissue Block Size' in mm and for rotating the tissue block (via sliders). Position can be adjusted by dragging the tissue block to the correct location within the 3D reference organ. An x-y view and a y-z view can be selected to check and correct the tissue position. The x, y, and z positions are also displayed to the user. Size, position, and rotation values can be reset by clicking on the corresponding circular arrow buttons.

Different procedures exist to capture relevant information; resulting data is submitted to diverse clinical record-keeping systems with different metadata schemas. Different organs—e.g., lung (156), breast (143), thymus (71), and pancreas (4)—have rather different needs and are subject to many standard operating procedures (SOPs) and checklists (15, 133). A closer look at these SOPs and checklists developed for different organs by different authors reveals the lack of common procedures and documentation standards. More importantly for HuBMAP, existing data captures only **partial or inconsistent spatial information** (i.e., the level of detail at which this information is captured varies across protocols).

Partial and inconsistent spatial information

Pathologists and other wet-bench workers typically use SOPs—written as protocols (14, 219) and published on protocols.io—to ensure reproducibility, establish relevant terminology, and share otherwise disparate materials and instructions in a consistent framework. Most importantly for our research, they use SOPs to capture specific workflows such as extracting tissue blocks from organs (41), tissue preservation through freezing (8, 219), or preparing specimens for further analysis (9). Some of these protocols require the lab worker to capture the spatial origin of tissue in reference to an organ and/or its dimensions: e.g., some SOPs involve pictures of dissected organs or tissue blocks on dissecting boards with markings for length and diameter units (8, 41, 115) (see Figure 22A), occasionally at different stages of the dissection process (114). The scale of the marker positions and the reported data is in the millimeter range. When using dissecting boards with markings is not feasible, some protocols supply abstract illustrations to show the extraction sites of tissue (98). The quality and purpose of these pictures vary; many are ad hoc, with inconsistent lighting and varying quality, or no pictures at all (165). In some cases, the authors provide no exemplary pictures but give a verbal description of how the donor organ has to be aligned and dissected (164). This causes many of the existing protocols to capture only **partial and inconsistent spatial information**. Manual annotations in these pictures offer a small amount of orientation with regard to the spatial provenance of the tissue block, but this kind of documentation lacks detail and reproducibility across teams and organs. Further, inferring the correct dimensions of a tissue block from a photo can be challenging, depending on the distance between the side of the tissue facing the viewer and the cutting board.

Limited computability of photos

A second issue with the current record-keeping practices for spatial origins of tissue blocks is that **images** of extracted tissue and/or organs, if present at all, **are not computable**. To be of value for the HuBMAP atlas, tissue spatial data must be provided in a format that is uniform across organs and can be used to correctly determine the size, position, and rotation of tissue blocks in relation to a 3D reference body. Images with spatial annotations do not support this, and advanced techniques such as computer vision algorithms cannot be trained and used due to the quality and limited quantity of existing images. While photos provide an efficient way of archiving general spatial information in the context of individual labs, they do not provide the precision and standardization required for reference atlas design. To overcome this limitation, we implemented an online service that lets subject matter experts (SMEs) size and register 3D tissue blocks within 3D reference organs to generate unified data across tissue types.

Challenges of 3D manipulation

The Registration User Interface (RUI) was developed to **address practical concerns** regarding tissue registration. However, there are several known challenges when manipulating 3D objects in 3D. We assume the SME is an able-bodied individual with two hands and a basic understanding of how to use photography equipment and a paper or digital documentation sheet. The SME places the tissue on the dissecting board, aligns it with the provided grid system, takes photos, and writes down annotations.

In the proposed RUI, there are various cognitive challenges as 3D manipulation is non-trivial. In our review of prior work, we focus on two methods to enable a user to manipulate a virtual object in 3D space: widgets and extended input devices. The de-facto standard in many 3D modeling applications is the use of a mouse and color-coded virtual **widgets** attached to the object, as discussed in Maya (17) and Blender (30). These widgets allow the user to perform position, rotation, and scaling operations. Schmidt, Singh and Balakrishnan (168) proposed a user-input-based extension to the traditional widget system; but note that even the most experienced participants in the evaluation study needed twice as long to complete the assembly tasks when using their system than when using the traditional version. A similar framework was proposed in 1995 by Bukowski and Séquin (46), who prototyped their interaction language for “pseudo-physical behavior” from the user, where the 2D motion of the mouse cursor is extrapolated into 3D motion in the virtual environment.

In addition to using widgets to perform 3D manipulation, there exists a body of literature about **input devices extending the standard computer mouse** for these tasks. Balakrishnan et al. (22) developed the “Rockin’Mouse”, a 4-DoF device that allowed users to control position and rotation without having to switch between modes. Their pilot study found that users were able to complete a set of block-matching tasks 30% faster with a Rockin’Mouse than with a regular mouse. In order to explore the design space of a multi-touch mouse, Villar et al. (209) presented a series of five prototypes using different touch input layouts. They found that ergonomics and form-factor were important design aspects for user satisfaction, although their study was aimed at gathering qualitative results rather than quantitative performance measures. For their GlobeMouse and GlobeFish, Froehlich et al. (95) separated position and rotation manipulation using a trackball (rotation) connected to an inner and outer frame (position). When tested against a commercial option in a study, they found that the completion times for their devices were significantly faster than for the commercial SpaceMouse, although they found a similarly strong learning effect for the three devices tested over the course of two sets of four tasks per device (with training sessions before each task). A commercial approach to the extended mouse is the aforementioned SpaceMouse (1), a six-DoF device that lets a user position and rotate a 3D object along and around all three axes at the same time using a self-resetting internal mechanism when no user input is given. A major issue for this advanced, modified hardware is the steep learning curve, and the fact that these are prototypes makes their widespread deployment and adoption not feasible.

While the aforementioned projects feature variations on the widely used mouse input device, recent efforts have focused on alternative input devices. Soh et al. (185) developed a simple hand gesture interface for the Microsoft Kinect to enable translation of rotation of 3D objects. Similarly, Lee et al. (125) used a webcam and a projector to transform a piece of cardboard into a movable, handheld 3D device that lets users rotate the projected 3D object. In a more recent paper, Mendes et al. (138) used the HTC Vive and Unity to design a system for custom translation and manipulation axes (MAiOR). In a user study comparing MAiOR to a regular six-DoF approach without separation of manipulation and rotation as well as a system with virtual widgets, they found that the approach with traditional widgets achieved the highest overall success rate but came at the cost of higher completion times with increasing task difficulty and confirmed that mid-air manipulations with VR controllers lack precision.

Overcoming the challenges of 3D manipulation

Building on and extending this prior work, the HuBMAP RUI aims to support scalable and computable tissue registration and data management. It lets experts use a nearby computer to digitally capture the size, position, and rotation of tissue blocks in relation to a reference organ, together with important metadata such as name, tissue ID, date, and time.

This paper presents the results of user studies that aim to determine and compare registration accuracy and speed for different user interfaces. Specifically, we compare the 2D desktop setup with two VR setups—using a sitting and a standing setup. All three user interfaces support the general registration task shown in Figure 23 and detailed in Section 5.2.1. In all three setups, the subject sees a reference organ kidney with a virtual purple target block on the left and a white tissue block on the right that needs to be matched in position and rotation with the target block. In all cases, the sizes of the tissue block and the target block are identical, which resembles the real-world scenario in which a tissue block has just been extracted from an organ.

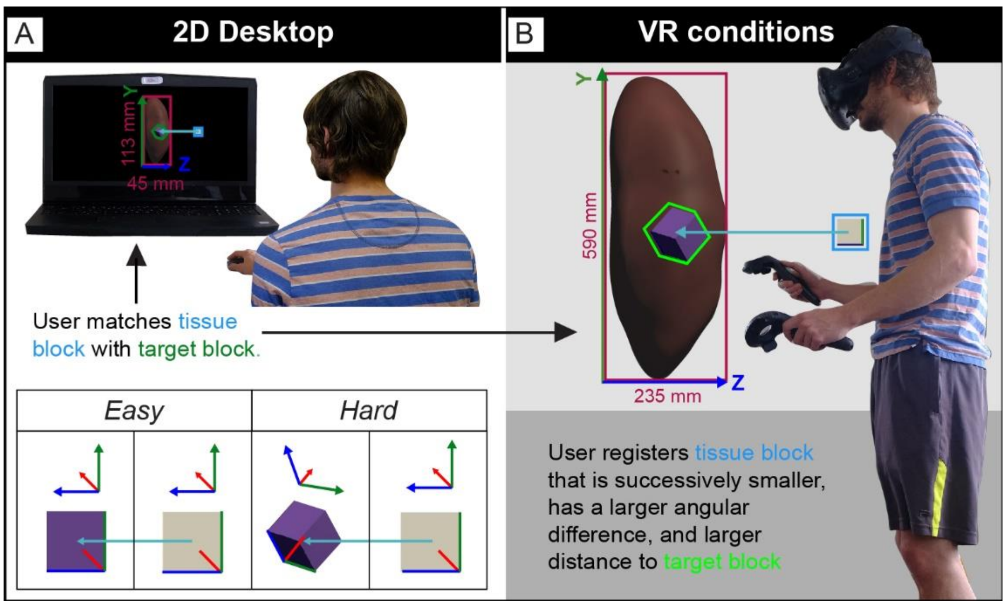

*Figure 23. The task setup in our user study. Reference organ with target block indicated (purple) and tissue block (white) to be registered into the target block. The light blue arrow indicates block centroid (mid-point) distance. Task difficulty increases as the tissue blocks get smaller, block rotation increases, and distance between the blocks increases. **A**: 2D Desktop setup. **B**: The two VR setups.*

The reference organ in our study appears in different sizes in the 2D Desktop setup and the two VR setups (see Figure 23). The kidney is 113 mm tall on the screen in the 2D Desktop condition and 590 mm tall in VR. We chose these different sizes to make use of the ability in VR to interact with objects that would not normally fit on a regular laptop screen. We elaborate on this (and its implications for data analysis) when presenting the results in Section 5.3. Three color-coded coordinate axes (implemented as long, thin cylinders running along the edges of each cube, conjoint at one of its corners) are used to indicate tissue and target block rotation. They are colored red, green, and blue for x-, y-, and z-axis, respectively. As we present the users with increasingly difficult registration tasks (see Figure 25), subjects must adjust not only the position but also the rotation of the tissue block to match the position and rotation of the target block. The focus on these adjustments mirrors the real-world need for these 3D manipulations to capture the spatial provenance of tissue blocks with regard to a reference organ in the RUI.

5.1.2 Requirements for registration user interface

Informal interviews and registration tests with revealed various requirements for the RUI. Requirements can be grouped into five categories and are discussed subsequently.

1. **Metadata Entry**: The RUI must support entry of data such as user name, organ name, tissue block size, and date and time of registration. This metadata must then be sent to a database for ingestion and usage in the HuBMAP data infrastructure and portal.
2. **Accuracy**: The RUI must support gross-anatomical-tissue registration at about **one mm** for position and about **20 degrees** for rotation accuracy.
3. **Training and Completion time**: The RUI should not require more than **five minutes** to learn, and each tissue registration should not take more than **one minute** to complete.
4. **Satisfaction**: The RUI should be easy to use, and subjects should feel a sense of accomplishment after they perform the registration task.
5. **Deployment**: The RUI should be usable on a computer in a lab, ideally right after tissue has been extracted. A typical lab computer uses a Windows or Mac operating system and runs Chrome, Firefox, or other web browsers. A typical window size is 1920 x 1080 (full HD) or 3840 x 2160 (4K) pixels at 72 DPI.

5.1.3 Research questions and hypotheses

Given the overall domain task and the requirements stated in Section 5.1.2, there are six research questions (RQ) this study aims to answer. We also present associated hypotheses (H) here:

RQ1: What position and rotation accuracy and completion time can be achieved with the three different RUI setups?

H1a: Users in VR Tabletop and VR Standup achieve significantly higher position accuracy than users in 2D Desktop.

H1b: Users in VR Tabletop and VR Standup achieve significantly higher rotation accuracy than users in 2D Desktop.

H1c: Users in VR Tabletop and VR Standup have significantly lower completion times than users in 2D Desktop.

RQ2: What are the error and bias, i.e. the deviations for each axis as well as the cumulative deviation for position accuracy (see Bias and error) in all three dimensions?

H2a: We do not expect any major bias for any setup in any dimension.
H2b: We expect the error to be greatest for the 2D Desktop setup due to its restricted input devices and limited viewing positions.

RQ3: How does task complexity (e.g., smaller tissue block size or more rotation, larger distance between tissue block and target) impact accuracy and completion time?

H3a: More complex tasks lead to lower position accuracy for all setups.

H3b: More complex tasks lead to lower rotation accuracy for all setups.

H3c: More complex tasks lead to higher completion times for all setups.

RQ4: What is the maximum performance level that a user can reasonably achieve, and how many practice tasks are required before performance levels out? That is, after how many tasks do users reach a plateau when accuracy or completion time do not significantly change anymore.

H4: VR users need a lower number of tasks to plateau than 2D Desktop users.

RQ5: What is the tradeoff between accuracy and completion time? For example, if users are asked to register fast, does accuracy decrease? If users are asked to register accurately, do they need a longer time to complete the task?

H5: In all setups, the more time users spend on a task, the higher position and rotation accuracy they achieve.

RQ6: How satisfied are users with the results achieved in the three different setups?

H6a: Users in both VR setups are more satisfied with their performance than 2D Desktop users.

H6b: There is no significant difference in user satisfaction between VR Standup and VR Tabletop users.

The paper is organized as follows: In Section 5.2, we introduce methods, including study design, task difficulty, and performance metrics. In Section 5.3, we present the qualitative and quantitative results of this study before interpreting the results with regard to the requirements from Section 5.1.2 and the research questions and hypotheses stated in Section 5.1.3. In Sections 5.4 and 5.5, we discuss results and present an outlook on planned future work.

5.2 Materials and methods

This section presents the overall study design, the three different hardware/software setups, task difficulty metrics and synthetic tasks generation, as well as human performance metrics, plateau, and a satisfaction score computation. A power analysis was conducted prior to running the experiment to determine the number of subjects required to achieve significant results.

5.2.1 Study design

We used a typical user study design featuring a study information sheet (SIS) in the beginning to get user consent, followed by a pre-questionnaire, tutorial and experiment tasks, and a post-questionnaire. All 42 subjects were run in person by one of the authors of this paper.

The main part of the experiment asked subjects to use one of three setups: 2D Desktop, VR Tabletop, or VR Standup. All three setups are shown in Figure 24, A-C. Users were randomly assigned to one out of these three setups. Different levels of task difficulty were used for the tutorial, Ramp-Up, and Plateau tasks (see Section 5.2.2). Tasks were identical for all users regardless of setups (performance metrics are detailed in Section 5.2.3). Users determined for themselves when a task was done and were provided the equivalent of a “Next” button in all three setups. More information can be found in videos showcasing all three setups from a user’s perspective on GitHub (https://github.com/cns-iu/rui-tissue-registration). The SIS, pre-, and post-questionnaire data was presented and gathered using an online Qualtrics form. The tutorial and experiment tasks used a setup implemented in the Unity game engine (200). The Qualtrics form, along with documentation of logged data formats and data analyses, can also be found on GitHub.

All subjects used the same Alienware 17 R4 laptop with a display diagonal of 439.42 mm (17.30 in), running Unity 2019.4 on Windows 10 with a secondary monitor attached for ease of configuring the individual steps of the experiment. The laptop had an Nvidia GTX 1070 with 32 GB RAM of memory. A 1080p webcam recorded audio and video. For the VR setups, we used a 2016 HTC Vive with two Vive controllers. We ran the application for all three setups straight out of Unity. Data was collected using a custom C# script, writing data to a CSV file at a frequency of 10 Hz every time the user pressed a button.

The research facilitator could observe the subjects’ viewpoint on the laptop display, which was recorded with a screen-capturing software. We conducted the study in a collaborative space in a public university building and took precautions to preserve our subjects’ safety. The usable space for VR Standup and VR Tabletop users was around 10 x 10 ft (3 x 3 m). 2D Desktop users sat at a 4 x 4 ft table (1.2 x 1.2 m).

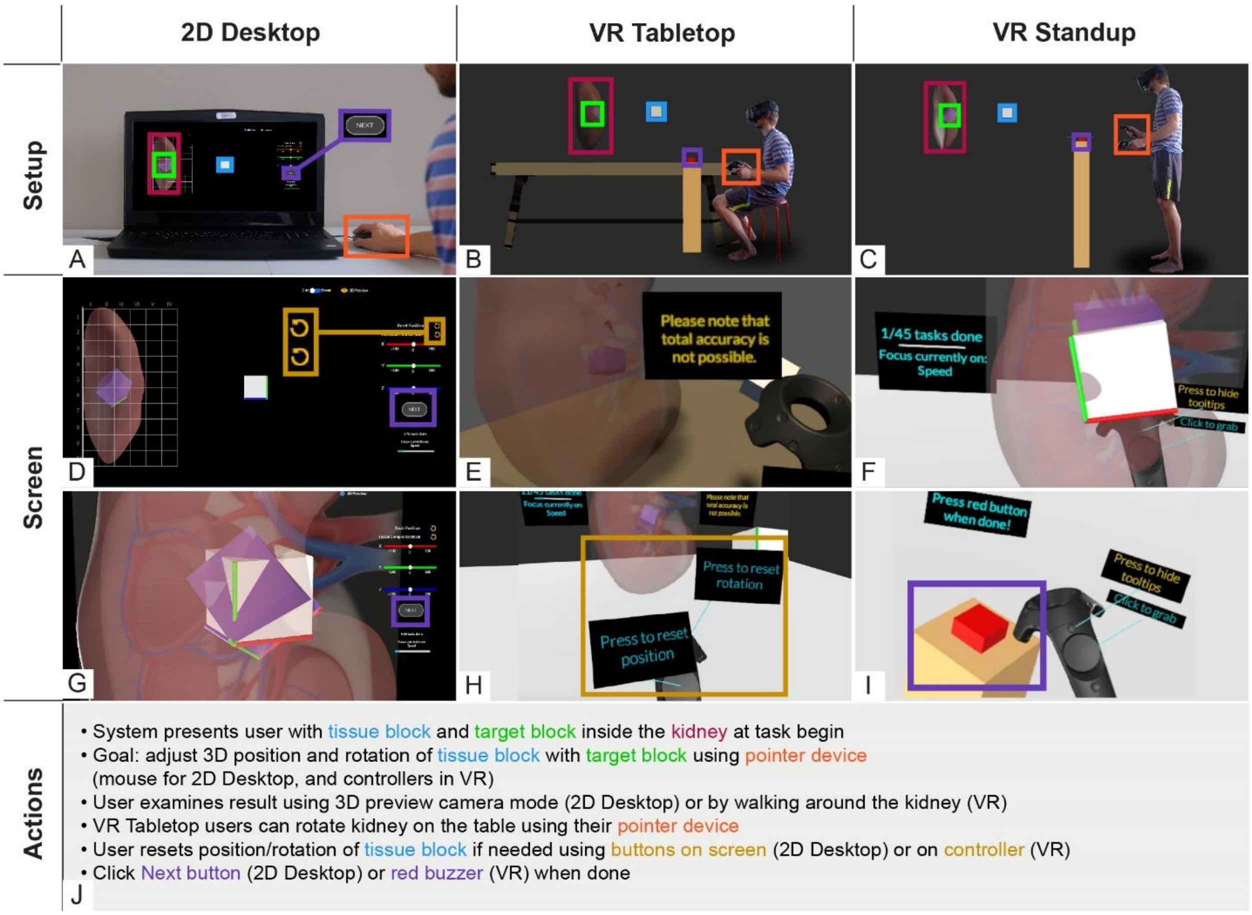

Figure 24. Setup, screen, and actions for 2D Desktop, VR Tabletop, and VR Standup. ***A-C****: Three RUI setups with a human subject.* ***D-I****: screenshots of the user interface.* ***J****: Required actions. The tissue block is outlined in blue, the target block in green, and the kidney—providing context and domain relevance—in pink. Tasks are submitted by selecting the purple NEXT/red button. The user could reset the position or rotation of the tissue block by pressing the corresponding yellow-brown virtual (2D Desktop) and physical buttons (VR).*

The three setups support nearly identical functionality as detailed subsequently.

The three setups

As described earlier, in both VR setups, the kidney was around 590 mm tall. In 2D Desktop, on the 1080p screen, the kidney appeared at a height of 113 mm. We chose these different sizes to make use of the ability in VR to display and interact with 3D objects in a much larger size than it would be possible on a standard laptop display. Figure 24, A-C, illustrates these differences with a human subject for scale.

2D Desktop

The screen in the 2D Desktop setup (see Figure 24A, D, G) consisted of a 3D work area covering most of the screen and featured a transparent model of a human kidney with an inlaid 2D image showing a schematic drawing of the vasculature with a 100 x 60 x 60 mm grid wrapping around the kidney. We added a progress bar on the bottom right with a text field indicating the number of completed tasks.

In terms of controls, the user could move the tissue block by clicking and dragging it with the left mouse button; they could also rotate it around each axis with the three sliders on the right side of the screen. Both position and rotation could be reset via button clicks.

The setup had two cameras the user could choose between: an orthographic camera (always aligned with either the x- or the z-axis) and a perspective ("preview") camera. The user could switch between the two by clicking the yellow eye icon on the center right, top edge of the screen. This preview camera could be rotated with relative freedom by clicking and dragging the left mouse button (see Figure 24G). The main camera, however, was less movable. The toggle switch in the top center allowed a movement of 90 degrees around the kidney's upward-facing axis, allowing the user to go back and forth between two predefined viewpoints with a smooth, animated transition. The usage of an orthographic camera is common in 3D modeling software as it makes the alignment of objects easier by taking away one dimension. The user could proceed to the next task by clicking a “Next” button.

VR Tabletop

In the VR Tabletop setup (see Figure 24B, E, H) the user was presented with a 3D model of a kidney and the same tissue block and target block as 2D Desktop users. As the name implies, the user sat at a table, both in the physical and virtual world. This allowed us to test whether simulating a physical work desk environment helped with the 3D alignment. The functionality provided through UI buttons and the mouse in the Desktop setup was implemented using a VR headset and VR controllers as pointer devices. The trigger button on the right hand allowed the user to grab and move the tissue block. Pressing the left touchpad or menu triggered a reset animation for the position or rotation of the tissue block, respectively. By default, tooltips were displayed atop the controllers, which could be turned off by pressing the right menu button. The user could rotate the kidney around its y-axis with the touchpad on their left hand and move the tissue slice with their right hand. The ability to turn the kidney was unique to VR Tabletop. The user could proceed to the next task by touching a virtual red buzzer on a stand at a height of around three ft (90 cm) above floor level.

VR Standup

This setup (Figure 24C, F, I) was similar to VR Tabletop; however, users stood in front of a reference kidney, able to explore it from 360 degrees while being assisted by the research facilitator for physical safety. The user could not rotate the kidney in this setup but was able to walk around the kidney to see it from different viewpoints.

Otherwise, the implementation was identical, with VR controllers as pointer devices and a virtual red buzzer to proceed to the next task.

Pre-questionnaire

After being welcomed to the experiment, subjects were asked to fill out a pre-questionnaire using an online survey software running on the same laptop we used for the actual tasks. This pre-questionnaire inquired about the subjects' prior experience with virtual reality and 3D video games and about their familiarity with different types of data visualizations such as graphs, charts, tables, maps, and networks. Additionally, we asked our subjects to disclose demographic information such as native language, job title, age, and gender. Items of specific interest for our user study were also whether users were right-handed or left-handed, their height, and whether they had a vision impairment. The complete questionnaire is available at https://github.com/cns-iu/rui-tissue-registration.

Tutorial task

After answering the pre-questionnaire, the subject was either presented with the experiment application in Unity (2D Desktop) or they donned the VR gear and got into position (VR Tabletop and VR Standup). They then listened to an approximately three-minute audio tutorial explaining the elements in the scene (kidney, tissue block, target block), what the tasks entailed, how to move and rotate the tissue block in a given setup, how to reset the position and location of the tissue block if needed, and how to submit task results and get a new task. The prerecorded audio ensured the same delivery of the content to all subjects. While the audio was playing, the subject was encouraged to practice moving and rotating the first tissue block, and to explore the

2D screen or 3D scene in front of them. We encouraged subjects in VR Tabletop and VR Standup to quit the experiment should they feel nauseous. The research facilitator monitored subjects at all times to ensure their physical safety. Having a facilitator present also helped many first-time VR subjects to correctly strap on the headset and move between VR and physical world without damage to the equipment or themselves.

Ramp-up tasks

Following the tutorial task, in the Ramp-Up phase, we asked subjects to solve 14 increasingly difficult tasks over time (see explanation of task difficulty in Section 5.2.2). As the task numbers increased, the size of the blocks to be placed became smaller and the rotational differences and distance between tissue and target block increased. After a pilot study with eight subjects, we decided to use 14 tasks to cover major difficulty levels while allowing every subject to finish the entire experiment in 60 minutes.

In the pilot study, some subjects spent unusually long times in the VR setups to achieve near perfect accuracy. To avoid this, we added three interventions: First, during the tutorial, we mentioned that 100% accuracy was not possible and asked subjects to use their best judgement when determining whether they were done with a task. Second, in the VR setups, we added a constantly visible text box next to the kidney with a reminder that 100% accuracy was not possible (see Figure 24E). Third, we gave subjects alternating audio prompts for odd tasks (focus on speed) and even tasks (focus on accuracy) in all three setups—this also let us explore the influence of task complexity on accuracy and completion time (RQ3, see Section 5.3.4), and tradeoffs in speed versus accuracy (RQ5, see Section 5.3.5).

Plateau tasks

Interested in understanding the number of tasks it takes before a user achieves their maximum performance in terms of accuracy and completion time, we asked subjects to register 30 blocks of identical size as fast as possible during the Plateau phase (RQ4, with results in Sections 5.3.2 and 5.3.3). The Plateau phase followed immediately after the last task of the Ramp-Up phase. We determined this number of tasks in pilot studies where we aimed to achieve a balance between subject exhaustion, total participation time, and detectability of a plateau. Given that subjects in the Desktop setup spent around three times longer on tasks as subjects in the VR setups (see Section 5.3.3), the number of tasks in the Plateau phase had to be high enough for us to detect a performance plateau while ensuring a timely finish of the experiment. A more complete description of Plateau phase task difficulty can be found in Section 5.2.2 and Figure 25.

Post-questionnaire

After finishing the registration tasks part of the study, the Unity application was closed, and each subject (now out of VR if part of VR Tabletop or VR Standup) completed a post-questionnaire about their experience. We included this post-assessment to learn how much users liked the registration interface, to determine what they would improve, and to compare user satisfaction across setups. Satisfaction score compilation is detailed in the section titled Satisfaction.

5.2.2 Task difficulty and stimuli generation

Task difficulty in the Ramp-Up phase of our study is a combination of the distance between tissue block and target block, the size of both blocks, and the angular difference between both blocks (see Figure 25). During the Ramp-Up phase, distance, angular difference, and size were continuously increased until the distance was 200% of the kidney height, the angular difference was 180 degrees, and the side length of the two blocks were only 5% of the kidney height. To generate stimuli used in the Ramp-Up phase, we used Lerp(), a native Unity method for linear interpolation (see Equation 1) to interpolate between start and end values. To increase angular difference over time, we used Slerp(), a different implementation of the aforementioned Lerp() function (for rotations), to gradually rotate the target block towards an end rotation of 0, 270, and 180 (around the x-, y-, and z-axis) using linear interpolation.

$$Lerp = \left\{ \begin{array}{ll} a, & , if\ t \leq 0 \\ b, & , if\ t \geq 1 \\ a + (b - a) * t & , if\ 0 < t < 1 \end{array} \right\}$$

Equation 1. Formula to compute task difficulty.

Distance and angular difference were smallest in the tutorial at 30% of the kidney height and 0 degrees rotational difference. Similarly, the length of each edge in both blocks was 20% of the kidney height, initially. Finally, in the Plateau phase, these values were consistent, with the distance and size values at around the same level of difficulty as the average Ramp-Up task (115% the kidney height for distance and 12.5% of the kidney height for size) but at maximum angular difference. Note that while we only show one Plateau task in Figure 25, there were 30 identical ones. Figure 25 also shows the tutorial (simplest), 14 increasingly difficult Ramp-Up tasks, and the 30 Plateau tasks (all of same task difficulty). Details on sizes, rotations, and distances used for the tasks are provided together with information on audio prompts.

	Distance	Angular difference	Size difference	Number of tasks	Prompt
Tutorial	*Smallest*		*Biggest*	1	Audio explanation of interface
	0.3 * kidneyHeight + offset	0 degrees + offset	0.2 * kidney Height		
Ramp Up	*Increasing*		*Decreasing*	14	Odd: speed Even: accuracy
	start: 0.3 * kidney Height + offset end: 2 * kidneyHeight	start: 0 degrees + offset end: 180 degrees	start: 0.2 * kidney Height + offset end: 0.05 * kidney Height		
Plateau	*Consistent*		*Consistent*	30	Speed
	1.15 * kidneyHeight	180 degrees	0.125 * kidneyHeight		

Figure 25. Task setup and levels of difficulty used in this study. The offset (computed via Equation 1) is a value that is added to gradually increase the distance and angular difference between the two blocks, and that is used to gradually decrease the size of the two blocks. Note that due to layout, only 13 out of the 14 Ramp-Up tasks are illustrated on the left.

The computation of Lerp() requires three values, where a is the start value (easiest), b is the end value (hardest), and t is an interpolation value between 0 and 1. For every task, t is computed by dividing the current task number by the total amount of Ramp-Up tasks (14). At task number 0 (i.e., the tutorial task), t evaluates to 0, which causes the function to return the start value. At task number 14 (the last Ramp-Up task), t evaluates to 1, prompting the function to return the end value. For any task in between, the function returns the start value with an offset value that increases over time. As input, Lerp() uses 3D vectors while Slerp() uses 3D rotations.

To ensure that task difficulty is identical across setups, we normalized each of the difficulty parameters (distance, rotational difference, size) by the height of the kidney in each condition. Figure 25 shows the tissue blocks and target blocks used throughout the experiment alongside the kidney for scale. The other columns indicate the different values for distance, angular difference, size, number of tasks, and audio prompts.

5.2.3 Performance metrics, plateau, and satisfaction score

To analyze survey and task data, we defined three performance metrics (position accuracy, rotation accuracy, and completion time) as well as a satisfaction score.

3D position accuracy

To answer RQ1 and RQ2, we needed to assess the position accuracy for each subject. Position accuracy equals the distance of the centroids of the tissue block and the target block, see light blue arrow in Figure 23. We compute the distance at run time using `Vector3.Distance()`, a static method in Unity that returns the distance between two points in 3D space. The position of both blocks and the centroid distance was collected at 10 Hz (i.e., 10 times each second).

To make use of the various possibilities for scaling in VR, the kidney was displayed in different heights across setups (but always with the same width-to-height-to-depth ratio). Measured from the lowest to the topmost vertex, the kidney in the two VR setups was 0.59 Unity scene units tall. In VR, scene units correspond to physical meters, so the kidney appeared at a height of 590 mm. Similarly, in the 2D Desktop setup, the kidney appeared at a height of 113 mm on the laptop display (see Figure 23). In order to compare position accuracy results between 2D Desktop and the VR setups, we normalized these values by dividing them by the height in which the kidney appeared to the user. When discussing the results in Section 5.3.2, we append the subscript "$_{norm}$" to denote normalized position accuracy values.

Bias and error

We also recorded raw position data for both blocks to compute *bias* and *error* for each tissue block placement (see Section 5.3.2). We define error as the median distance from every placed tissue block from the target block. This can be computed for all three dimensions, enabling us to describe position accuracy in a higher precision than just relying on a one-dimensional distance value. Bias, on the other hand, is the three-dimensional Euclidean distance $d(p,q)$ with the Cartesian coordinates for p being the target centroid position normalized to 0 and the coordinates of q being the median errors in the x, y and z-dimensions.

3D rotation accuracy

Rotation accuracy equals the angular difference between the two tissue blocks at task submission (see Figure 23). For ease of analysis, it was reduced to an individual number between 0 (exact same rotation) and 180 (diametrically opposite rotation). We used Unity's built-in Quaternion.Angle() function to compute this angle. Angle() takes two orientations, each consisting of three angles, expressed either as Euler angles or Quaternions, and returns a single float value between 0 and 180.

This means that several combinations of different rotations between tissue block and target block could yield the same angular difference. In order to preserve as much detail about the subject's action as possible, equivalent to the position, we logged the rotation of both blocks throughout the experiment as well. This allowed us to analyze the angular difference for all three axes (see Section 5.3.2).

Completion time

Completion time refers to the amount of time between the submission of a task and the submission of the previous task. Completion time is measured in seconds.

Performance plateau

During the Plateau phase (see section titled Plateau tasks and Figure 25), subjects perform 30 identical tasks, providing a unique opportunity to identify if and when a subject achieves a performance plateau. A plateau of a performance variable (task completion time, centroid accuracy, or rotation accuracy) is reached when the deviation of the performance variable does not exceed the mean performance of the subject until the end of the Plateau phase. As mean performance, we consider the average performance in a moving window of 20 tasks of the subject to reduce the influence of possible performance outliers. This width of the moving window supplies a stable mean by considering a certain inertia in performance improvement without including at all times the extreme values that can often be found towards the beginning and the end of the Plateau phase. For each subject, we analyzed after which task the performance stabilized by iterating through a recursive process, in which the relative deviation of the last task of the Plateau phase is calculated. If it does not exceed one (thus if the deviation of the performance variable in this task is not higher than its mean) we iterate this calculation for the previous task until we arrive at a task where the relative deviation is larger than 1. We consider all tasks after this (until the last task of the phase) to be on a performance plateau. If a subject reaches a performance plateau, we take the average performance (for example, mean completion time per task) of all the tasks that are completed after reaching this plateau.

Satisfaction

To assess user satisfaction, we included a corresponding item in the post-questionnaire via a five-point Likert scale from one (not at all satisfied) to five (very much satisfied), with three being a neutral value, and we report results aggregated by setup. This pertains to RQ6 (with results presented in Section 5.3.6).

5.3 Results

This section presents subject demographics, performance and satisfaction for all three setups, and a comparison of results plus discussion of requirements and research questions presented in Sections 5.1.2 and 5.1.3.

5.3.1 Demographics

We solicited 43 subjects for in-person user study appointments between 30 and 60 mins. We had to drop one subject from the analysis for not meeting the age requirement for participation, leaving us with 42 subjects. Subjects spent an average of 43 minutes with the experiment, including pre- and post-questionnaire.

The gender split in our experiment was almost exactly 50/50, with 20 female and 21 male subjects and one subject preferring not to specify gender. In terms of age, 10 were between 18 and 20 years old, 29 were between 21 and 30, one between 31 and 40, and two between 51 and 60. There were 34 English, four Chinese, two Bengali, one Russian, and one Spanish native speaker. All subjects except one were right-handed.

In terms of vision impairments, 20 indicated near-sightedness, four far-sightedness, three preferred not to answer, two reported astigmatism, one reported to be both far- and near-sighted, and one presbyopia. 11 subjects reported perfect vision. All subjects were allowed to wear glasses during the experiment.

5.3.2 Accuracy

To answer RQ1 and RQ2, we analyzed the data for differences in position and rotation accuracy in the tissue block placements between the three setups for the Plateau phase. Results are plotted in Figure 26.

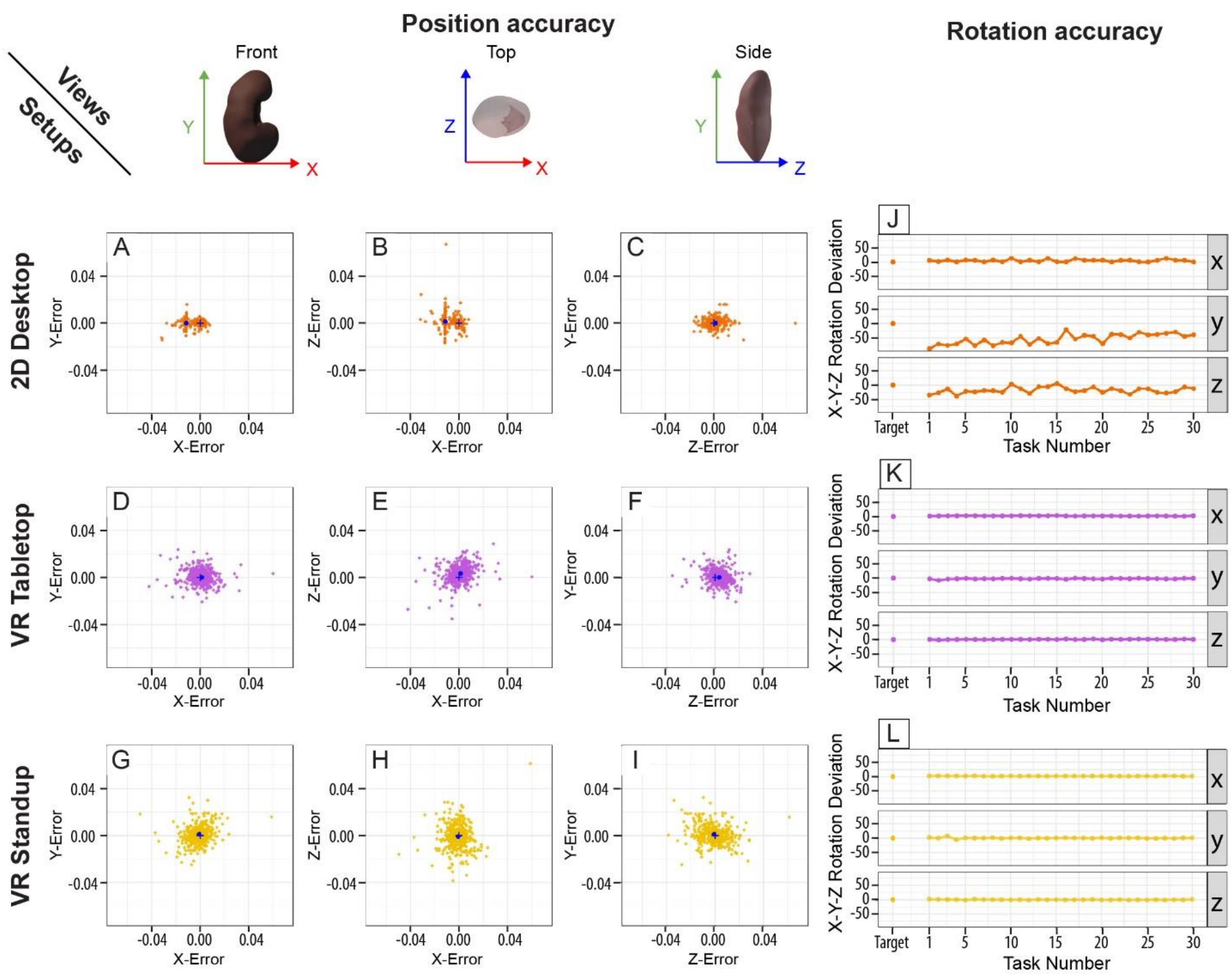

Figure 26. Graphs for position and rotation accuracy. ***A-I****: Scatter graphs showing the (kidney height) normalized error for position accuracy during the Plateau phase. Each dot represents one of the 30 tissue block placements. The blue cross at the origin of each scatter graph shows the location of the target block. The blue dot shows the average of all centroids (bias).* ***J-L****: Line graphs with rotation accuracy for each axis (x, y, z).*

When comparing **distance** between the centroids of the blocks for each user after they had reached their position accuracy plateau during the Plateau phase**,** we found **no significant difference** between 2D Desktop, VR Tabletop, and VR Standup subjects when normalized by the height of the kidney (2D $\text{Desktop}_{\text{norm}}$ = 0.0118, VR $\text{Tabletop}_{\text{norm}}$ = 0.0114, VR $\text{Standup}_{\text{norm}}$ = 0.0125), prompting us to **reject H1a**. Investigating the error for each axis, we found one error that stood out: 2D Desktop subjects tended to place the tissue blocks towards the negative space of the x-axis (median $\text{x-error}_{\text{norm}}$ = -0.01138; see Figure 26, A and B). This error for the 2D Desktop setup was significantly higher than the y- and z-errors ($p < 0.001$), prompting us to **confirm H2b**. We need to emphasize that the errors and biases are extremely minor. In fact, this median $\text{x-error}_{\text{norm}}$ for 2D Desktop (-0.01138) corresponds to just 1.138% of the kidney height of 113 mm, or 1.29 mm, which, in terms of gross-anatomical registration accuracy, is more than sufficient.

Further, this error could possibly be ameliorated through a change in camera control for the user. It is possible that the x-error occurs due to the main camera in the 2D Desktop setup being aligned with either the x-axis (side view of the kidney) or the z-axis (front view), oriented towards the positive x-axis space. This could have caused subjects to have a bias on that axis. We explain a planned improvement of the user interface in Section 5.4. The $\text{x-error}_{\text{norm}}$ for 2D Desktop caused a bias ($\text{bias}_{\text{norm}}$ = 0.01146) about three times larger than the bias for VR Tabletop ($\text{bias}_{\text{norm}}$ = 0.00372**)** and about 7.7 times larger than VR Standup **(**$bias_{norm}$ = 0.00148), prompting us to **confirm H2a**.

In terms of **rotation accuracy**, subjects in both VR setups outperformed 2D Desktop subjects with median rotation accuracies of **16.3 degrees** (2D Desktop), **4.3 degrees** (VR Tabletop), and **5.0 degrees** (VR Standup) during the Ramp-Up phase. The median Plateau levels were **5.88 degrees** (2D Desktop), **3.89 degrees** (VR Tabletop), and **4.67 degrees** (VR Standup). While the slight improvement for the VR setups can likely be attributed to the learning effect (since the Plateau phase came after the Ramp-Up phase), the jump in accuracy for 2D Desktop users stands out. We assume that many subjects became more familiar with the rotation sliders over time and were able to memorize the values for each axis as all tasks in the Plateau phase were identical. Figure 26, J-L, shows the rather severe differences in rotation accuracy not only between the setups but also between individual axes for 2D Desktop subjects during the Plateau phase. The mean deviation around the x-axis was relatively small (**4.9 degrees**) but rather exorbitant for the y-axis at (**-54.15 degrees**). We can see a clear upward trend for y-axis as subjects progressed through the experiment and improved over time (see Figure 26J, middle line graph). Given these results, we **accept H1b**.

5.3.3 Completion time

Task completion time for the three setups and both phases is shown in Figure 27 using a series of boxplots.

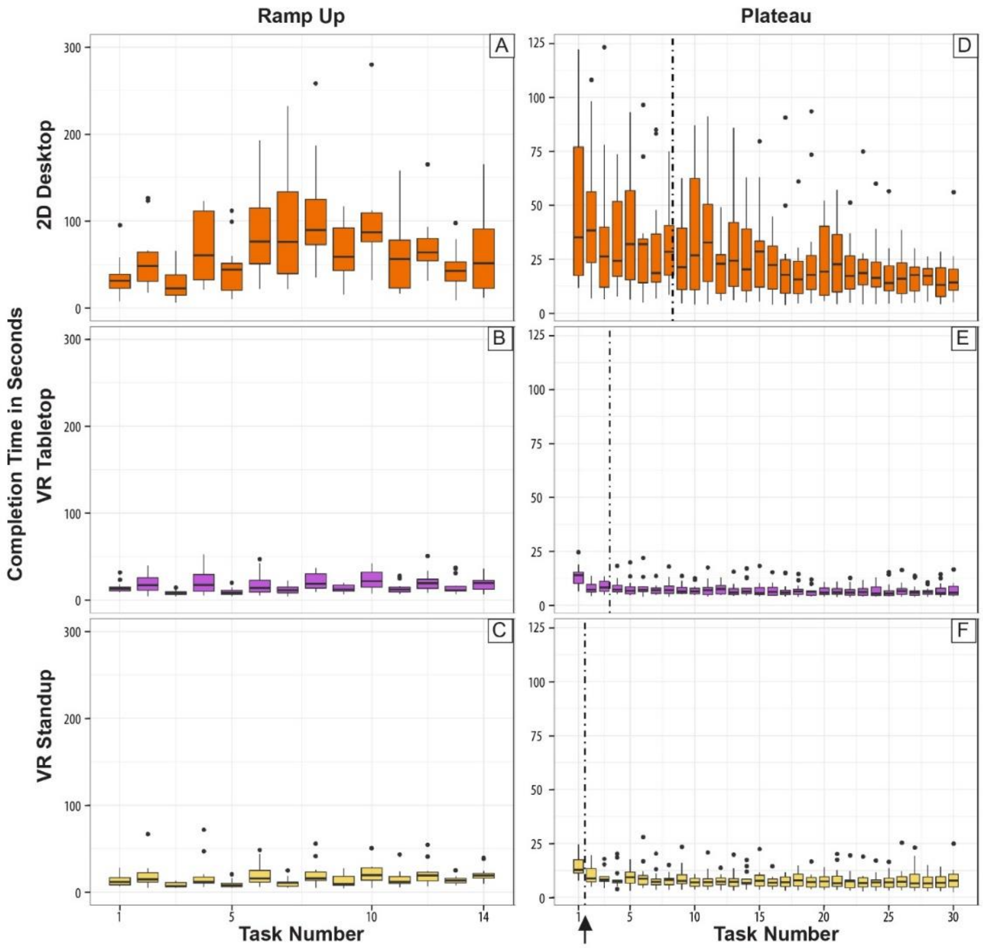

*Figure 27. Completion time for both phases and all three setups. **A-C**: Ramp-Up phase. **D-F**: Plateau phase. The vertical dash-dot line (black arrow) indicates after what task the plateau was reached, on average.*

In the Ramp-Up phase, we found significant differences between the 2D Desktop and both VR setups but no difference between VR Tabletop and VR Standup. On average, subjects needed **67.3 seconds** for a placement task in 2D Desktop but only **16.5 seconds** in VR Tabletop and **16.3 seconds** in VR Standup, yielding a significant difference in completion time. The results of the VR setups do not differ from each other significantly. Further, in Figure 27, one can clearly see the fluctuating medians for the completion time depending on the task. During odd tasks, subjects were given a prompt to focus on speed; during even tasks, we asked them to focus on accuracy. This is mirrored in the graphs for all three setups but especially so for VR Tabletop. We discuss this in more detail in Section 5.3.5.

Regarding RQ4, in the Plateau phase, the median Plateau level for 2D Desktop users was **22.6 seconds** after **8.3 trials** versus **7.1 seconds** after **3.43 trials** for VR Tabletop and **7.39 seconds** after just **1.5 trials** for VR Standup. Thus, it takes 2D Desktop subjects longer to reach a completion time plateau. Figure 27, D-F, shows the distribution of completion times during the Plateau phases. Given these findings, we **accept both H1c** (lower completion times for both VR setups) and **H4** (2D Desktop subjects need more trials to reach completion time plateau).

5.3.4 Influence of task complexity on accuracy and completion time

To answer RQ3, we computed the impact of task complexity on task accuracy and completion time during the Ramp-Up phase. Figure 28 shows position accuracy in mm on the y-axis (i.e., centroid distance), completion time in seconds on the x-axis, and task difficulty (circle area size) for each setup.

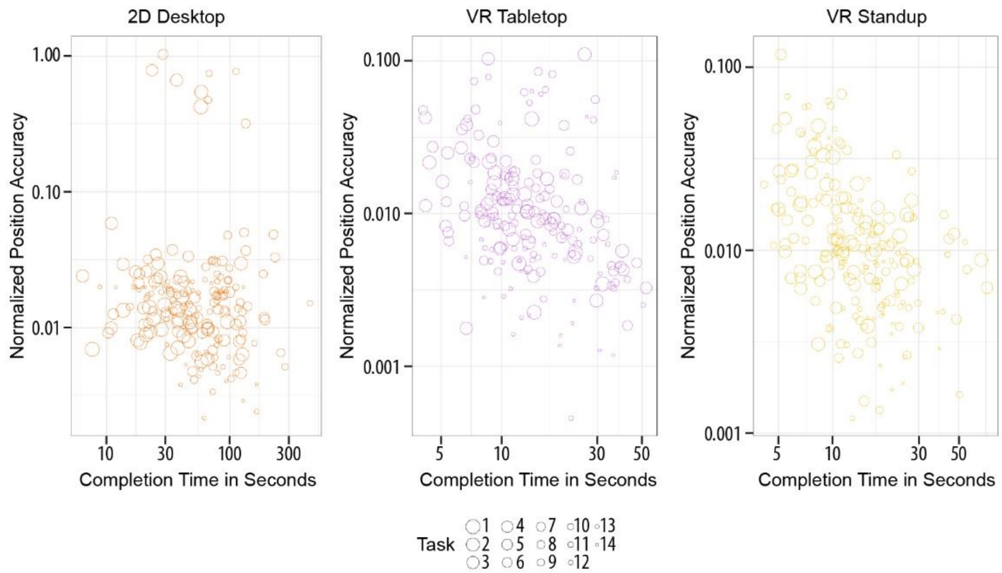

Figure 28. Position accuracy vs. completion time dependent on tissue block size, with a log-log scale.

For VR Tabletop and VR Standup, we can see a strong cluster of records at **13.3 seconds** and **13.4 seconds**, respectively (median completion time). This is less apparent for 2D Desktop (median = 54.1 seconds) where there is more than one cluster. As becomes apparent from Figure 28, we found no significant correlation between task complexity and position accuracy for any setup, requiring us to **reject H3a**. We did, however, find a significant and positive Pearson correlation between task complexity and rotation accuracy, for all setups (2D Desktop: 0.457, $p < 0.001$; VR Tabletop: 0.167, $p < 0.05$; VR Standup: 0.231, $p < 0.01$). We thus **accept H3b**. Finally, for completion time, we only found a significant, positive correlation for the 2D Desktop setup (0.163, $p < 0.05$) and thus **reject H3c**.

5.3.5 Tradeoff in speed versus accuracy

Next, we wanted to understand whether there was a gain in accuracy when spending more time on a task in the Ramp-Up phase (see RQ5). Here, the results vary greatly per setup. For 2D Desktop, we found no significant Pearson correlation between completion time and any accuracy measures. For VR Tabletop, we only found a significant negative Pearson correlation between position accuracy, expressed as centroid distance ($r = -0.18$, $p = 0.01$). If controlled for instructions the subject received at the onset of the task (focus on speed vs. on accuracy), however, it becomes evident this correlation is only significant for speed tasks ($r = -0.2$, $p = 0.05$), not for accuracy tasks. Finally, for VR Standup, we identified a significant negative Pearson correlation for both position ($r = -0.33$, $p = 0.0$) and rotation accuracy ($r = -0.22$, $p = 0.001$), regardless of instructions. Given these results and the evident differences between the setups, we **reject H5**.

Note that these results are surprising as we typically see an alignment between the VR setups; however, they diverge substantially here. A possible explanation could be the larger degree of freedom for movement afforded by the VR Standup, where subjects could walk around the kidney, crouch below it if needed, and spend more time on finding a workable angle. Naturally, this setup also required the most space.

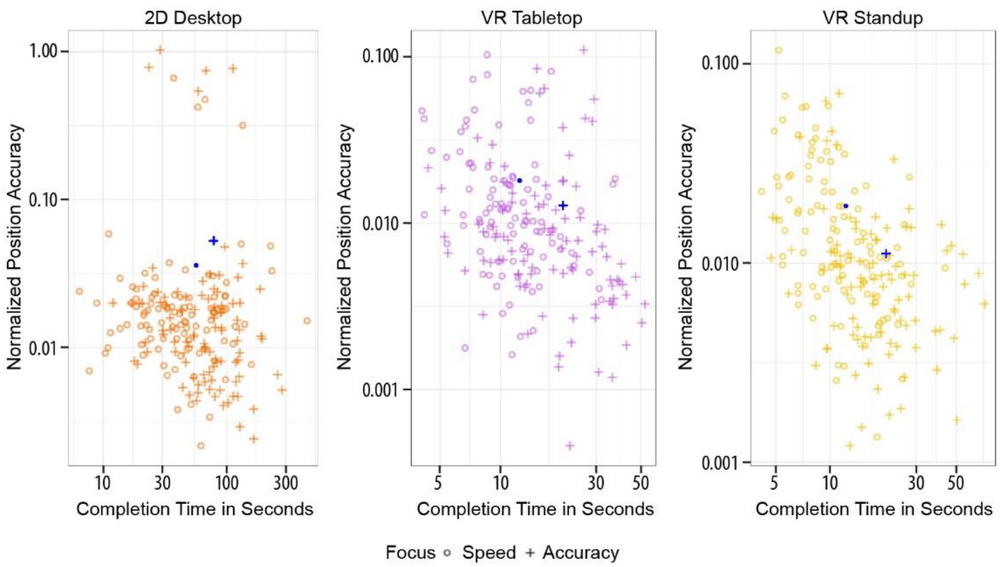

Figure 29. Position accuracy vs. completion time dependent on instructions. The blue circles and blue crosses mark the average completion time and position accuracy for speed and accuracy prompts, respectively.

We then performed a test to see whether subjects followed the prompts given to them when starting a new task. Figure 29 shows position accuracy by completion time. For VR Tabletop and VR Standup, we see a tendency for longer completion times for tasks with accuracy prompts. The same pattern is evident in the boxplots in Figure 27**B** and Figure 27**C** that follow an up-and-down pattern, depending on whether the task number is odd (speed) or even (accuracy). However, none of these differences in completion time and position accuracy for the two prompts are significant, and the pattern is even less present for 2D Desktop users.

5.3.6 Satisfaction

Finally, to address RQ6, we analyzed and graphed subjects' self-reported satisfaction using the post-questionnaire data (see Figure 30).

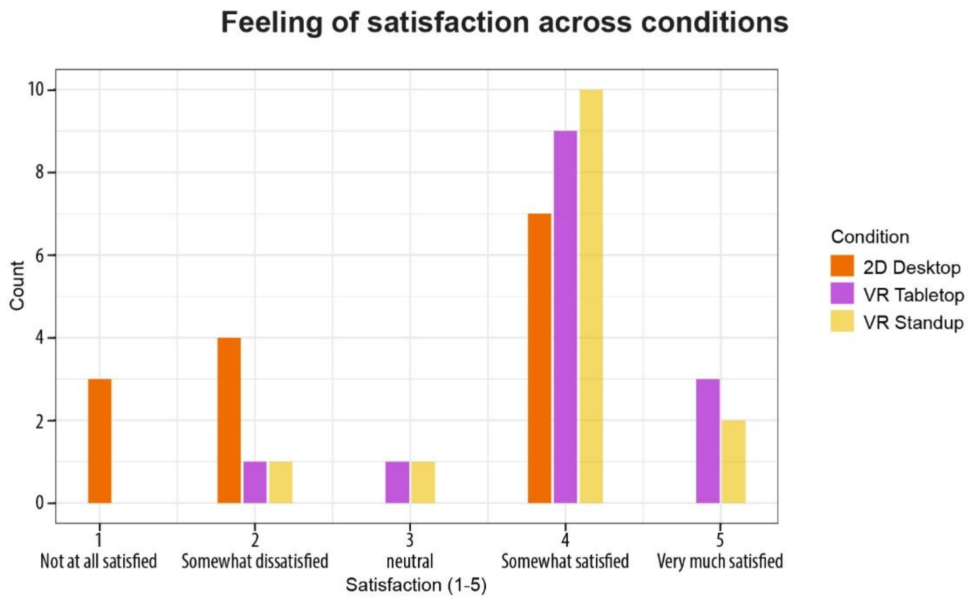

Figure 30. Grouped bar graph of overall user satisfaction.

Subjects used a five-point Likert scale with one (not satisfied at all) to three (neutral) to five (very much satisfied). With a mean of **3.6** across all setups, the satisfaction was on the positive side. We then performed a pairwise Kruskal-Wallis test (with adjusted significance level for alpha inflation correction). The Kruskal-Wallis test is a non-parametric test that allows to check whether more than two non-normally distributed samples are drawn from the same distribution—i.e., it assesses whether data samples differ significantly from each other (137). This yielded a significantly lower satisfaction for 2D Desktop users (mean = 2.79) compared to those in VR Tabletop (mean = 4) and VR Standup (mean = 3.93). The result of the VR setups does not differ significantly from each other. We thus **accept both H6a** and **H6b.** We found no correlations between satisfaction and prior experience with 3D software, first-person shooters, or VR.

5.3.7 Prior exposure to VR and 3D applications

Across the 42 subjects, there were only minor differences in previous experiences with 3D or VR applications. 24 subjects had used a VR headset before; 34 subjects had played video games in the past 12 months, 28 of whom had played first-person shooters (FPS). Further, 22 subjects had used 3D modeling software before. The largest differences appeared in previous exposure to VR between VR Standup and VR Tabletop (6 subjects vs. 10 subjects, respectively), but these did not result in a significant difference in performance. After running a comparison test, we found no differences in the distributions of completion time and accuracy measures grouping by sex, color blindness, vision impairment, age group, and right-/left-handedness. Additionally, we found no correlations of demographic variables, prior exposure to VR, or 3D applications with performance variables.

5.4 Discussion

This paper reported the results of a user study with 42 subjects involving 14 increasingly complex and 30 identical tissue block registration tasks across the 2D Desktop, VR Tabletop, and VR Standup setups. Our findings focused on comparing three different setups for the RUI in terms of **accuracy** (**position**, **rotation**), **completion time**, and **satisfaction**.

Contrary to our expectations, many of our predictions were not confirmed in the study. We expected the VR Tabletop and VR Standup subjects to outperform 2D Desktop users in all of these metrics; however, we only found this to be true for rotation accuracy (**H1b**), completion time (**H1c**), and satisfaction (**H6a**), but not for position accuracy (**H1a**). From our analysis (see Sections 5.3.2 and 5.3.3), we conclude that the VR users are about **three times as fast** as Desktop users and about **a third more accurate** in terms of **rotation** for a sequence of 30 identical tasks, but **similarly accurate** for **position when normalized for the kidney height**.

We argue that several factors contributed to the high position accuracy recorded for 2D Desktop subjects:

Restriction to two axes: 2D Desktop users only moved the tissue block in two dimensions at a time, and the main camera was always aligned to either the x- or the z-axis. We modeled this functionality after the "quad view," which is common in 3D modeling software. It allows the user to see a 3D object from three orthographic perspectives with an additional window showing a 3D view, facilitating more precise 3D alignment. This restriction for 2D Desktop users might have played a role in their high position accuracy. In prior work, the lack of such restrictions for 3D manipulation has been shown to be a source of frustration for novice users (168). Similarly, Masliah and Milgram (135) showed that even with advanced input devices, users separate translational (i.e., position) and rotation control when performing virtual docking tasks.

Precision of the mouse: The mouse proved to be a superior tool for performing fine adjustments, and the hand-eye coordination required to align the blocks seemed achievable for most subjects.

Separate manipulation of position and rotation: While position and rotation adjustments were performed by different tools in 2D Desktop (mouse and rotation sliders, respectively), VR Tabletop and VR Standup users performed both simultaneously (with their VR controllers). It is perfectly possible that many VR users achieved high position accuracy early on but worsened their result by subsequently adjusting the rotation of the tissue block.

Our analysis of position accuracy yielded an important insight for the continued development of the RUI. As explained in Section 5.3.2 and Figure 26, we observed an error in 2D Desktop tissue block placements, suggesting a tendency of users to place the tissue blocks according to the camera view (i.e., side or front) utilized at the time. A potential solution for this recurring error would be to implement **more than two predefined camera views**, thus giving the user multiple perspectives from which to view the reference organ.

This high position accuracy, however, was somewhat offset by the significant difference in rotation accuracy and completion time between the VR setups and the 2D Desktop setup. Yet, despite this inferiority, the 2D Desktop implementation meets the requirements outlined in Section 5.1.2. The tasks of the Plateau phase most closely resemble a real-world usage scenario, where multiple registrations are being performed in succession. With a median position accuracy of **1.3 mm given the kidney height on the laptop display**, 2D Desktop users got close to the goal of one mm for position accuracy. Similarly, at a median of **5.89 degrees**, the goal of rotation accuracy by 15 degrees is well met. Further, at **22.6 seconds,** the median task completion time plateau for Desktop users was within an acceptable range. In a real-world context, where the accuracy requirement is not as pronounced as it was in this study, we can expect that a reasonably accurate registration can be achieved in less time. In future studies, the research on accuracy from human tissue registration presented here will serve to support so-called Stage 2 registration at the single-cell level using image registration software and machine learning.

Finally, another goal was to make the RUI experience satisfying. In this regard, the 2D Desktop implementation was clearly lacking with a significantly lower self-reported satisfaction score than either of the VR setups (**H6a**), between which we found no significant difference (**H6b**). However, it is important to remember that this study only crudely approximates a real-world usage scenario, where the high level of accuracy and completion time suggested in this study is likely not necessary, resulting in less pressure on the user to keep adjusting their tissue blocks. This is corroborated by the fact that more time invested does not result in higher accuracy for the 2D Desktop setup (**H5**), making it more "forgiving" to users who choose to spend less time on getting a "perfect" registration. Additionally, the ease of use and wide availability of high-resolution 2D screens and computer mice is likely an advantage for users who have never experienced VR before.

Additionally, we can assume that 2D Desktop technology is less likely to cause technology frustration as 2D computer monitors, of various resolutions and size, and mice are widely available, easy to service, and use. As VR equipment becomes cheaper, less bulky, and easier to set up, VR setups may catch up, but at the time of this writing, 2D Desktop setups hold a clear advantage in this regard.

5.5 Conclusions

The insights gained in this study inform the continued development of the RUI Desktop setup as part of the HuBMAP Ingest Portal, see the recently released RUI 1.5 (65). The revised RUI is optimized for Google Chrome, Firefox, and the latest

(Chromium-based) version of Edge. By February 2021, 45 tissue blocks have been registered with the RUI (15 for left kidney, 11 for right kidney, 15 for spleen, and 4 for colon). The average size for the 26 kidney blocks is: 22.7 mm x 16.7 mm x 5 mm (H x W x D) for the left kidney, and 23.7 mm x 11.7 mm x 6.7 mm for the right kidney.

Going forward, we envision two types of user studies exploring 3D manipulation further.

First, we plan to run studies in a more “in the wild” setting (144). This would allow us to consider variables that are hard to test in a lab setting with mostly novice users, and result in more accurate data about user performance and satisfaction in a true production setting. This would likely be a more focused study with a smaller sample of subject matter experts at their place of work (i.e., a wet lab or adjacent data processing facility), and would enable us to evaluate the performance of the 2D Desktop RUI in a realistic usage scenario.

Second, it would be valuable to test how interventions could help users improve their performance during the experiment (e.g., between the Ramp-Up and Plateau phases). Specifically, we aim to run a study with a “reflective” phase where the user sees a visualization of their own performance data from previous tasks before completing a second set of tasks. Our goal is to use the human ability to recognize patterns and trends visually to test if different types of interactive data visualizations can help users formulate strategies to improve their performance in terms of position accuracy, rotation accuracy, and completion time. Given the detailed telemetry data collected from RUI users (especially those in VR), a natural next step would be to add an intervention where users can see their own movement as well as the position and rotation of the tissue and target blocks over time, thus enabling them to detect problems and strategize more efficient solutions for future tasks.

6. User Study 2: Using VR Data Visualizations to Improve Time, Accuracy, Satisfaction in 2D Desktop vs. VR RUI

6.1 Introduction

Due to decreasing cost and an increasing amount of hardware choice, VR has become a popular entertainment tool. Devices like the Valve Index (https://store.steampowered.com/valveindex) and the Oculus Rift S (https://www.oculus.com/rift-s/) offer a wide variety of content that can run independently of the platform they were purchased on: movies, video applications, virtual desktops, and an ever larger number of video games. More recent devices such as the Oculus Quest 2 (https://www.oculus.com/quest-2/) offer VR at even lower prices and without the need for a strong PC or laptop for rendering. Software development kits (SDKs), among them OpenVR (204) and SteamVR (205), allow developers to produce, test, and deploy content to a wide variety of VR devices in a vendor-agnostic and unified pipeline using game engines like Unity (200) or Unreal Engine (85). Further, there is an increasing market for coaching and training new and old employees for retail, maintenance, and administrative jobs.

Depending on the hardware and the needs of the application, users of VR equipment generate position and rotation data at a rate of up to 120 Hz, and every button press can be logged and associated with a time stamp via telemetry. In addition to these physical variables, additional data can be derived via computation at runtime or in later analysis, allowing designers and researchers to measure a user's performance and behavior when completing tasks such as arranging objects or navigating spaces. The novelty of VR, while demonstrably exciting and invoking a feeling of presence in users (23), brings with it challenges due to its unfamiliarity to many users. With the basis of training being repetition and improvement over time, methods to assess and improve one's performance are necessary. The visual primacy of VR, along with the ready availability of user data, makes data visualization a good tool to allow users to gain insights into their own data.

In this and the following chapters, we describe two user studies where we developed interventions to improve VR performance for manipulation (RUI VR, see chapter 6) and navigation tasks (Luddy VR, see chapter 7). These VR visualizations were developed using the DVL-FW (31, 32, 35), a theoretical toolset to interpret, construct, and teach data visualizations. The DVL-FW comes with a series of seven typologies to categorize, among others, visualization types (such as graphs and maps), visual encodings via graphic symbols (such as points, lines, volumes) and graphic variables (such as color hue and size), and interactions with data (such as filter as well as link and brush). We use the DVL-FW to describe the data visualization interventions with an abstracted terminology that expresses both traditional, 2D data visualizations (like bar graphs and line graphs) and advanced VR visualizations such as the ones presented here. Of special interest is the implementation of four interaction types (filter, navigate, animate/replay, link and brush): We enabled the subjects to filter their data by time stamp or graphic symbol (RUI VR) and task number (both studies); users could navigate freely around their data, which was displayed in its original spatial context on a 1:1 scale (RUI VR) and minimized (Luddy VR); it was possible to play back the data by time stamp in different speeds via a time slider (RUI VR); and subjects could select bars in a bar graph and then apply filters correspondingly to view only specific tasks based on their completion time. The goal of these studies was to test whether significant differences in performance and satisfaction were measurable between the control and experiment cohorts and to determine the effects between behavior in the intervention and performance in the subsequent set of tasks.

In chapter 6, we describe the research questions, implementation, study design, and results of the RUI VR study. Likewise, in chapter 7, we present the subsequent Luddy VR study. Finally, in section 7.10, we compare the implementations of the data visualization interventions for these studies to determine the cause for our results. We conclude with a series of design implications for future data visualization interventions to improve performance in VR before ending the chapter with an overview of limitations and next steps.

This research was conducted in unison with a user study about differences in performance and satisfaction between three implementations of the same interface (43). In this initial study, 42 participants, split across three setups, performed 14 increasingly difficult and then 30 identical 3D matching tasks either using a VR HMD while standing or sitting, or with a traditional 2D screen on a laptop. In this user study, we added a second cohort of 42 subjects (for a total of 84) while introducing an intervention between the Ramp-Up (increasing difficulty) and Plateau phases (identical difficulty). During this “Reflective phase”, the participants could first learn the controls and familiarize themselves with a visualization of the best-performing subject from the control cohort in the same setup (2D Desktop, VR Tabletop, VR Standup) and then explore their own data with the goal of enabling them to turn their insights into action.

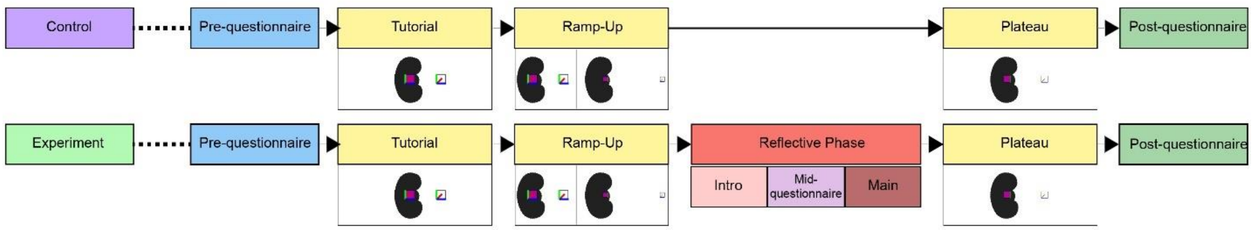

Figure 31. Study procedures for control (top, corresponds to original RUI study) and experiment cohort (bottom).

Upon arriving at the research site, the participants answered a pre-questionnaire to collect demographic data alongside information about their prior exposure to virtual reality, video games, and 3D applications in general (such as 3D modeling software). Additionally, their answered questions about whether they were left- or right-handed and if they suffered from any vision impairments (such as near-/far-sightedness or color-blindness).

Subsequently, the subjects completed the first part of the experiment proper, consisting of a tutorial task, followed by the first 14 tasks (Ramp-Up phase). Following that, the study procedure for both cohorts started to differ: The control cohort went on to the second set of tasks (30 identical ones, called "Plateau" phase). The experiment cohort, meanwhile, performed the Reflective phase.

6.2 Research questions and hypotheses

RQ1: When presented with a VR visualization of their own movement data from the **Ramp-Up** phase during their **Reflective phase**, do users in the **experiment** cohort have a better performance in the following **Plateau phase**, measured in accuracy and completion time, compared to the **control** cohort who does the Plateau phase **without** a **Reflective phase**?

H1: There will be a **significant** difference in completion time and accuracy for Ramp-up and Plateau phases between control (without Reflective phase) and experiment group (with Reflective phase). However, this will only occur for the VR Standup and VR Tabletop setups; the 2D Desktop users will not be able to gain significant gains in accuracy and completion time over their peers.

RQ2: Between-subject and within-condition, across cohorts, are there significantly different usage patterns between the experiment and control cohort during the Plateau phase?

H2: Users in the experiment group that previously spread out across a larger area in the Ramp-Up phase will use less space and concentrate on an overall smaller work area in the Plateau phase. They will also use less space in the Plateau phase on average than the control group.

RQ3: In the Reflective phase, how do users apply the interactive tools?

H3a: Most users will use the time slider to scroll through around 1000% (=10 times) the time span of their dataset.

H3b: The most selected location for the play head of the slider will be towards the very end of the timecode in the dataset.

H3c: Users will spend the majority of time with the kidney turned on as the presence of a reference organ is highly useful to understand the data overlay. The kidney is turned on by default.

RQ4: Within the experiment group, in the Reflective phase, are metrics on user actions and interactive tool usage, measured through telemetry, correlated with higher performance in the Plateau phase?

H4a: More distance traveled has a negative effect completion times in the Plateau phase.

H4b: More distance traveled has a negative effect on distance (higher position accuracy) in the Plateau phase.

H4c: More head rotations have a negative effect on completion times in the Plateau phase.

H4d: More head rotations have a negative effect on distance (higher position accuracy) in the Plateau phase. This may be due to high-performing users feeling more comfortable in 3D environments in general, and VR specifically, enabling them to move around their own data more fluently in the first place.

RQ5: In the mid-questionnaire between the intro and main part of the Reflective phase, what is the relationship between the task score for this questionnaire and the performance in the Plateau phase for 2D Desktop users?

H5a: There **is a** significant negative correlation between task score in the mid-questionnaire and position accuracy in the Plateau phase in terms of distance.

H5b: There **is a** significant correlation between task score in the mid-questionnaire and position accuracy in the Plateau phase in terms of error and bias.

H5c: There **is a** significant negative correlation between task score in the mid-questionnaire and rotation accuracy in the Plateau phase.

H5d: There **is no** significant negative correlation between task score in the mid-questionnaire and completion time in the Plateau phase.

H5e: The majority of users agree or strongly agree that the subject shown to them in the Reflective phase was highly fast and accurate.

6.3 Study design

In order to allow the users in the experiment cohort to inspect their own data from the Ramp-Up phase, we created a separate Unity application with the same base map, i.e., kidney and buzzer, as the Ramp-Up phase (for users in the VR Tabletop and VR Standup setups). 2D Desktop users were presented with a line graph visualization of the distance and angular difference between the tissue and target blocks. In this section, we outline what implementations of the Reflective phase looked like for each setup, the visual encoding, the interactivity, and the mid-questionnaire that concluded the Reflective phase before subjects continued with the Plateau phase. The Reflective phase consisted of two parts: an **intro** and a **main** part. In the intro, the user was shown a visualization of the best-performing subject in the control cohort of their setup in terms of completion times and position as well as rotation accuracy. A ~6 minutes tutorial (~3 minutes for 2D Desktop) introduced the goal of the Reflective phase, the visual encoding, the interactivity (for VR subjects), and prompted the user to derive strategies for faster and more accurate placement going forward. Following the intro, we presented the user with a mid-questionnaire (see section 6.3.4). Subsequently, the user was shown their own data in the main part of the Reflective phase. While we measured the time spent in the Reflective phase by the user, we did not impose a minimum or maximum time limit on the user.

6.3.1 Reflective phase across setups

In this section, we outline the Reflective phase implementation for all three setups. Note that for each subject, we omitted the data from the tutorial task (Task #0).

2D Desktop

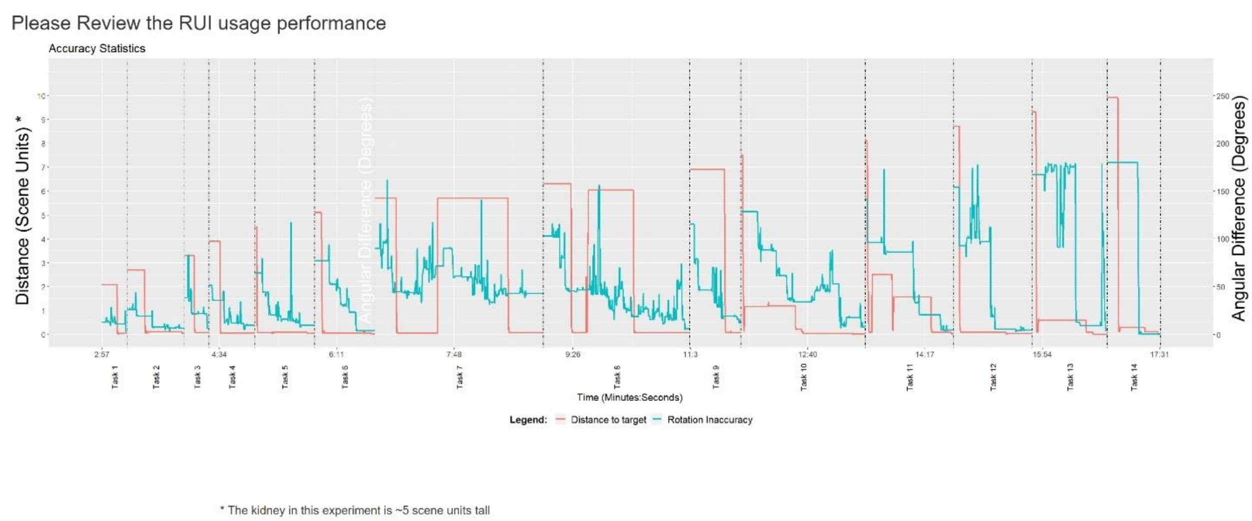

Figure 32. Line graph of distance between tissue and target block (orange) and angular difference (green) for the best user in the control cohort for 2D Desktop.

2D Desktop users were shown the line graph in Figure 32. On the x-axis, we plotted the elapsed time in seconds as well as task numbers; on the y-axis, we added two scales: distance between the two blocks (left side, measured in Unity scene units) and the angular difference (right side). Additionally, we inserted vertical dot-dash lines to indicate the end of one task and the beginning of the next one. This static visualization was created using R and Shiny after loading the CSV files with data from the subject's Ramp-Up phase.

VR Tabletop

In the VR setups, we used the inherently spatial reference system of the virtual environment to produce 3D dot density maps, encoding the headset, hand, and tissue block positions over time. Figure 33 shows the Reflective phase setup for a user in the VR Tabletop setup. They were seated at the same virtual and physical tables as they were during the Ramp-Up phase, ensuring a 1:1 mapping. They were allowed to move around during the Reflective phase and inspect their data form multiple angles.

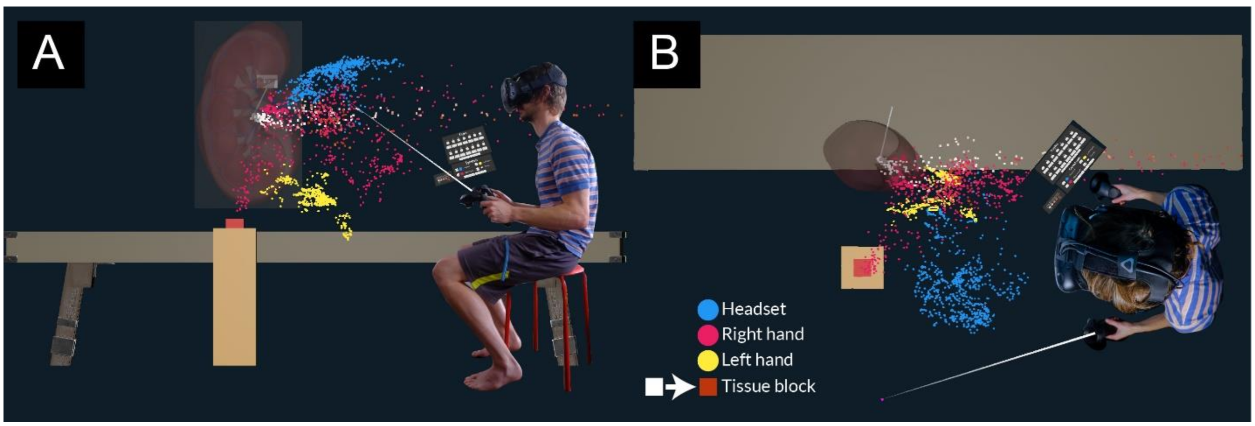

*Figure 33. Reflective phase setup for VR Tabletop. **A**: side view. **B**: top view.*

Figure 34 shows top and side views for three users with distinct movement patterns. A difference is noticeable, for example, in the movement of the left hand (yellow dots). In Figure 34, user A occupies a relatively small area compared to user C, with user B being somewhat in the middle. Similarly, user A has a higher vertical spread of dots with less density (bottom view), while users B and C cover less height. Contrast this, one the other hand, with the movement pattern of the best-performing user from the control cohort in the VR Tabletop setup, whose data we see in Figure 33. This user works in a highly compressed area, avoiding to cover large distances or lingering outside their small work area any longer than absolutely needed. They also minimized the movement of their left hand, as becomes apparent from the small spread of yellow dots.

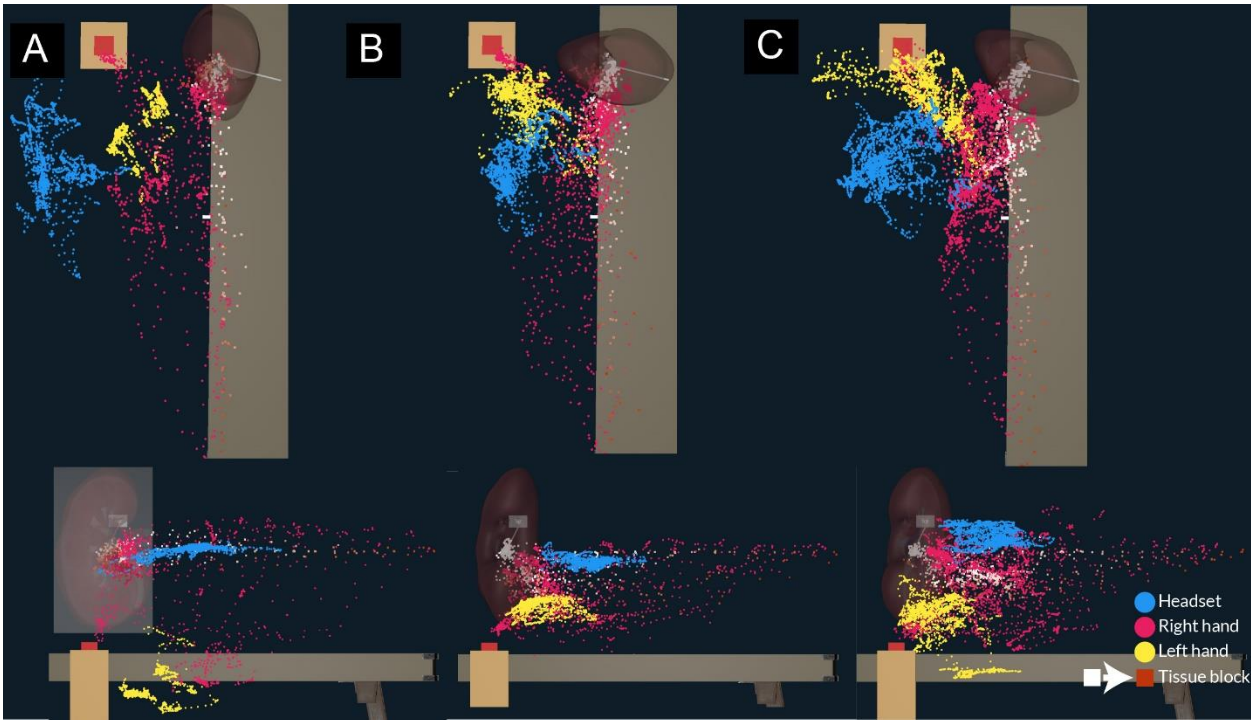

*Figure 34. Top view visualizations for three subjects with various spatial usage patterns (Tabletop). **A**: concentrated work on one axis. Middle: plenty of back-and-forth along the z-axis. Right: wide spread around the z and x-axis.*

VR Standup

Similar to subjects in the VR Tabletop setup, VR Standup users explored first the data of the best-performing users from the control cohort and then their own. Figure 35 contains side (A), top (B), and back view (C, with some transparency added for clarity).

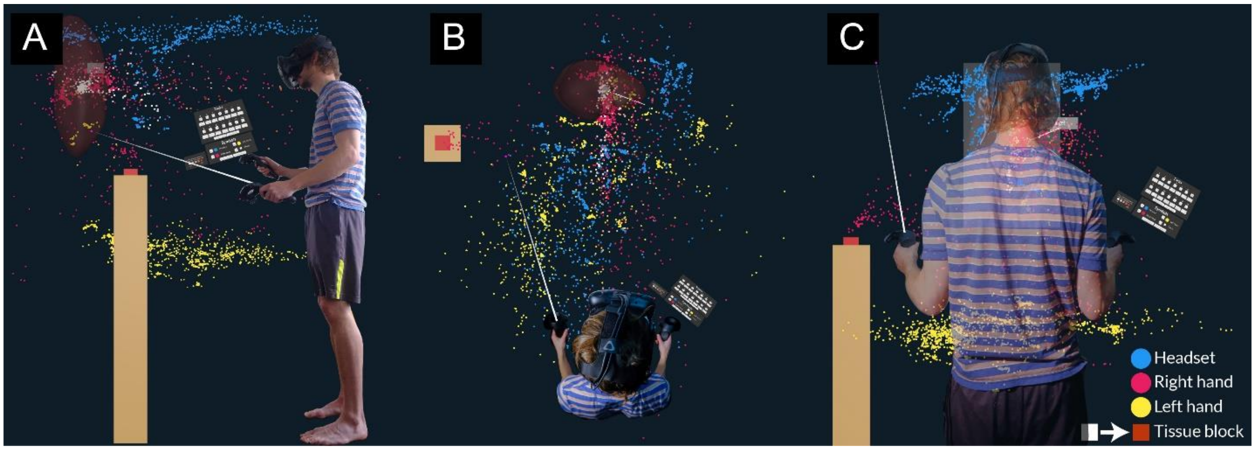

Figure 35. ***A****: side view of Reflective phase (VR Standup) scene with data overlay of best-performing user in control cohort.* ***B****: top view, same dataset (transparency was added to the user to avoid occlusion).*

Figure 36 shows the data for three distinct users. User A moved their head (blue) predominantly along one axis, minimized movement with the left hand (yellow), and kept the area they cover as small as possible. User B, on the other hand, moved back and forth along the z-axis (between the target block and the tissue block) frequently and at great speed, visible from the many blue dots distributed along that axis. However, they mostly stayed on one side of the z-axis. User C, however, not only moved back and forth frequently, they also switched from one side of the axis to the other multiple times.

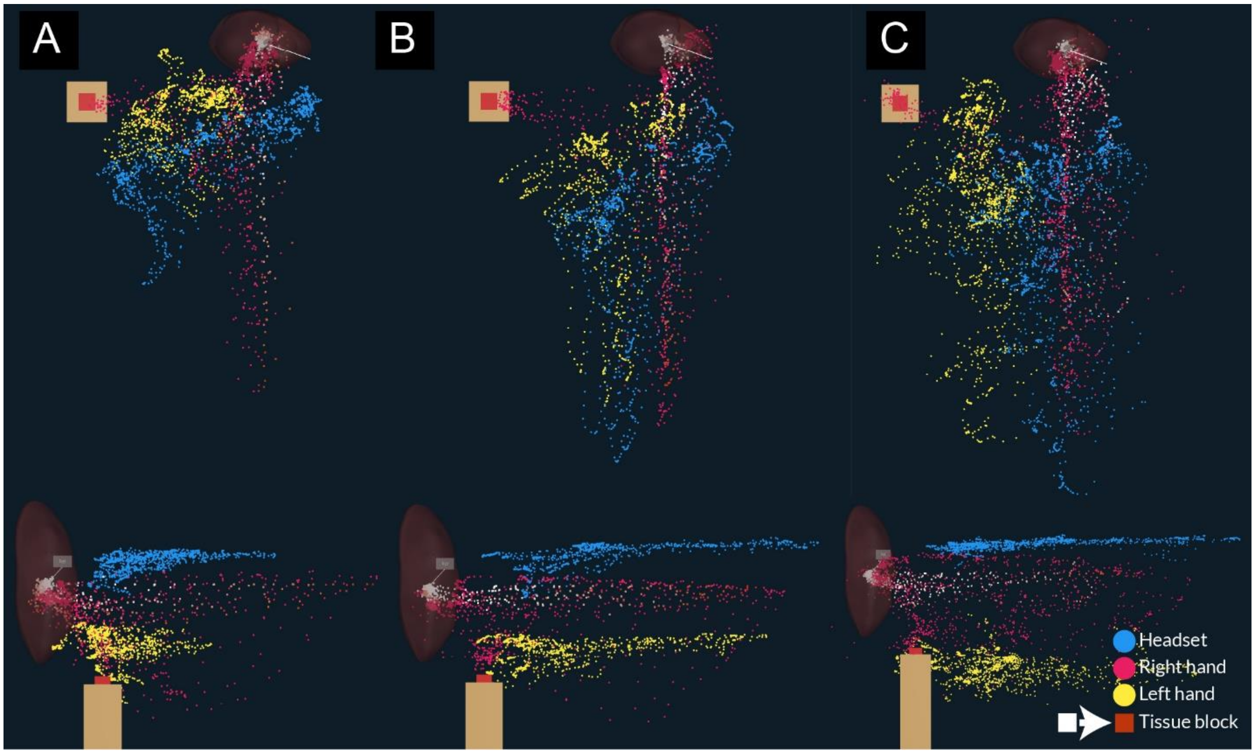

Figure 36. Top view visualizations for three subjects with various spatial usage patterns (Standup). Left: concentrated work on one axis (see headset). Middle: plenty of back-and-forth along the z-axis. Right: wide spread around the z and x-axis.

Just like in the VR Tabletop setup, users in the VR Standup setup were allowed to explore the 3D dot density map freely by walking around the space while using the kidney and buzzer as a base map.

6.3.2 Visual encoding

Like the Reflective phase implementations for the three setups, the visual encodings applied to each setup differed as well. In this section, we describe the graphic symbol and graphic variable pairings (31) used to encode the telemetry and task performance data from the Ramp-Up phase.

2D Desktop

In the 2D Desktop setup, the graphic symbol line encoded two data records: position accuracy, expressed as the distance between the tissue block and the target block over time, and the angular difference, i.e., the difference in rotation between the two blocks, expressed as a single value from 0 (same rotation) to 180 (diametrically opposed rotation). The graphic variables x-y position and color hue encode position and rotation accuracy, respectively. Additionally, various linguistic and pictorial symbols provide additional information to the user: axes are properly labeled; a note at the bottom of the graph indicates the height of the kidney so that reading the position accuracy values becomes easier; vertical dot-dash lines mark the beginning of a task and the start of the next; white gridlines help with reading values off the y-axes. As a temporal visualization, the x-axis contains the elapsed time in minutes and seconds (since the end of the tutorial task).

VR Tabletop and VR Standup

For the two VR setups, we used a straightforward visual encoding scheme for the user's HMD and controllers: blue for the HMD, pink for the right controller, yellow for the left controller, white to orange for the tissue block over time, see Figure 37.

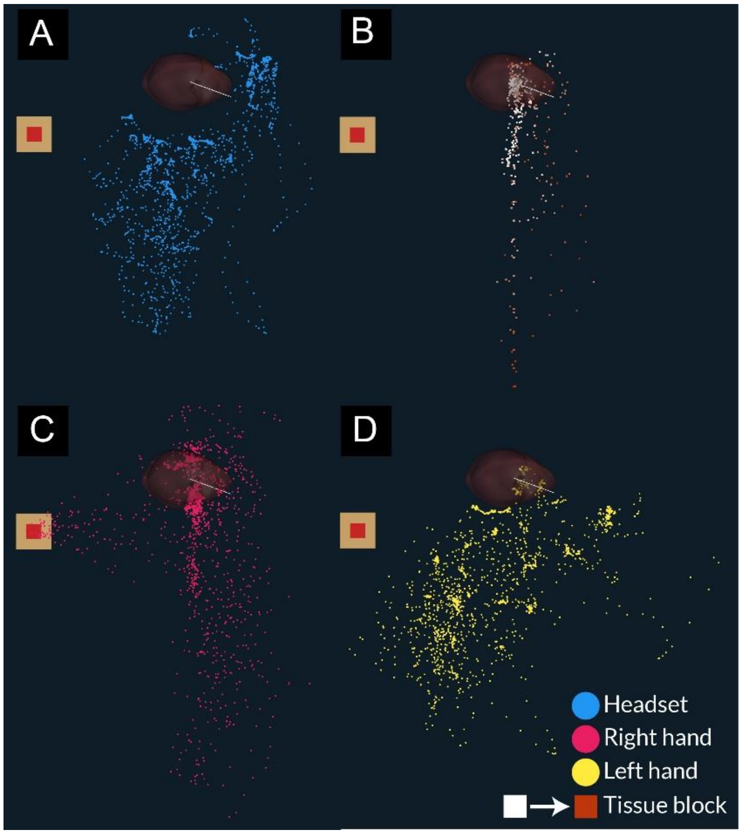

Figure 37. Data overlay for the VR setups separated by color. ***A****: HMD.* ***B****: tissue block.* ***C****: right controller.* ***D****: left controller.*

The graphic symbol volume, together with the graphic variable color saturation, encodes the angular difference between the tissue block and the target block and was indicated to the user in a legend, see Figure 38. For all graphic symbols, the graphic variable x, y, z-position encoded the x, y, z-position if the corresponding device (HMD, controllers) or virtual object at a given moment in time.

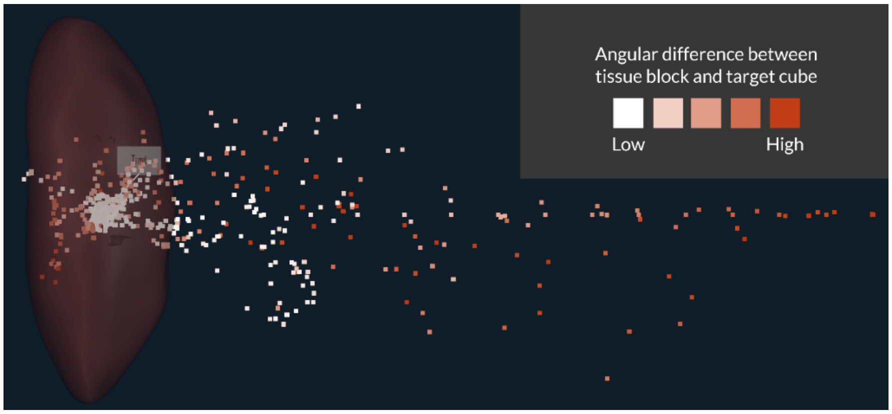

Figure 38. Distribution of tissue block locations over time, with angular difference between the tissue and target block encoded with a sequential color scheme.

The resulting visual encoding allowed users to quickly identify areas of concentrated activity. Frequently visited areas of space were thus indicated by a higher density of dots. Every user, by nature of the experiment, produced such a hot spot around the "Next Task" buzzer (see Figure 39).

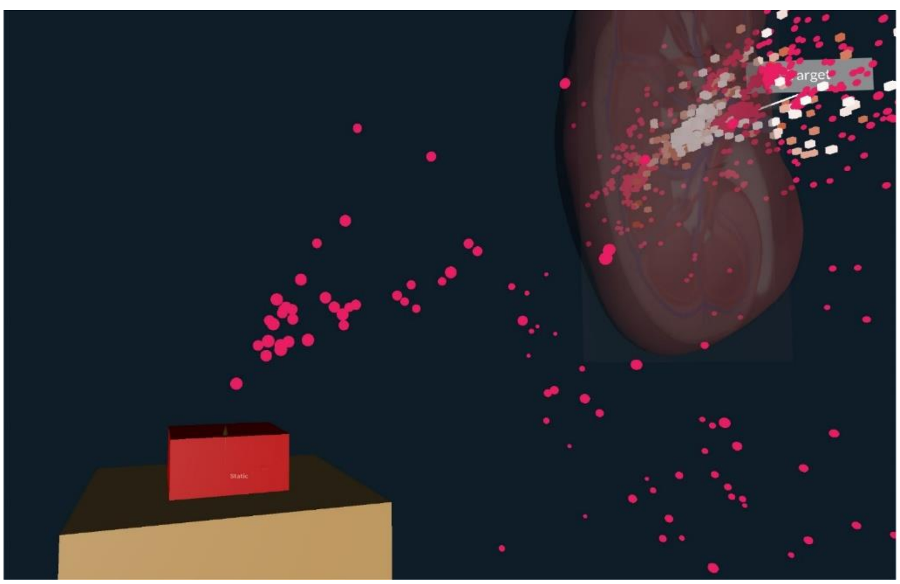

Figure 39. The user's repeated pressing of the virtual red buzzer produces a hot spot.

6.3.3 Interactivity

The aforementioned areas of concentrated activity were visualized in an aggregate view that the user encountered when they first entered the stage. By default, all data was shown; as a consequence, patterns in the movements were easier to spot, but the large number of dots generated by the user over the course of the Ramp-Up phase also led to visual clutter (see Figure 40). To allow the user to remove various layers of data through filtering of their choice, we implemented two interactions: **filter** and **animate**.

*Figure 40. Distribution of tissue blocks around the target cube (labeled) with graphic symbols for right hand and kidney visible (**A**) and invisible (**B**).*

The area around the target block position shown in Figure 40A tended to amass a large amount of data records due to frequent user activity when fine tissue block placement was performed. A mix of pink dots and white-orange cubes visualized the user's right hand placing the tissue block while minimizing the angular difference. By using their controller, the user could remove parts of the base map (the kidney) as well as parts of the data overlay by a series of features: **graphic symbol type**, **time stamp**, and **task number**. In Figure 40B, for instance, the user has removed the graphic symbol for the right hand (pink dots) to declutter the display around the target.

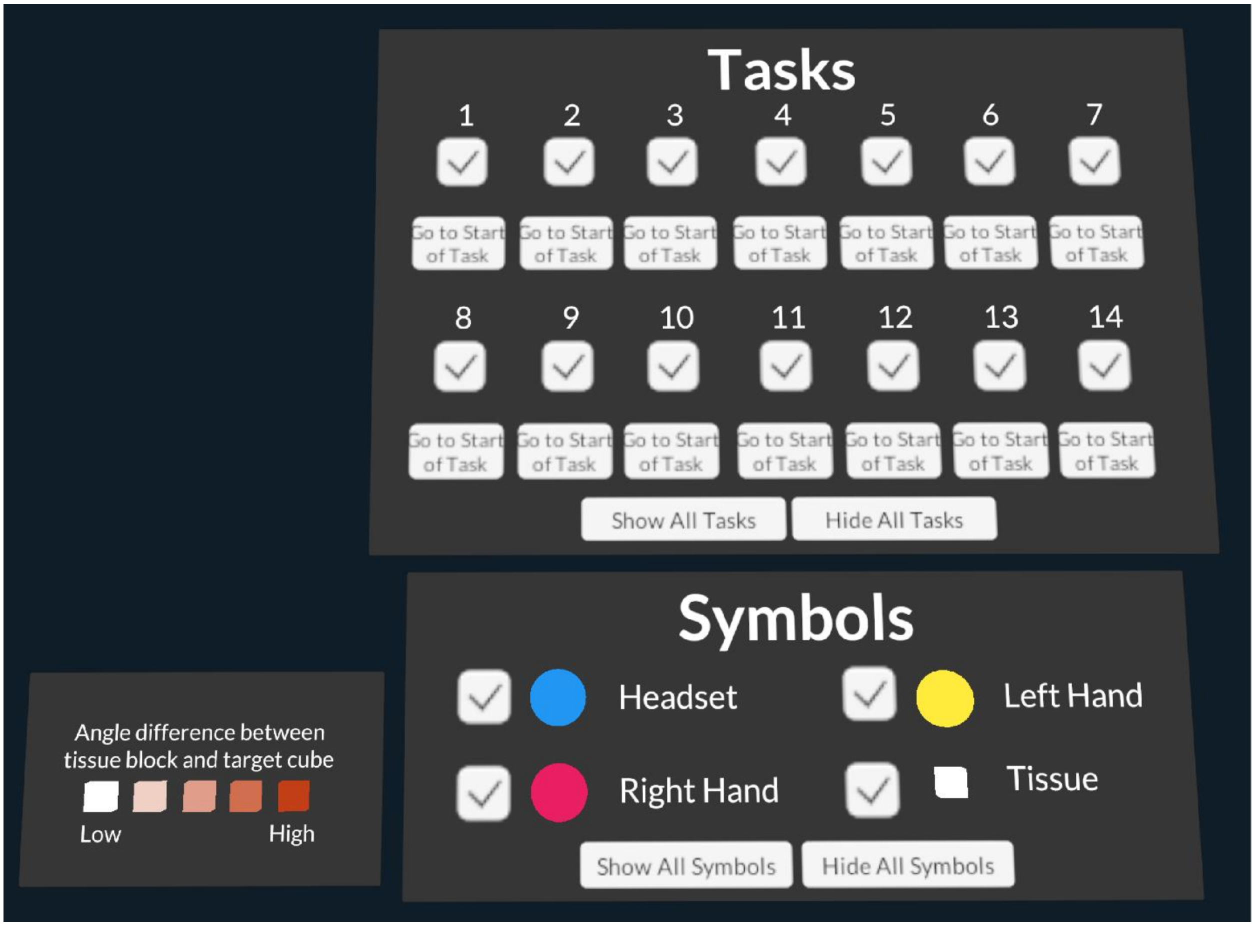

Figure 41. Filter menu to turn parts of the data overlay on and off by task number or graphic symbol type.

Figure 41 is a screenshot of the interactive legend presented to the user on top of their right controller. It consisted of three sections: **graphic symbols**, **tasks**, and a **static legend** for the angular difference. The graphic symbols part allowed the user to turn parts of the data overlay on and off by entire types of data records encoded by these symbols. Checkboxes enabled the user to **filter** by graphic symbol type and task number.

Similarly, Figure 37 above shows an entire dataset separated by graphic symbol type. Figure 42 below, on the other hand, shows the dataset in chunks: tasks 1-5 (A), tasks 10-11 (B), and only task 14 (C). For both of these filters, we added “Show All” and “Hide All” buttons for the user’s convenience.

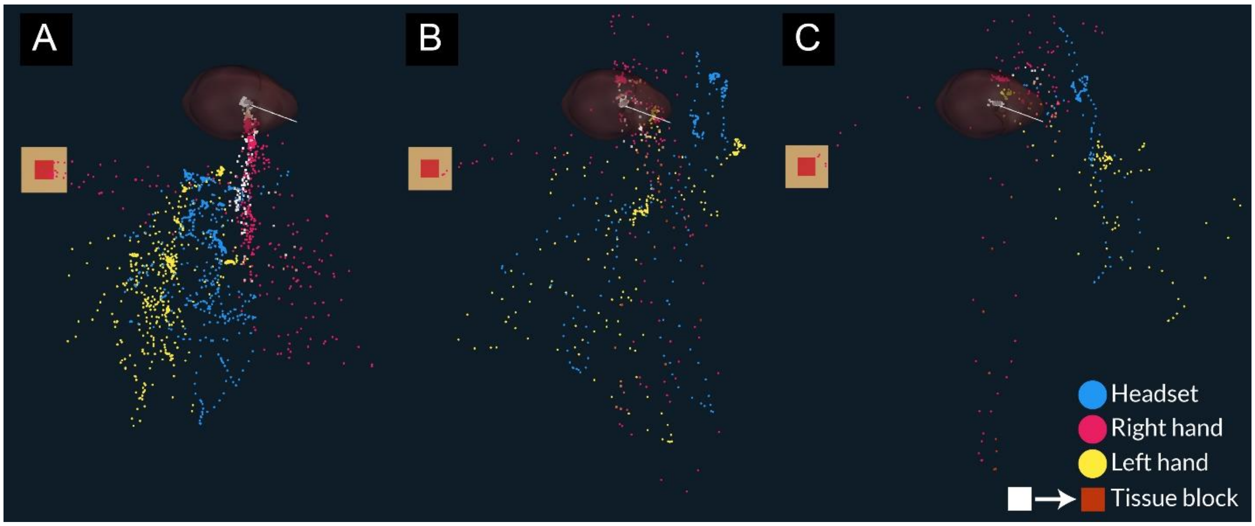

Figure 42. VR Standup user with three different stages shown. A: Tasks 1-5. B: Tasks 10-11. C: Only task 14.

Lastly, we allowed the user to show and hide data records by time stamp. Specifically, we implemented a time slider on top of the user’s left controller (see Figure 43).

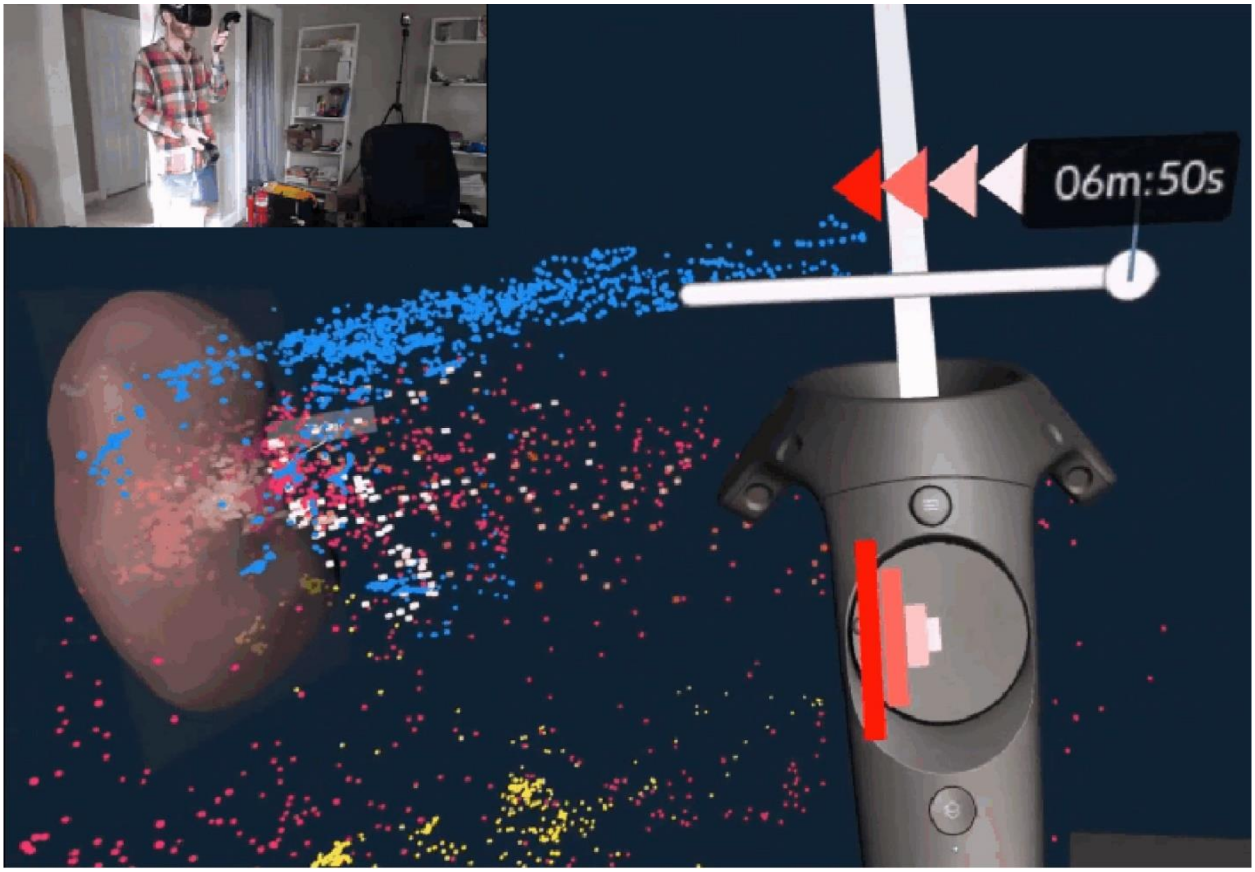

Figure 43. Time slider to skip forward and backward in time by using via the thumbpad on the VR controller.

The time slider consisted of a slider area with a play head, similar to what one would find in a video editing program. The user could move the play head along the slider by putting their thumb onto the trackpad on the left controller (see Figure 44).

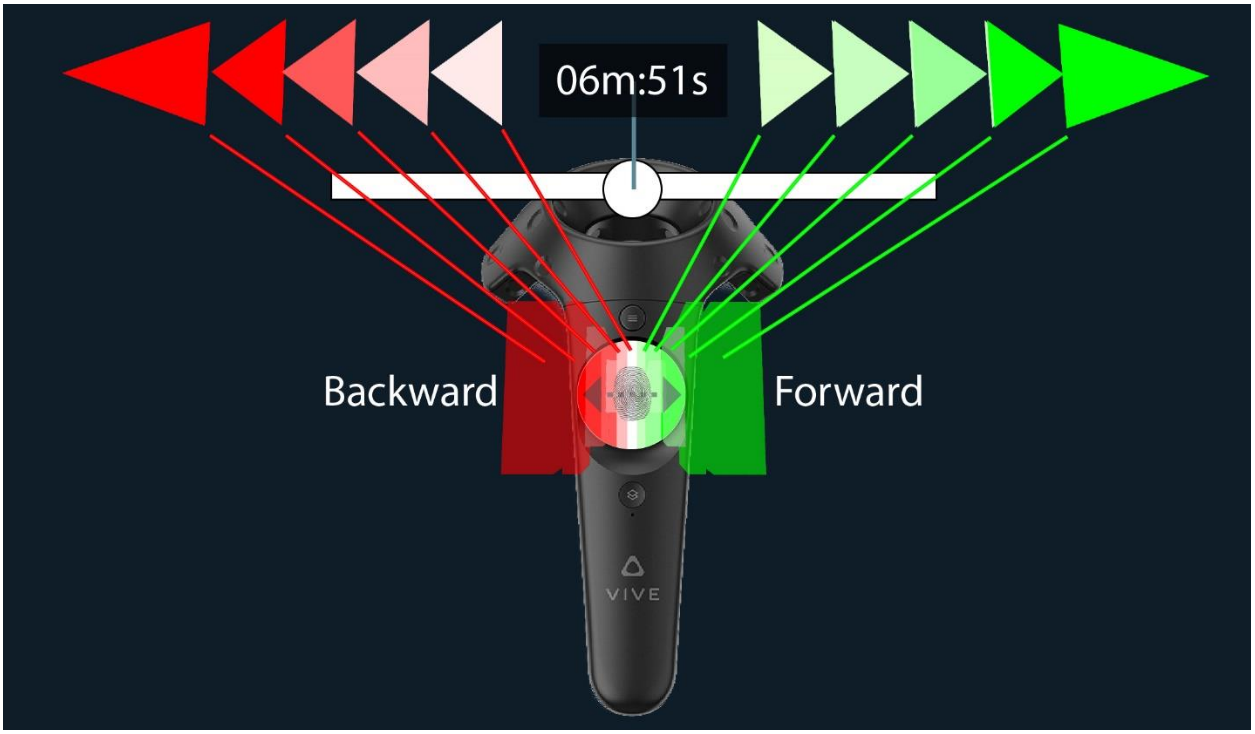

Figure 44. Close-up of the left controller with play head speed zones.

Placing the thumb onto the right and left half of the trackpad let the user skip forward and backwards through the dataset by time stamp, respectively. This allowed the user to replay the dataset at various speeds depending on their horizontal distance from the center of the touchpad. Additionally, they could activate a fast-forward and fast-backward mode when additionally pressing the trigger button on the back of the controller. To indicate to the user the current speed at which they were skipping through the dataset, green or red arrows were displayed next to the current time stamp (in minutes and seconds since the beginning of task 1). In parallel, 3D green and red blocks were displayed over the touchpad in the user's virtual view (see Figure 43).

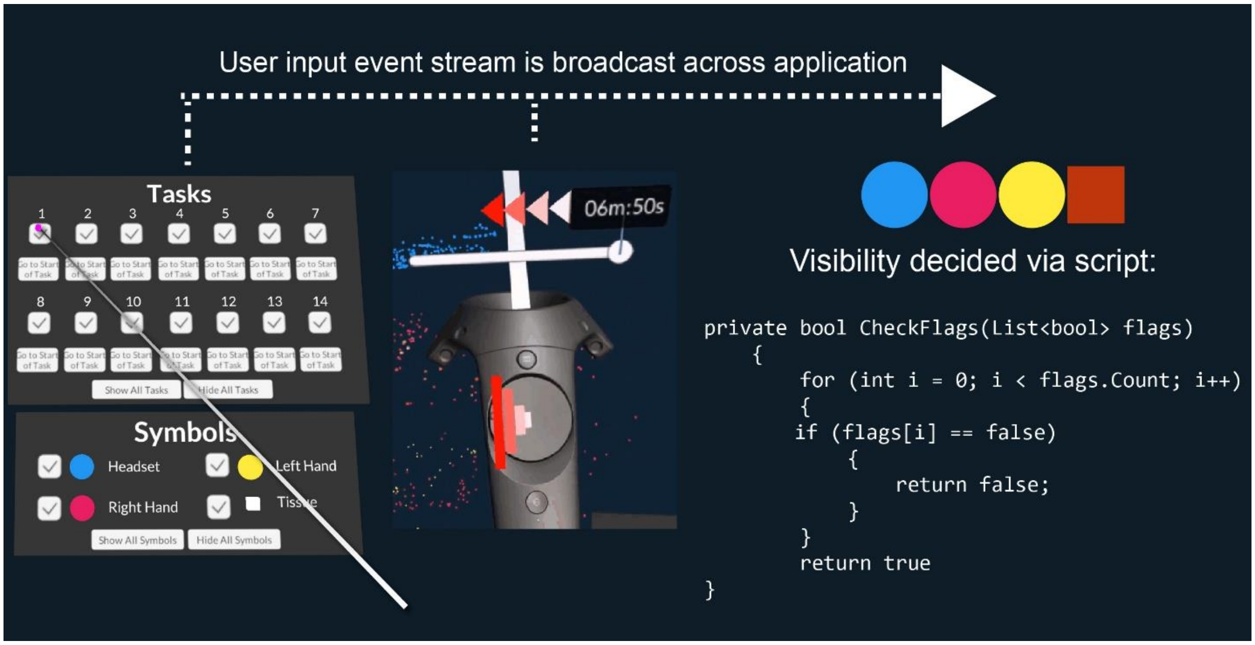

Figure 45. Abstract image of the event listener for each graphic symbol. A series of Boolean values is passed into the CheckFlags() function, for example, whether the currently shown time stamp is later than the associated time stamp of the graphic symbol. If all arguments evaluated to true, the graphic symbol was displayed.

The combination of task, graphic symbol, and time stamp filter formed an interactive system that let the user quickly isolate individual tasks, replay tasks, as well as isolate specific elements of the scene (such as the tissue block), and enabled them to switch between aggregate and focused views in a way that felt natural and immersive. When implementing this interactive system, we used the native C# event system to broadcast any changes to the UI elements to the graphic symbols in the scene. The graphic symbols had behaviors attached to them that then evaluated whether all conditions were met for them to be shown or hidden (see Figure 45). Event systems are highly useful when implementing 3D scenes with many interoperating actors (as is common in video games), and their presence and easy availability in Unity and C# was helpful in making our application responsive, fluid, and interactive.

6.3.4 Mid-Questionnaire

The 2D Desktop setup was not interactive and we thus collected no data about the interaction between the user and the visualization. Instead, we analyzed the user's understanding of the line graph visualization we presented them with through a series of questions. Specifically, we asked them about the visual encoding before prompting them to identify maxima of distance and angular difference in the line graph presented to them. At this point, the line graph contained data from the best-performing user from the control cohort of the experiment. Subsequently, we gave the users more retrieval tasks, e.g., to identify the task with shortest and longest completion times.

For the VR setups, we included one question about the visual encoding of the tissue blocks in the Reflective phase, giving the subjects three options: distance between the tissue and target block, angular difference between the two (correct answer), and completion time.

6.4 Metrics

To assess completion time, position accuracy, rotation accuracy, and satisfaction, we used the same metrics as in our RUI user study (43). In this section, we elaborate on how we assessed space usage for the Ramp-Up and Plateau phases.

Utilizing telemetry data, we wanted to determine if the space usage patterns of users for the control and experiment cohorts during the Plateau phase were significantly different. For this step, we isolated the Plateau phase data for all subjects and then compared between the cohorts. We limited this analysis to subjects in the VR Tabletop and VR Standup setups.

For example, we may find that experiment subjects traveled a significantly shorter overall distance in the Plateau phase after seeing their spread in the Reflective phase. Similarly, they could have discovered that they were covering too wide an area and went on to work more alongside a simple line in the Plateau phase. Exemplary space usage patterns of three control cohort subjects in the Ramp-Up phase are shown in Figure 46. The visualizations in this section were made using Kepler.gl (198).

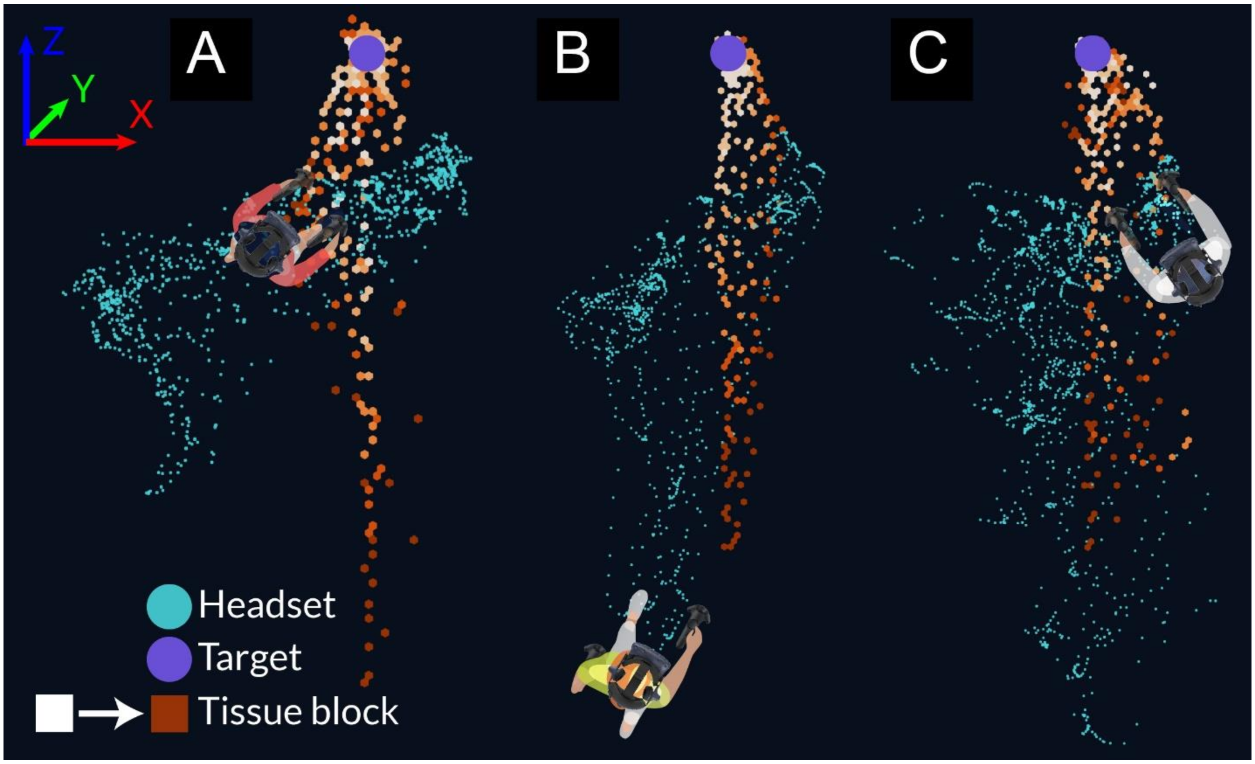

Figure 46. Overhead view (x-z plane) of three participants (A, B, C) with unique movement patterns during the ***Ramp-Up*** *phase.*

As apparent in Figure 46 above, different users display different spatial usage patterns. User works mostly along one well-defined axis in close proximity to the target, starting farther on the left side and working their way towards the target as the start distance between tissue and target block becomes larger, reaching the tissue block by extending their arms. User B, on the other hand, rather than staying close to the target, moves backwards along the z-axis (running between the target and tissue block), thus covering a much wider area with their movements. These movement happened with some regularity in that the user would walk back along the z-axis, turn around, grab the tissue block, and then walk back toward the target block. Finally, user C shows little distinguishable patterns. Notably, as opposed to users A and B, they spent more time on the right side of the axis between the target and tissue blocks overall.

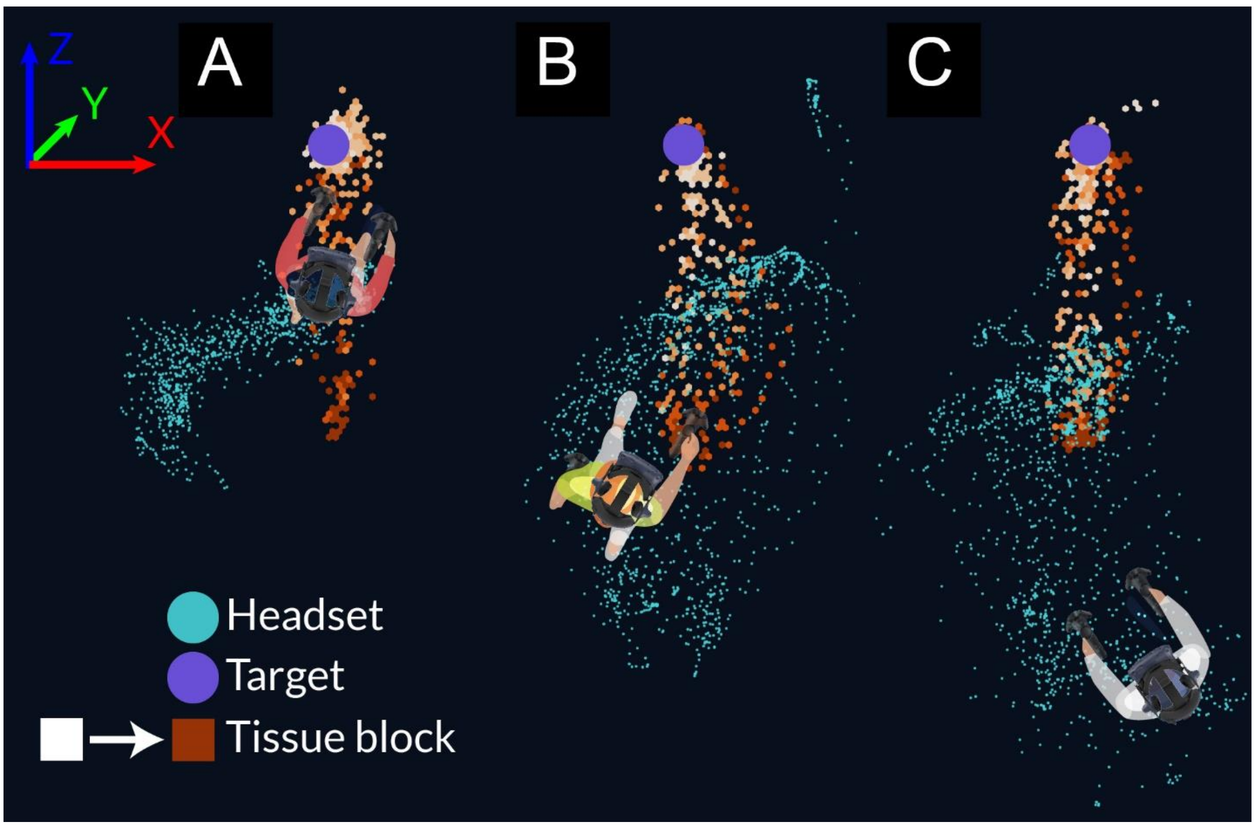

Figure 47. Overhead view (x-z plane) of three participants (A, B, C) with unique movement patterns during the ***Plateau*** *phase.*

The differences in spatial usage patterns become even more apparent for the Plateau phase in Figure 47. During the Plateau phase, these control subjects showed distinct movements. User A, like in the Ramp-Up phase (see Figure 46), limits themselves to a small area close to the target, relying on their arms to reach the tissue block. Users B and C, on the other hand, cover wider areas, frequently positioning themselves "behind" the tissue block when it appears to "carry" it forward into the target. Similarly, for user A, it becomes apparent how they perform all the rotation adjustment very close to the kidney, hence the many white dots just around the target. Users B and C, on the other hand, perform much rotational adjustment before reaching the target.

Figure 48 shows three more users; one of them (user D) was the best-performing subject in that cohort and setup by completion time and accuracy. User D shows a distinct movement profile. Taking up lots of space but also moving frequently with few clusters during the Ramp-Up phase, they go around the kidney for the harder tasks. This becomes even more apparent in the Plateau phase, where they spend the majority "beyond" the kidney. This allowed them to have a clear view of the handles indicating the rotation of the target they had to match with their tissue block. In stark contrast, users E and F seemed almost timid in that they limited themselves to staying in front of the kidney at all times, both during the Ramp-Up and the Plateau phases. The different in movement becomes even more apparent during the Plateau phase, where they stay mostly in one spot and execute what appears to be swinging motions with their arms in order to place the tissue block.

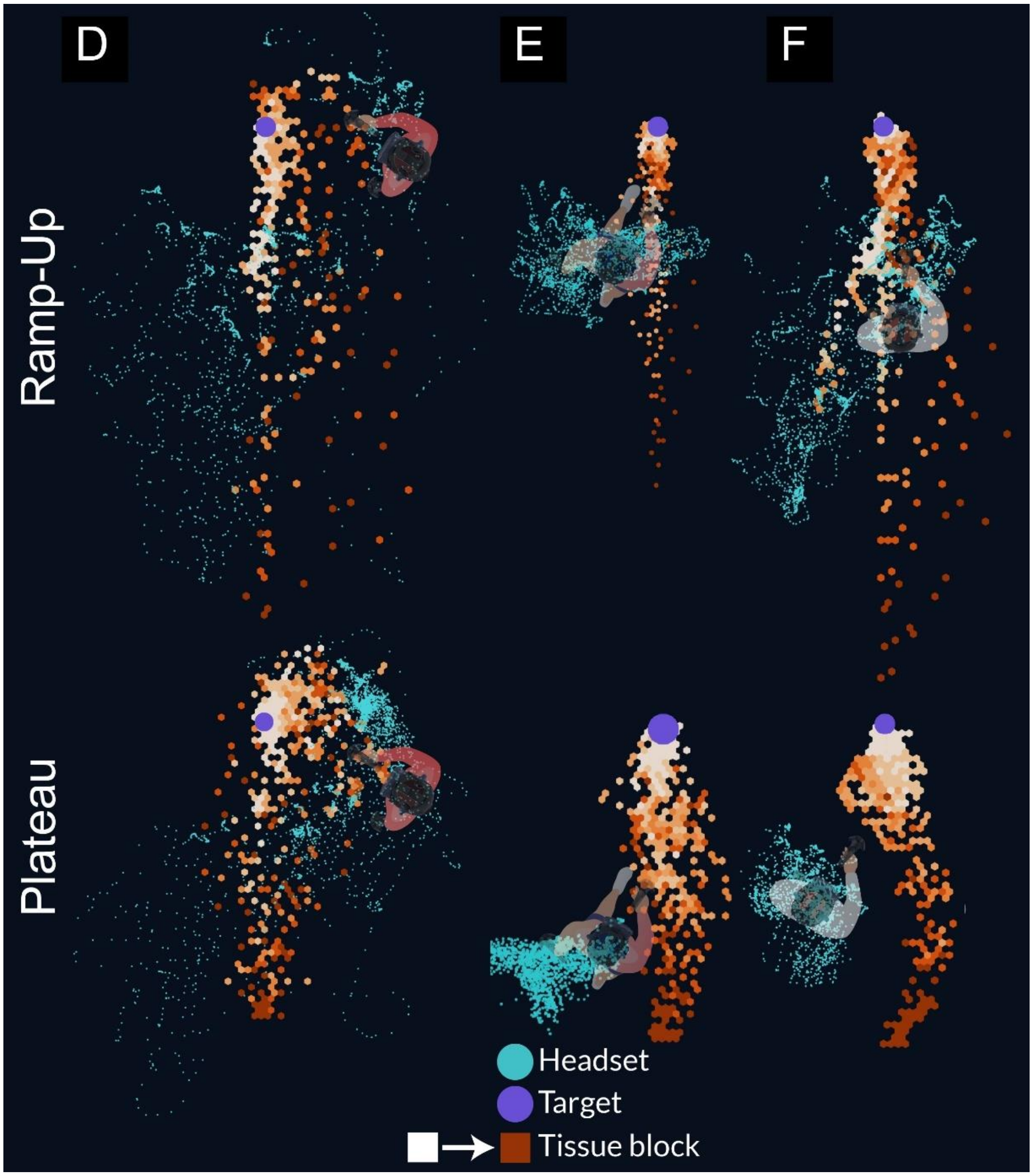

Figure 48. Three more users from the VR Standup control cohort (D, E, F). User D was the best-performing user by completion time and accuracy for that cohort.

To describe the spatial usage by the VR Tabletop and VR Standup users, and to answer RQ2, we analyzed the telemetry presented so far using three metrics.

First, the **convex hull** yielded the contour of all the recorded headset or controller locations in the experiment. Calculating the area of the convex hull allowed us to quickly compare the extent to which different subjects occupied the physical space around them.

Second, the **number of clusters** in a user's telemetry data enabled us to compare whether subjects tended to spend a lot of time within a confined area or whether they tended to spread out into the space. If the number of clusters was low, we could assume that the subject mostly worked in a fixed location, while a larger number of clusters could hint at more agile subjects.

Third, we developed a metric specifically for the analysis of this dataset is what we refer to as the **angle of attack**, see Figure 49 below.

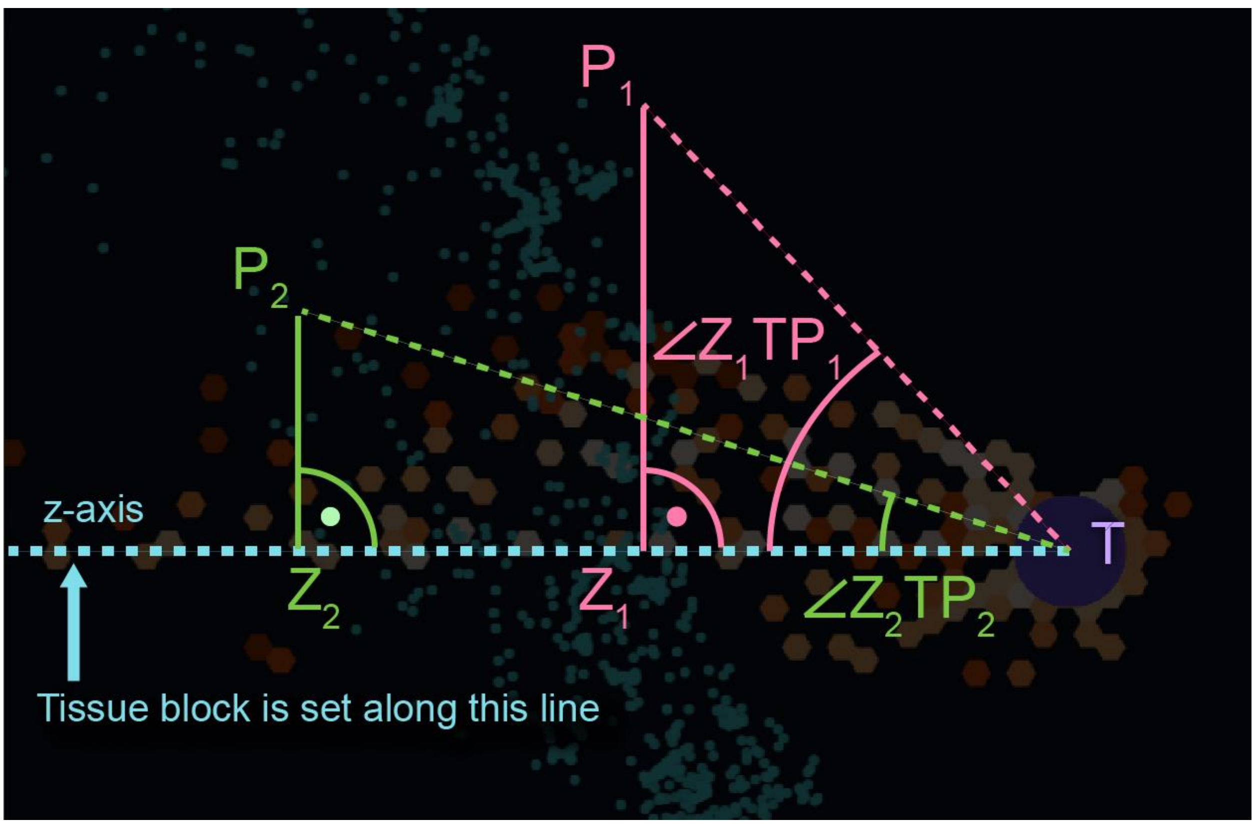

Figure 49. The angle of attack for two exemplary data points P_1, P_2 for one user's head. The blue line is the imaginary line from the target to the start positions of the tissue block (parallel to the global z-axis). We can compute angles $\angle Z_1TP_1$, $\angle Z_2TP_2$, $\angle Z_nTP_n$ *from the (x, z) positions of any point P_n.*

In the background of Figure 49, we see a bird's eye view of a user's headset position during the Ramp-Up phase in the control cohort (darkened, light blue dots). The white-orange hexagons (also darkened) are aggregate hexbins of tissue block positions over time. In the foreground, two exemplary right triangles are shown that share the unchanging location of the target cube as a corner T. Likewise, we can draw right triangles Z_nTP_n for any of the n headset locations. The hypotenuse is formed by TP_n, with TZ_n and the adjacent and Z_nP_n as the opposite. The adjacent lies on the "work axis", i.e., the line on which the tissue block is shifted backwards algorithmically as task complexity is increased. The work axis is parallel to the global z-axis. The angle Z_nTP_n is different for any headset location P. We call this angle the "**angle of attack**" that informs us about the distance at which the user works along the work axis. In the example above, P_1 generates a higher angle of attack than P_2. By calculating the mean of all angles of attack for a given user, we can create a single metric to compare how close individual users were to the main work axis. If we calculate the angle of attack for all recorded headset positions, we can compare between users in identical setups and across cohorts in terms of how far or how close they were to the work axes throughout the study.

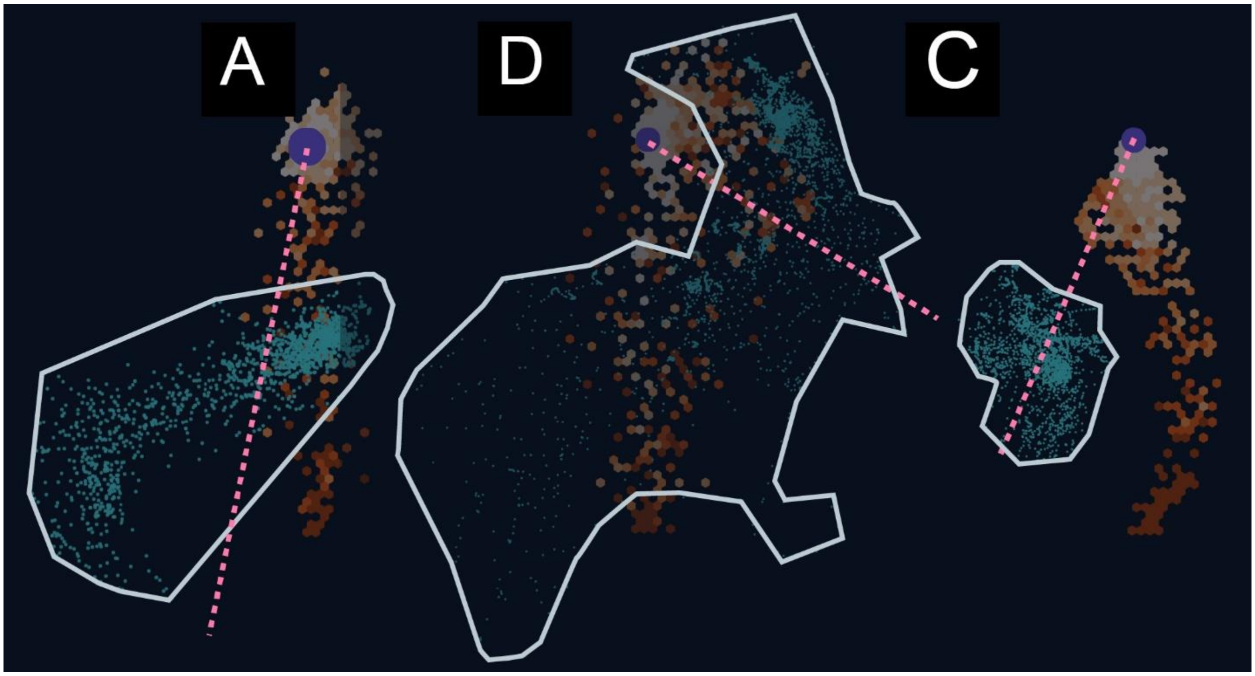

Figure 50. Profiles for three subjects with illustrative convex hull and angle of attack overlay.

Taken together, we can then use the three metrics described above to describe a user's behavior with few numbers. Figure 50 illustrates the average angle of attack, the convex hull, and the number of clusters for user A, D, and C from Figure 47 and Figure 48, demonstrating widely different approaches to the Plateau phase tasks. Both users A and C covers a much smaller area than user D; however, user A has a much smaller average angle of attack than user C. User D, on the other hand, travels around the kidney (which is also mirrored by the many hexbins in that area) and also shows a much large angle of attack.

6.5 Results

To answer the research questions described in Section 6.2, we treated users in this study as the experiment cohort and the users from our RUI user study was the control group. As shown in Figure 31, both cohorts received identical treatments, the only difference being the presence of the Reflective phase for the experiment cohort.

6.5.1 Influence of Reflective phase on performance and satisfaction

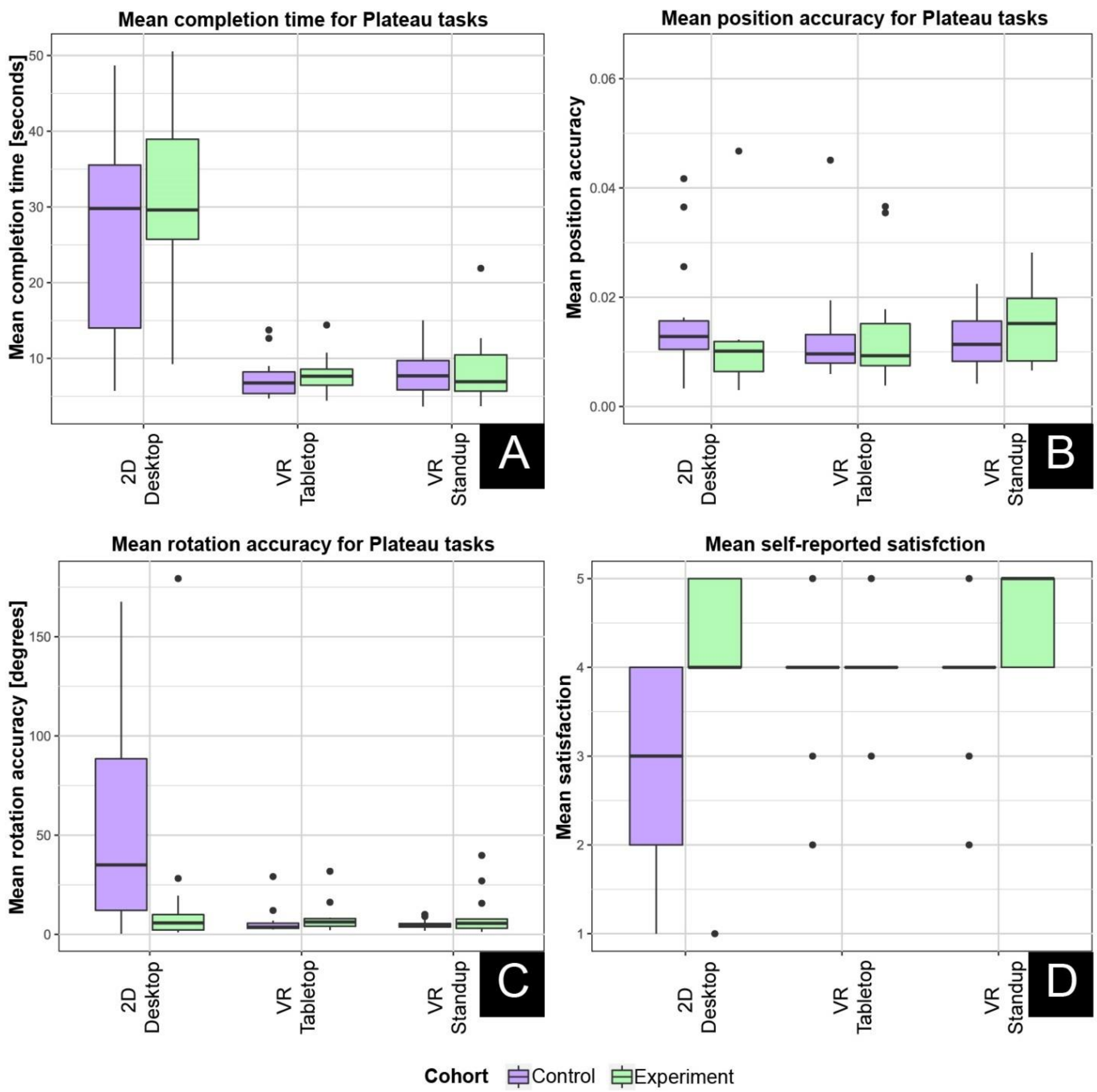

Figure 51. Mean completion time and position/rotation accuracy for Plateau tasks in all setups for both cohorts.

We found **no difference** between the cohorts for completion time, position, and rotation accuracy for the two **VR setups**. However, in the experiment group of the **2D Desktop** setup, we found a **highly significant** higher rotation accuracy (Mann-Whitney-U-Test, $p = 0.031$) as well as a **slightly significant** difference in position accuracy (Mann-Whitney-U-Test, p = 0.085). This means that the Reflective phase (line graph) for the 2D Desktop users did indeed help users outperform the Desktop users in the control cohort. However, we found no difference in terms of completion time for 2D Desktop. Based on these findings, we have to **reject H1**.

Surprised by these findings, we performed further analyses for the users' self-reported **satisfaction**, and found a **significantly higher** feeling of satisfaction in the experiment group of the **2D Desktop** setup (Mann-Whitney-U-Test, $p = 0.04$) and in **VR Standup** (Mann-Whitney-U-Test, $p = 0.016$). In Section 6.5.4, we examine the user behavior in the Reflective phase of the VR Standup setup to determine what factors may have contributed to this higher level of satisfaction with one's performance.

6.5.2 Change in space usage

All the space usage metrics did not differ significantly for the Ramp-Up phase of the control and the experiment group. The results of comparing the Plateau phases of the experiment and the control group are as follows (oriented towards a significance to the 5%-level).

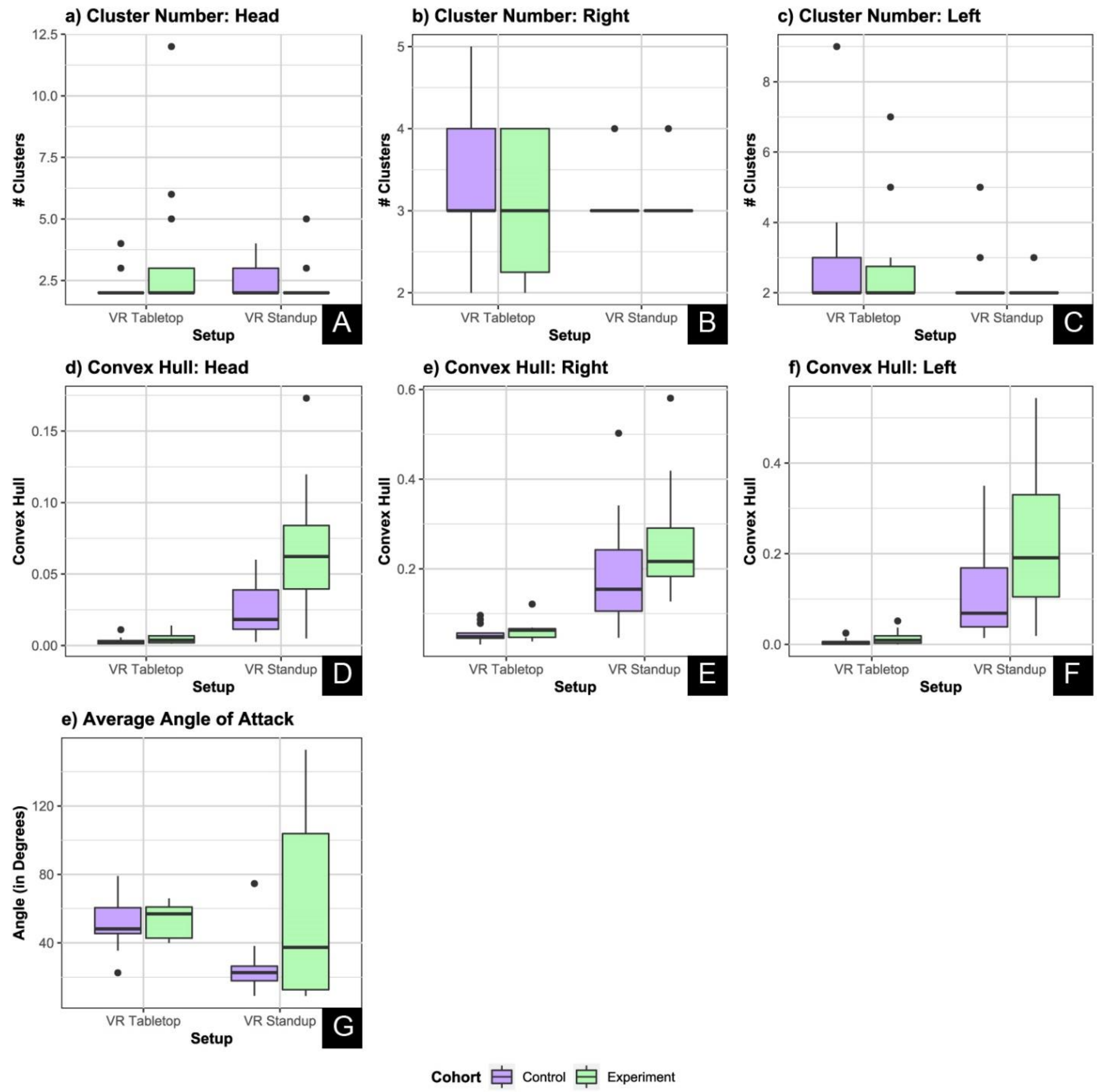

Figure 52. Spatial usage between cohorts for VR Tabletop and VR Standup.

We found **no difference** between the control and experiment cohorts for the **head, right hand, left cluster numbers, convex hull (left hand),** and **average angle of attack for neither VR Standup nor VR Tabletop**. However, it becomes apparent from Figure 52 that there is much greater variance for the experiment cohorts in VR

Tabletop in terms of the number of clusters of the right hand and in terms of average angle of attack in VR Standup.

However, for the **convex hull (head)** and **convex hull (right hand)**, we found **significant differences** for the VR Standup setup (Welch-Test). Specifically, subjects in the experiment cohort formed a larger convex hull with their head and right-hand movements than control cohort users. Presumably, these users, due to spending more time in VR and thus becoming more familiar with virtual space, performed their Plateau tasks with more spatial confidence. Many users who may have been timid at first gained courage. However, we detected neither of these differences for the VR Tabletop setup, where the user's movement was more restricted. We thus have to **reject H2**.

Further, we aimed to find out whether a **change** in space usage behavior (relative to the Ramp-Up phase) has an **effect** on performance in the Plateau phase. We thus calculated a Pearson correlation matrix for the **change** of space usage variables in comparison to the Ramp-Up phase ("**delta_**", see Figure 53**)**.

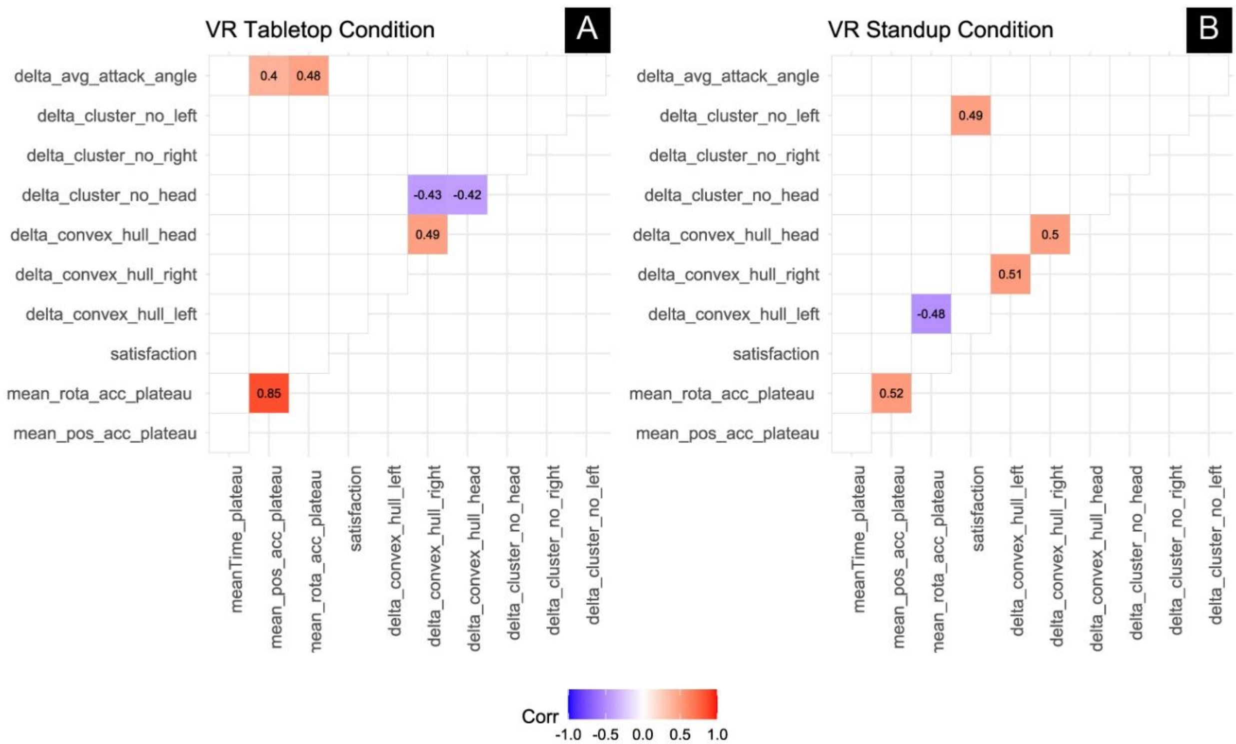

*Figure 53. Pearson correlation matrix between the change in usage variables ("**delta_**") and performance metrics as well as self-reported satisfaction. Only significant correlations are shown; insignificant ones are left blank.*

For the **VR Standup** setup, we see that a change in space usage does not change the Plateau performance in most cases. Only an i**ncreased convex hull (left hand)** is correlated with a **higher rotation accuracy.** This might be an indicator that subjects who kept their left hand less steady and adopted a more natural flow might have been more able to gauge the quality of their tissue block rotation more, but this connection is tenuous at best. Interestingly, an **increased number of clusters (left hand)** is positively correlated with the **feeling of satisfaction**.

This came as somewhat of a surprise to us as we expected that subjects who looked behind the kidney in the Plateau phase would be more able to see the poles indicating the rotation of the target block and would thus save precious time by identifying a suitable vantage point.

For the **VR Tabletop** setup, we even found that an **increase** in the **angle of attack**, i.e., moving their body more in the direction of the kidney, is correlated not only with a lower **position accuracy** but also a lower **rotation accuracy**.

To check the robustness of these results in cases of extreme values, we calculated the point-biserial correlation of change in the performance parameters and a dummy variable of whether users had an increase of the angle of attack of more than 30 degrees or users that moved behind the kidney, i.e., who had an angle of attack greater than 90 degrees. Both show **no significant correlation** with the performance metrics.

These findings point to the different roles of the angle of attack for these two setups. VR Standup users generally had more freedom in their movement through space and could choose their angle of attack without any significant influence on their performance. For VR Tabletop users, on the other hand, it may have made sense to have a setup that forces them to a more acute angle of attack. For example, instead of setting the work axis, i.e., the line between the target and the tissue block at task start, parallel to the edge of the table, a more angular approach may have prevented some users from increasing their angle of attack too much and getting too close to the backside of the kidney.

6.5.3 Tool usage (descriptive statistics)

We recorded a variety of metrics from VR subjects in the experiment group during the Reflective phase. This phase consisted of two parts: an intro where the user explored the best-performing user's dataset (with an audio tutorial about the interactive tools and goals of this phase) and, subsequently, the main part where they explored their own data.

Table 8 gives the number of observations (N), mean, standard deviation (SD), median, min, and max value for a series of metrics from the intro as well as the main part of the Reflective phase for both VR setups together and separate.

Table 8. Summary statistics of tool usage during Reflective phase.

Variable	Setup	N	Mean	SD	Median	Min	Max
reflective2_amountKidneyTurnoff	All	28	1.643	2.483	1.000	0.000	11.000
	VR Tabletop	14	1.643	1.946	1.000	0.000	5.000
	VR Standup	14	1.643	3.003	1.000	0.000	11.000
reflective2_avg_number_tasks_visible	All	28	6.215	3.992	5.333	1.238	14.000
	VR Tabletop	14	6.416	3.957	5.643	1.255	14.000
	VR Standup	14	6.014	4.164	5.149	1.238	14.000
reflective2_avg_task_visible	All	28	7.716	1.259	7.485	5.548	11.542
	VR Tabletop	14	8.092	1.194	7.754	6.925	11.542
	VR Standup	14	7.341	1.251	7.157	5.548	9.643
reflective2_cluster_no_head	All	28	3.393	3.119	2.000	2.000	14.000
	VR Tabletop	14	2.500	1.345	2.000	2.000	7.000
	VR Standup	14	4.286	4.084	2.000	2.000	14.000
reflective2_cluster_no_left	All	28	3.250	2.413	2.000	2.000	13.000
	VR Tabletop	14	3.786	3.093	2.500	2.000	13.000
	VR Standup	14	2.714	1.383	2.000	2.000	7.000
reflective2_cluster_no_right	All	28	3.571	3.259	2.000	2.000	15.000
	VR Tabletop	14	3.786	3.167	3.000	2.000	14.000
	VR Standup	14	3.357	3.455	2.000	2.000	15.000
reflective2_convex_hull_head	All	28	0.195	0.260	0.088	0.003	1.001
	VR Tabletop	14	0.048	0.075	0.025	0.003	0.299
	VR Standup	14	0.342	0.296	0.206	0.028	1.001

reflective2_convex_hull_left	All	28	0.865	1.777	0.353	0.039	9.017
	VR Tabletop	14	0.146	0.100	0.116	0.039	0.368
	VR Standup	14	1.584	2.331	0.760	0.321	9.017
reflective2_convex_hull_right	All	28	0.620	0.591	0.457	0.046	2.330
	VR Tabletop	14	0.180	0.140	0.154	0.046	0.573
	VR Standup	14	1.060	0.536	0.993	0.387	2.330
reflective2_degree_headrotationY	All	28	7760.250	5769.767	6638.196	3367.796	32753.357
	VR Tabletop	14	7086.072	2512.690	6978.933	3367.796	12462.243
	VR Standup	14	8434.427	7864.391	5271.406	3401.142	32753.357
reflective2_distance_lefthand	All	28	40.497	18.613	36.125	11.712	81.941
	VR Tabletop	14	33.540	15.515	28.106	11.712	71.736
	VR Standup	14	47.453	19.355	50.230	18.599	81.941
reflective2_distance_rawslider	All	28	3.915	2.993	3.301	0.022	12.068
	VR Tabletop	14	5.148	3.619	4.868	0.364	12.068
	VR Standup	14	2.681	1.492	2.871	0.022	5.150
reflective2_distance_righthand	All	28	39.860	27.179	34.344	11.246	126.756
	VR Tabletop	14	29.254	19.971	20.814	11.246	74.955
	VR Standup	14	50.466	29.885	45.569	12.534	126.756
reflective2_distance_traveled	All	28	34.421	18.036	28.314	10.083	81.413

	VR Tabletop	14	26.878	12.494	24.651	10.083	57.660
	VR Standup	14	41.964	19.924	38.200	13.635	81.413
reflective2_head_upDownY	All	28	8.155	4.563	6.564	2.620	20.754
	VR Tabletop	14	7.145	4.030	6.068	2.620	17.641
	VR Standup	14	9.165	4.981	6.803	3.890	20.754
reflective2_mean_rawSlider	All	28	0.756	0.141	0.728	0.487	0.983
	VR Tabletop	14	0.787	0.138	0.771	0.584	0.979
	VR Standup	14	0.725	0.143	0.683	0.487	0.983
reflective2_time_toggle_filter_usage	All	28	0.406	0.380	0.326	0.000	0.985
	VR Tabletop	14	0.523	0.412	0.680	0.000	0.985
	VR Standup	14	0.288	0.317	0.187	0.000	0.872
reflective2_time_without_kidney	All	28	0.395	0.376	0.417	0.000	0.962
	VR Tabletop	14	0.469	0.387	0.564	0.000	0.962
	VR Standup	14	0.322	0.364	0.208	0.000	0.903
reflective2_total_time_spent	All	28	430.696	216.311	389.771	195.529	1284.828
	VR Tabletop	14	464.629	149.821	501.839	247.949	696.611
	VR Standup	14	396.763	268.800	358.766	195.529	1284.828

Table 9. Definition of variables for user behavior from the Reflective phase.

Variable	Definition
convex_hull_[INPUT DEVICE]	convex hull of the recorded position coordinates of the left hand, right hand and the head during the Reflective phase
cluster_no_[INPUT DEVICE]	number of clusters of position data as a result of k-means clustering (optimal number of clusters determined by silhouette analysis)
total_time_spent	time spent in the reflection phases
distance_[INPUT DEVICE	cumulated movement of left hand, right hand and head (in meters)
degree_headrotationY	Cumulated total degrees of head rotation around the y-axis
head_upDownY	Cumulated total head movement up and down the y-axis
mean_rawSlider	Average raw slider position ranging from 0 to 1 for each subject
amountKidneyTurnoff	Total number of times the kidney visualization was turned off
time_without_kidney	Share of reflective time spent with kidney turned off
time_toggle_filter_usage	Share of reflective time spent with other filter toggles used
avg_task_visible	Average of task numbers that were visible during reflective phase
avg_number_tasks_visible	Average numbers of tasks that were visible at the same time

All the values reported below have been recorded in the **main** part of the Reflective phase.

On average, users spent **464.63 seconds** (VR Tabletop, *SD* = 149.82 s) and **396.76 s** (VR Standup, *SD* = 268.8 s) in the main part of the Reflective phase. Unsurprisingly, VR Tabletop users traveled less with their headsets (**26.88 meters**, *SD* = 12.49 m) than VR Standup users (**41.96 m**, *SD* = 19.92 m). Notice, however, the large **range 114.22 m** between the subject with the **most right-hand movement** and the subject with the **least** (both in the VR Standup setup)**.** Likewise, the subject with the most **distance traveled for the headset** in the VR Standup setup was measured at **81.41 m** (vs. **13.64 m** for the least traveled), yielding a **range of 67.77 m**.

The bigger freedom of movement probably also to VR Standup users rotating their heads more, with **8434.43 degrees** vs. **7086.07 degrees** for VR Tabletop, equaling around 26 and 20 theoretical complete head rotations, respectively. Further, on average, VR Tabletop users spent **46.9%** of their time in the main part of the Reflective phase without the kidney, compared to **32.2%** for VR Standup users, prompting us to **confirm H3c** (users will spend the majority of time with the kidney turned on). Because VR Tabletop users could rotate the kidney when completing their tasks, kidney rotations were shown in the Reflective phase, which is users may have had more of an incentive to leave the kidney visible. The average number of tasks simultaneously visible was **6.42** (VR Tabletop, *SD* = 3.96) and **6.01** (VR Standup, *SD* = 4.16). Finally, with regards to time slider usage: VR Tabletop users moved the slider **almost twice as much on average** as VR Tabletop users. Specifically, VR Tabletop users scrolled through **5.15 times** the time span of their dataset, compared to **2.68**

times for VR Standup. Since both of these values are far off from the 10 times we predicted in **H3a**, we need to **reject H3a**.

Likewise, the mean position of the raw slider on a scale from 0 (first time stamp, beginning of the dataset) to 1 (last time stamp, end of the dataset) was similar for both setups at **0.79** for VR Tabletop and **0.73** for VR Standup users, requiring us to **reject H3b** (predicting that the most selected location for the slider would be towards the very end of the dataset).

In Section 6.5.4, we investigate the influence of metrics during the Reflective phase on performance in the Plateau phase.

6.5.4 Metrics during Reflective phase and influence on Plateau phase performance

We further wanted to understand the relationship between metrics for user behavior as well as interactive tool usage and performance in the Plateau phase and self-reported satisfaction. As discussed in Section 6.5.1, we detected **no significant difference** between control and experiment cohort for the two VR setups in terms of their Plateau performance. While for RQ1, we aimed to detect if there was a difference in Plateau phase performance, in this section, we focus on whether there are **measurable VR behavior traits** in the Reflective phase that **exert an effect** on the performance in the Plateau phase in any way. This could help us pinpoint what specific elements of the Reflective phase could be adjusted to improve user performance in subsequent studies or during the development of a Reflective phase for real-world VR training. We discuss **effects** (see Table 10**)** between metrics during the Reflective and Plateau phases for the experiment cohort to understand how behavior in the Reflective phase influences performance in the Plateau phase in order to answer **RQ4**.

Table 10. Regression table for relationship between tool usage as well as behavior in Reflective phase and performance in Plateau phase. Significant effects are highlighted yellow. Standard errors are given in italics below each effect size.

Variable	Time	Normalized Position Accuracy (Distance)	Rotation Accuracy (Angular Difference)	Satisfaction
convex_hull_left	-0.00426	0.00087	2.23984***	-0.00218
	0.1024	*0.0009*	*0.6966*	*0.0551*
convex_hull_right	-0.10238	0.00079	2.68827**	0.06548
	0.3129	*0.002*	*1.3039*	*0.1071*
convex_hull_head	0.29969	0.00104	1.23162	0.1835
	0.5822	*0.0053*	*1.5989*	*0.4664*
cluster_no_left	-0.07976	-0.00023**	-0.30223	0.03488***
	0.3704	*0.0001*	*0.7223*	*0.0027*
cluster_no_right	-0.28293*	0.00074**	0.43080***	0.03476
	0.1644	*0.0003*	*0.029*	*0.0206*
cluster_no_head	-0.01434	0.00078*	-0.33586	0.08262***
	0.1829	*0.0004*	*1.3199*	*0.0249*
total_time_spent	-0.00285***	0.00001	0.00104	-0.00003***
	0.0006	*0*	*0.0039*	*0*
distance_lefthand	0.00745	-0.00002*	0.01688	0.0014
	0.0083	*0*	*0.0203*	*0.0016*

distance_righthand	0.00537	-0.00003	0.00509	0.00242
	0.014	*0*	*0.0064*	*0.0022*
distance_traveled	0.00845	-0.00004**	-0.01332	0.0019
	0.0074	*0*	*0.0111*	*0.0037*
degree_headrotationY	-0.00007	0	-0.00010*	0.00001***
	0.0001	*0*	*0.0001*	*0*
head_upDownY	0.04830*	-0.00010***	-0.13422	0.01787*
	0.0242	*0*	*0.101*	*0.0101*
mean_rawSlider	9.00896***	-0.02615***	1.65585	-1.75738*
	0.5026	*0.007*	*18.2541*	*1.0011*
distance_rawslider	-0.13367**	0.00032	-0.3294	-0.02021**
	0.0621	*0.0004*	*0.5474*	*0.0098*
amountKidneyTurnoff	-0.12624	0.00009	0.53339	0.01049
	0.3022	*0.0005*	*0.5338*	*0.0346*
time_without_kidney	-4.41309	0.00876**	3.53305*	-0.63312***
	4.3396	*0.0033*	*2.0591*	*0.0953*
time_toggle_filter_usage	-1.31098**	0.00935***	2.05651	-1.08556***
	0.6081	*0.0013*	*2.3138*	*0.2277*
avg_task_visible	1.01021	0.00107	0.14262	-0.20498***
	1.7188	*0.0014*	*0.4511*	*0.0333*

avg_number_tasks_visible	0.062	-0.00003	0.30905	-0.02092
	0.1001	*0.0001*	*0.3333*	*0.046*

***significant at the 1% level, **significant at the 5% level, *significant at the 10% level

First, we find that spending **more time** in the Reflective phase (*total_time_spent*) has a **significant negative effect** on **task completion time** in the Plateau phase (users become faster), without jeopardizing **position** or **rotation accuracy**. However, it has a **negative effect** on the feeling of **satisfaction**. This might be, because the Reflective phase presents an analytical mode as opposed to the more task-focused and almost playful modes like the Ramp-Up and Plateau phases where users actually get to move tissue blocks and interact with virtual objects in a more natural way. Additionally, many users, when confronted with their own data, may have found their performance in the Ramp-Up phase to have been lacking.

Second, not seeing the **base map**, i.e., the kidney, for extended periods of time in the Reflective phase (*time_without_kidney*) has a **significant negative** effect on **satisfaction** and a **positive effect** with both distance (position) and angular difference (rotation) in the Plateau phase, resulting in lower **position** and **rotation accuracy.** This is a wholly undesirable. Seeing one's data without the proper context seems to be a major issue not only for performance but also enjoyment of the entire VR experience. The integrity of the base map or reference system thus seems conducive to accuracy and user satisfaction for this experiment.

Third, viewing **later, more complex tasks** (*avg_task_visible*, likely with data more spread out due to the larger distance between the tissue block and the target block) also has a **negative effect** on satisfaction, likely leading to **confusion** for many users through **display clutter**. Likewise, spending more time analyzing **later tasks** in the Reflective phase indicated by the time slider (*mean_raw_slider*, likely with previous tasks visible) has a significant **negative** effect on **centroid inaccuracy**, but a **positive** effect on **task completion time** and a **less significantly negative** effect on **satisfaction**. It may thus be useful to not show aggregated data by default. In the Reflective phase, the user saw started with the time slider at 100% (last time stamp, end of the dataset), with all tasks and graphic symbols visible. Slider usage (*distance_rawslider*) was shown to have a **negative** effect on **completion time** but also on **satisfaction**.

Fourth, we identified metrics that enhanced both **accuracy** and **satisfaction** metrics. The **total number of degrees of head rotation around the y-axis** (*degree_headrotationY*) had a negative effect on angular difference (**rotation accuracy)** but **no effect** on distance **(position accuracy)**, as well as **higher satisfaction** without influence on completion time, prompting us to **reject both H4c and H4d**. Thus, encouraging users to look around in their environment seems to be a good design guideline for similar interfaces, and the ability to see data with depth using head movements is a key element of VR. Additionally, the **number of clusters** of the **left hand** (*cluster_no_left*) had a **positive** effect on **satisfaction** and a **negative** effect of **distance** (higher position accuracy). Similarly, the **head movement up and down the y-axis** (*head_up_DownY*, i.e., exploring the data from different vertical viewpoints) had a **positive** effect on **satisfaction** and a **negative** effect on **distance** (higher position accuracy) but also led to led to an **increase** in **task completion time**. We have to **reject H4a**, because while the **total distance** traveled by the **head** as well as the **left**

hand was **negatively** correlated with **distance** (**confirming H4b**), it had no effect on completion time.

6.5.5 Mid-questionnaire and correlation with performance in Plateau phase

Finally, in RQ5, we were interested to check the relationship between the task scores in the mid-questionnaire, which we presented to users between intro and the main parts of the Reflective phase, and performance metrics in the Plateau phase (for 2D Desktop users only). We had hypothesized that there would be significant negative correlations between task score and position accuracy in terms of distance between the tissue and target blocks (**H5a**), error and bias (**H5b**), and rotation accuracy in terms of angular difference (**H5c**). Further, we predicted that there was no significant correlation between the task score and the completion time (**H5d**). Lastly, we assumed that the majority of users would agree or strongly agree that the subject shown to them in the intro part of the Reflective phase was highly fast and accurate (**H5e**). Figure 54 shows the distribution of right and wrong answers to questions in the mid-questionnaire as well as a stacked bar graph of our users' assessment of the performance of the best subject in their respective setup shown to them during the intro part of the Reflective phase.

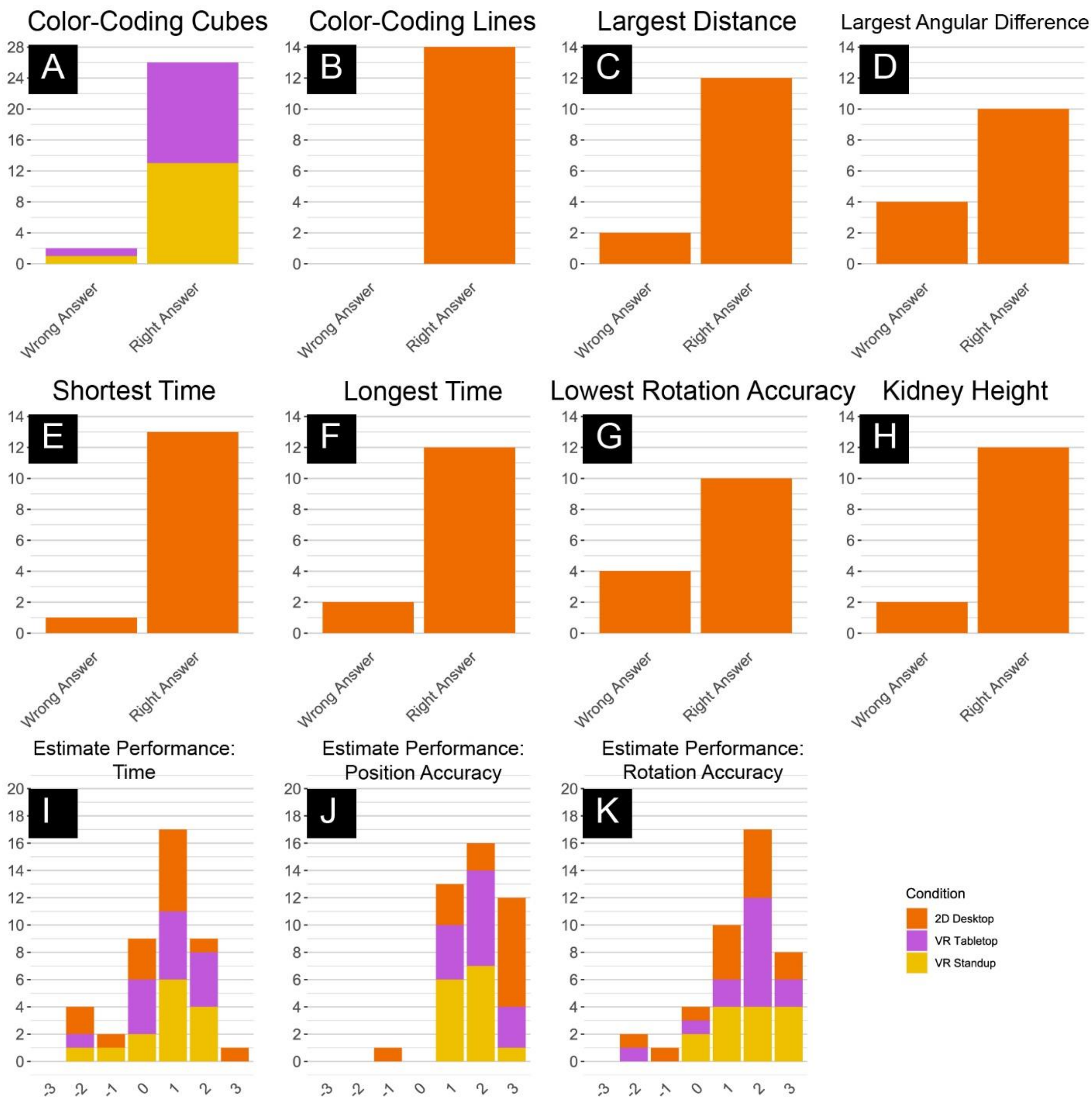

Figure 54. Bar charts showing the distribution of right and wrong answers for the mid-questionnaire.

As becomes apparent from Figure 54, VR Tabletop and VR Standup subjects, with one exception in each setup, were all able to determine that the color of the tissue block in the Reflective phase encoded the rotational difference between the target and tissue blocks. Similarly, all 2D Desktop users correctly stated that the lines in the line graph indicated the distance and rotational difference between the blocks. The majority of 2D Desktop users also answered all the following questions correctly; we found the maximum number of wrong answers (four each) when we asked them to indicate a) the point in the line graph with the maximum rotation difference, and b) the task with the highest rotational difference at task submission. Finally, all subjects were presented with a 7-point Likert scale (from -3 to +3) for their level of agreement about whether the data shown to them constituted a high performance for completion time, position accuracy, and rotation accuracy (see bottom row in the stacked bar graph in Figure 54). Notably, the average level of agreement with these statements was much higher for position ($mean_{2D\ Desktop}$ = 2.1428, $mean_{VR\ Tabletop}$ = 1.9285, $mean_{VR\ Standup}$ = 1.6428) and rotation accuracy ($mean_{2D\ Desktop}$ = 1.2142, $mean_{VR\ Tabletop}$ = 1.5714, $mean_{VR\ Standup}$ = 1.7142). This was probably due to the explicit visual encoding of these two metrics in the data, even more so for 2D Desktop users with their explicit line graphs (see Section 6.3.2). On the other hand, users were less convinced of the quality of the user's performance with regards to completion time ($mean_{2D\ Desktop}$ = 0.4285, $mean_{VR\ Tabletop}$ = 0.7857, $mean_{VR\ Standup}$ = 0.7857). Notably, completion time was never explicit encoded in the data, especially for VR users, who had to estimate the task completion time by looking at time stamp intervals between tasks. Likewise, 2D Desktop users could gauge the completion time by looking at the distance between the vertical dot-dash lines that encoded the end of one task and the beginning of the next, but this kind of visual encoding was less explicit than the lines indicating position and rotation accuracy.

Finally, to test our hypotheses, we checked for correlations between task scores and performance in the Plateau phase (see Table 11).

Table 11. Correlation check for RQ5 hypotheses.

Hypotheses	PerformanceVar	Correlation
H5a	Normalized Mean Position Deviation	-0.74 ***
H5b(1)	X-Error	-0.38
H5b(2)	Y-Error	-0.45
H5b(3)	Z-Error	-0.22
H5b(4)	Bias	-0.6 **
H5c	Mean Rotation Deviation	0.08
H5d	Mean Time	0.15

***significant at the 10% level, **significant at the 5% level, *significant at the 1% level

We found significant correlations between the task score and position accuracy and bias, allowing us to confirm **H5a** and **H5b(4)**. **H5b(1-3)** have to be **rejected**, because no correlation exists between test results and any error specific to x, y, or z-axis. Further, because there is no correlation between task score and rotation accuracy, we have to **reject H5c**. Additionally, we **accept H5d**, because there is indeed no correlation between task score and mean completion time. Finally, we also **confirm H5e** (most users accurately state that the data shown to them in their intro part of the Reflective phase comes from a high-performing subject. VR subjects correctly answered that the data shown to them belonged to a high performer in larger numbers while the answers of 2D Desktop subjects were more mixed. Additionally, the high performance in the position and rotation accuracy measures was more easily identified by our users than the completion time (see Figure 54).

7. User Study 3: Improving Completion Time, Memory, and Satisfaction for Traversing Virtual Buildings Using VR Data Visualizations

Following the RUI VR Reflective phase study, we designed a second experiment involving the navigation of virtual buildings. We took a model of Luddy Hall, the home of the School of Informatics, Computing, and Engineering at Indiana University in Bloomington, IN, USA. The model was designed by Philip Beesley Architect Inc. (http://www.philipbeesleyarchitect.com/). Since this model was built as a scaffold for a public art piece to be installed by the architect's studio (https://cns.iu.edu/amatria.html), it came in two parts: a simpler version of the entire building where most structures were just hinted at, without any materials; and a more detailed version of just the atrium of Luddy Hall. For the entire experiment, the user spent time only in the atrium of Luddy Hall. Detailed screenshots are shown in Figure 55.

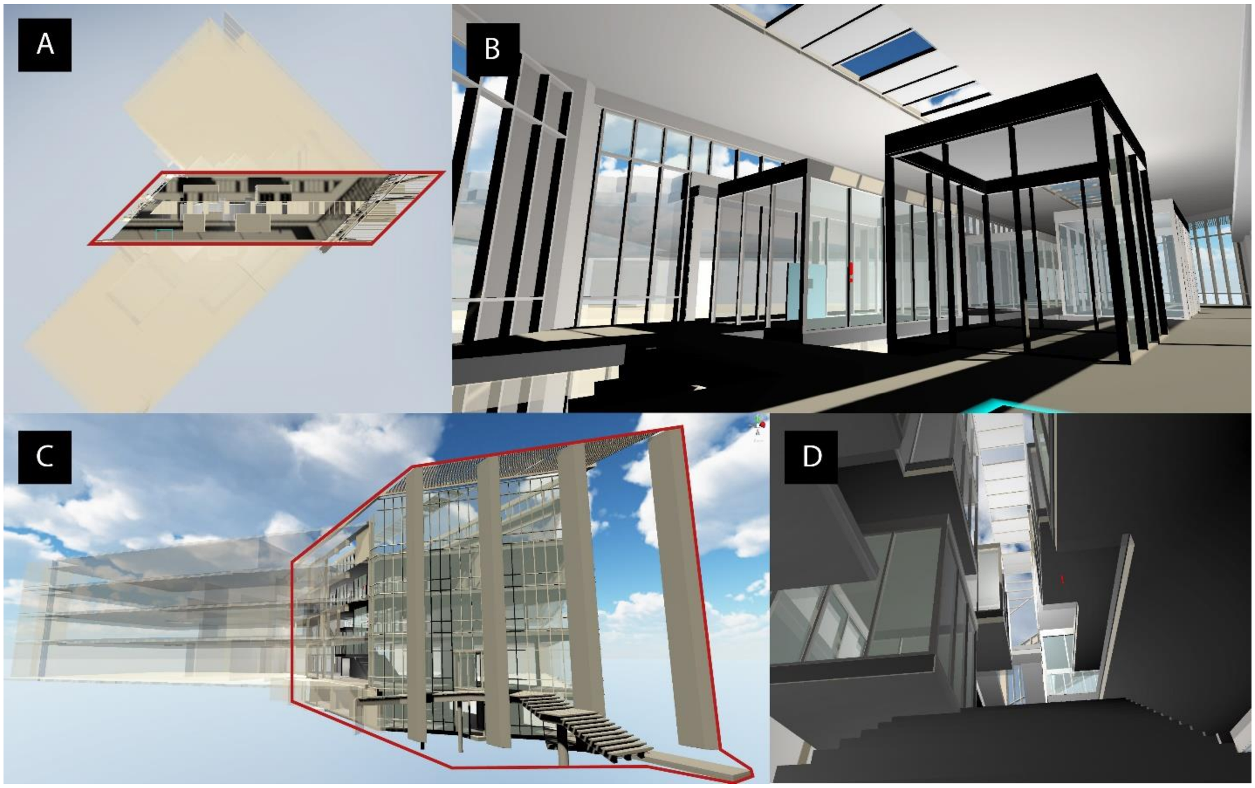

Figure 55. Exterior and interior shots of the Luddy Hall model. ***A****: top view with atrium highlighted orange.* ***B****: 4th floor near star case, start point for all navigation tasks.* ***C****: side view, atrium highlighted.* ***D****: view from first floor up the stair case.*

The goal of this study was to identify whether we could observe performance improvements between a control and experiment cohort. Both cohorts performed a series of navigation tasks using a VR HMD and controllers for two sets of 24 tasks (including tutorials) with the option of three different navigation methods, with a break in-between (control) and a Reflective phase to inspect their performance from the first trial in order to formulate strategies for improvement in trial 2. We implemented three common navigation choices in VR (walking, teleporting, free-flying), which we explain in more detail in Section 7.4.

7.1 Research questions and hypotheses

For our study, we aimed to answer the following research questions and provided the following hypotheses.

RQ1: Is there a difference in completion time between the control and experiment cohorts during trial 2?

H1: The experiment cohort achieves significantly lower completion times than the control cohort during VR Trial 2.

RQ2: Is there a difference in the rate of change in completion time from trial 1 to trial 2 between the two cohorts? That is, when computing the differences in completion time per trial and per subject, and then compare these values between the cohorts, is there a significant difference?

H2: The experiment cohort achieves significantly larger changes in completion times between trial 1 and trial 2.

RQ3: When asked questions about the tasks and the virtual building after taking a break (control) and completing their Reflective phase (experiment), is there a difference in score between the two cohorts?

H3: The experiment cohort achieves higher scores in the mid-questionnaire than the control cohort.

RQ4: What are the preferred choices of navigation methods during the last round of tasks?

H4a: Subjects prefer teleporting when finalizing a task within sight of the start position.

H4b: Subjects prefer free-flying when finalizing a task out of sight of the start position.

H4c: Subjects prefer walking just as they finalize a task.

RQ5: Is there a difference in self-reported satisfaction between the two cohorts at the end of the experiment?

H5: There will be no significant difference in satisfaction between the cohorts.

7.2 Study design

We developed our research design for two cohorts (control and experiment), which, for the most part, performed identical steps (see Figure 56). The major difference was the addition of a Reflective phase for the experiment cohort. The study was designed to last approximately 30 to 45 minutes per subject.

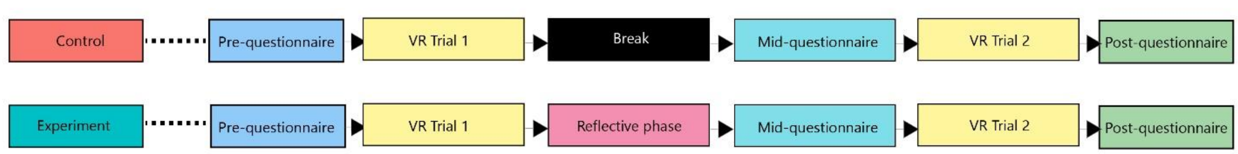

Figure 56. Luddy VR study design. Both cohorts repeat the same set of VR tasks twice ("VR Trial 1" and "VR Trial 2"). Control users take a break before answering the mid-questionnaire; experiment users get a Reflective phase.

When arriving at the research site in Luddy Hall, the subject was asked to sit down at a table with a laptop running a survey. The survey began with a study information sheet before presenting a pre-questionnaire to obtain information about the subject's demographic background as well as prior experience with VR, video games, 3D applications in general, data visualizations, and their familiarity with Luddy Hall. Following that, each subject put on an HTC Vive HMD and took two VR controllers. During the following VR Trial 1, they performed a total of 24 tasks in four rounds.

Following that, the control group took a break from the study. The research facilitator encouraged them to stand up and walk around the research area. The experiment group, on the other hand, stayed in VR and was presented with a Reflective phase, where they saw their own data visualized as 3D trajectories across a miniature version of the building. We describe this in more detail in Section 7.5.

After the break (control) and the Reflective phase (experiment), subjects from both cohorts sat down at the laptop again to fill out a mid-questionnaire, where we asked them how many tasks they had completed in total, how many floors the building had, and more questions. The mid-questionnaire is discussed in Section 7.6.

Following the mid-questionnaire, all subjects donned the VR gear again for VR Trial 2, where they repeated the same 24 tasks from VR Trial 1, but without any audio tutorials. We still had them perform the tutorial tasks but excluded them from data analysis, see Section 7.9.

Finally, all subjects completed a post-questionnaire, where we asked them to rate their own performance, state their preference for the navigation tasks, and indicate their satisfaction with their performance. After successful completion of all parts of the study, each subject was remunerated with a $20 gift card.

7.3 Task difficulty

At the core of this study were two sets of 24 navigation tasks in VR. In this section, we describe what the individual task entailed, where they were located, and what was required from the user. An overview of all 24 navigation tasks can be found in Figure 57 (B). During each of the first three rounds, only one navigation method was possible, starting with walking, then going to teleporting, and ending with free-flying. In the fourth round, the subject could choose which navigation method they wanted to use, and they could change it at any time. The first task in every round served as a mini-tutorial where a pre-recorded voice explained the scope and goal of the experiment as well as the controls of the currently active navigation method to them.

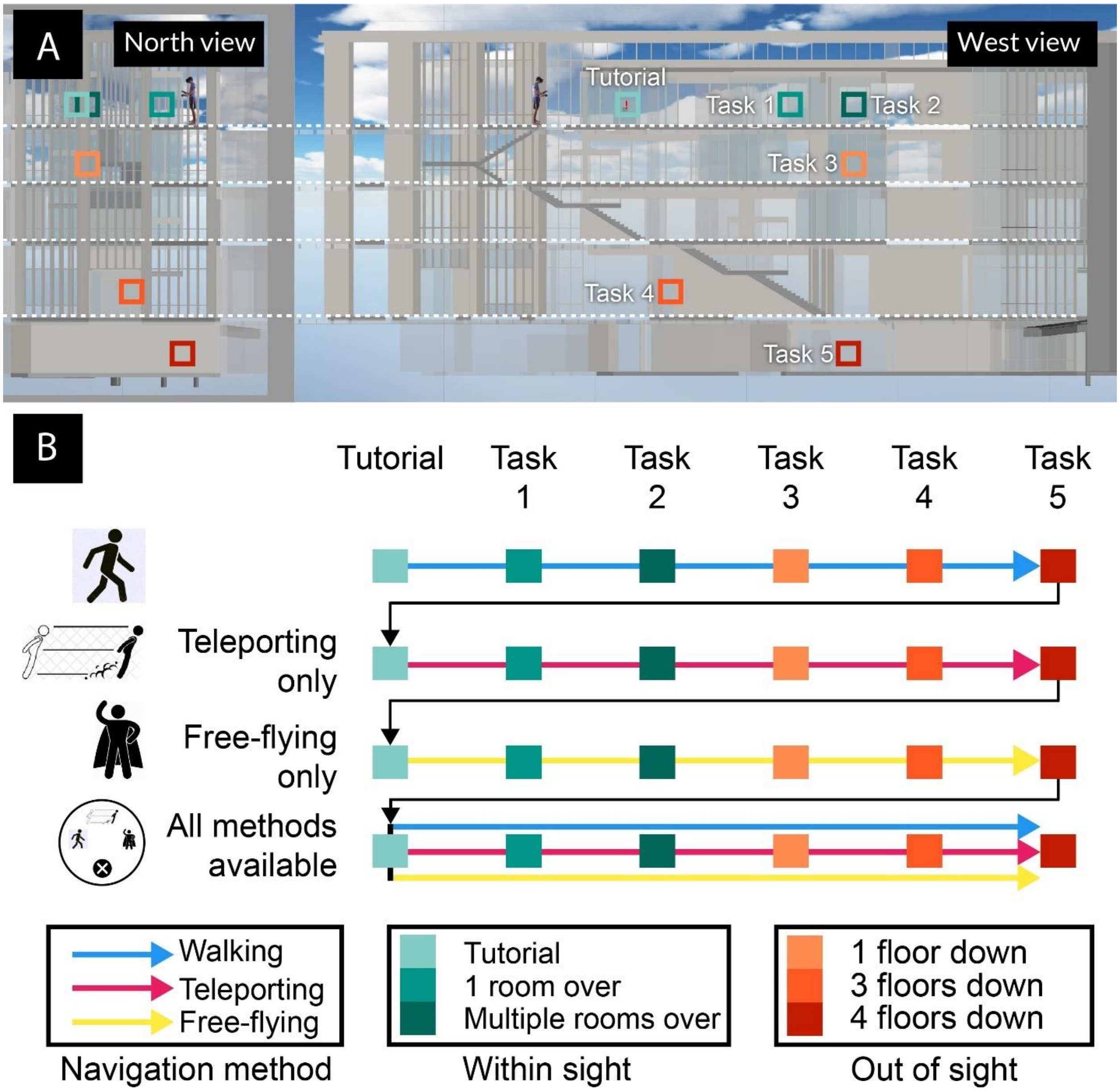

Figure 57. This distribution of tasks across Luddy Hall and task sequence. ***A****: the five navigation tasks (plus tutorial) at their locations in two aligned cross-section views of the building.* ***B****: all 24 tasks in sequence with color-coded difficulty level and possible navigation methods.*

Luddy Hall is an angular building with a central staircase leading through an atrium.

A variety of study rooms are built into the atrium, adjacent to the stair case. This setup makes Luddy Hall a rather challenging environment for virtual (and physical) navigation.

All 24 tasks entailed navigating from a fixed start position (see the human figure in Figure 57, A) to a fixed target position. The start position was identical across all tasks, and the target position for each task did not change between rounds. The start position was on the fourth floor at the top of a staircase leading through the atrium (see Figure 55 above). The target positions for tasks 1 and 2 were on the same floor and within sight, in two study rooms. Tasks 3-5 were on increasingly lower levels, with task 3 being inside another study room, task 4 being in a small office space under the staircase on the ground floor, and task 5 being in a classroom in the lower levels of the building. We used the increasing distance between the start position and a task target position as well as whether a target was within sight or out of sight to increase the difficulty over time.

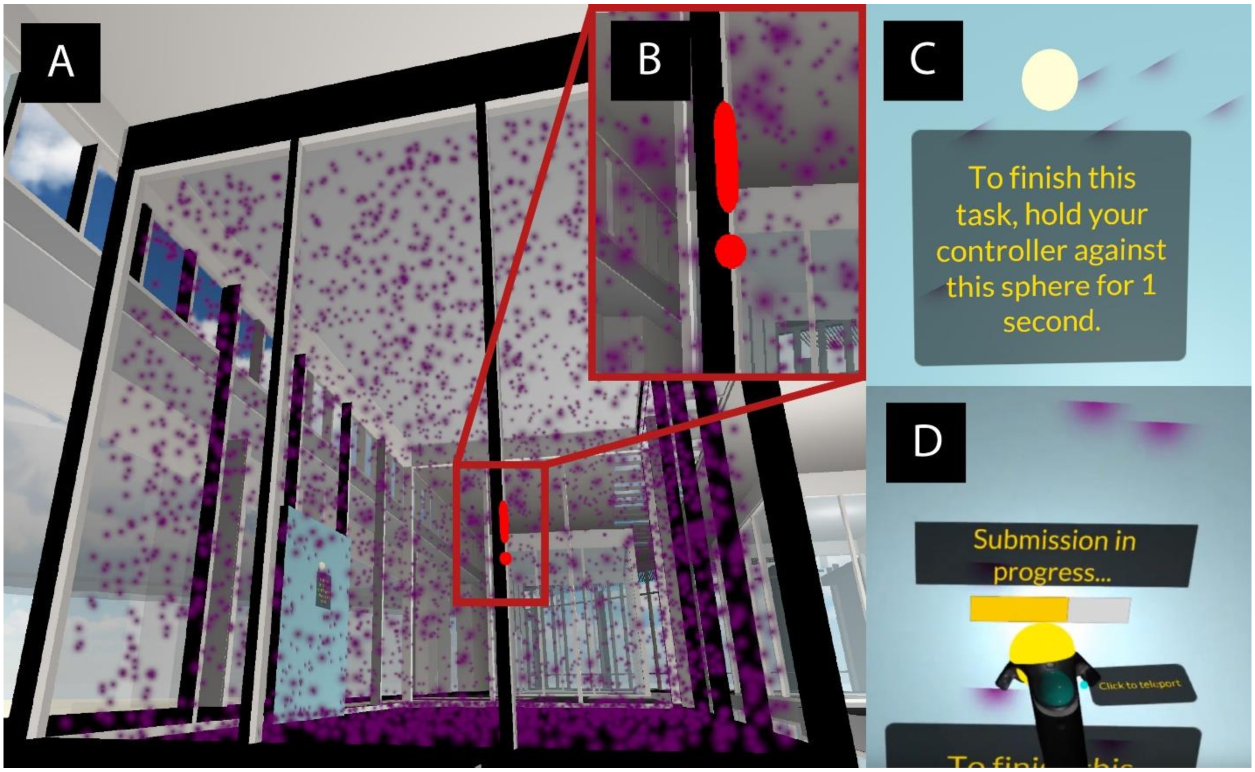

Figure 58. Anatomy of a task room. ***A****: the tutorial task room from the outside.* ***B****: note that the red exclamation mark is rendered on top of the wall in front of it.* ***C****: task submission instructions.* ***D****: the submission in progress.*

Inside each task room, there was a flat, blue panel with a white sphere inside it, with a diameter of about 10 cm. The center of the sphere was about 120 cm above the floor. Below the sphere, a text panel instructed the user to finish the task by holding their controller against the sphere for one second. When the user arrived at a task, they had to locate the panel, approach it, interact with the sphere; subsequently, they were transported back to the start position. A fade effect with a duration of one second smooth the transition between standing in the task room and being back at the start position.

As has become evident, the six different task rooms were distributed all across the building. In order to indicate to the user where to navigate next, a red exclamation mark was placed as a waypoint inside each task room. We used a custom shader to ensure that the waypoint was rendered on top of any other surfaces in the scene (see Figure 58, B), effectively imitating best-practice navigation aids from video games, especially first- and third-person ones. To help the user locate the task room once in sight, we added a flurry of around 4000 little, twinkling, purple particles to each room.

7.4 Navigation methods

For this study, we aimed to determine the usefulness of three navigation methods in virtual environments. Before we describe the movement, controls, and features for all three methods (and the navigation choice mechanism in round four of VR Trials 1 and 2), we briefly describe how movement in this user study is generated.

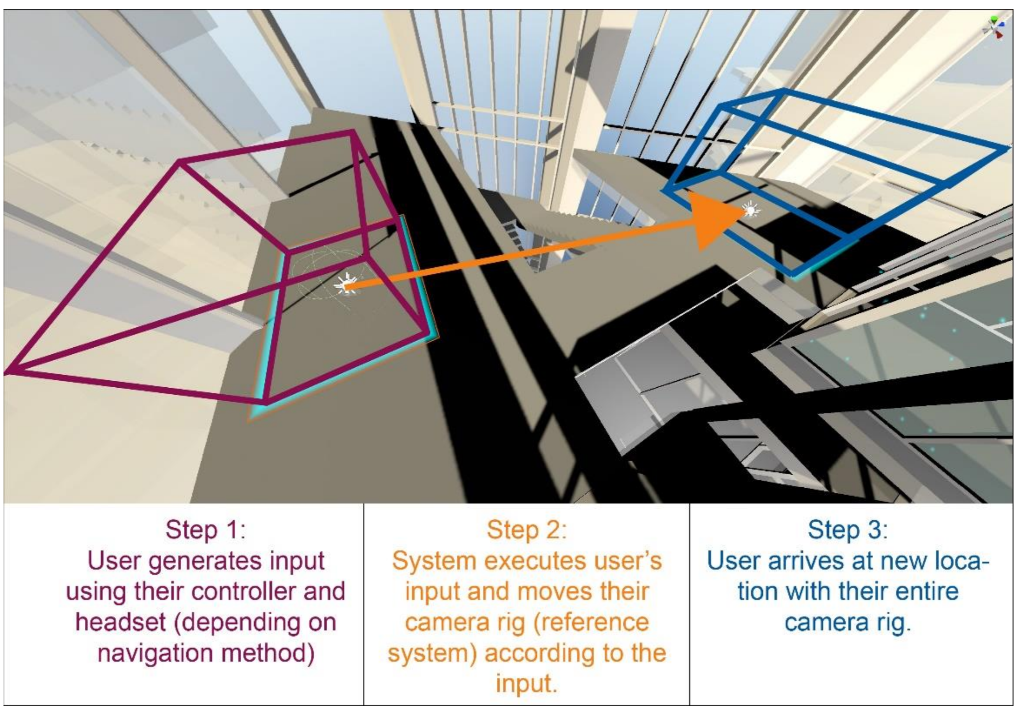

Figure 59. Illustration of basic navigation.

In Figure 59, we show the three basic steps at the basis of every navigation action. The user provided input via their right controller and their HMD (Step 1). Our custom navigation scripts them took the user input and translated the camera rig with the user inside accordingly (Step 2), leaving the user at the new location (Step 3). We describe controls and features for each method in Figure 60, and the calculations for speed and direction of the translation in detail in Figure 61.

In addition to moving their entire camera rig with them (see cuboid shape in Figure 59), they could also move around physically while the camera rig stayed in place. This, however, only allowed them to cover small distances, limited by the size of the play space, see Section 7.8.

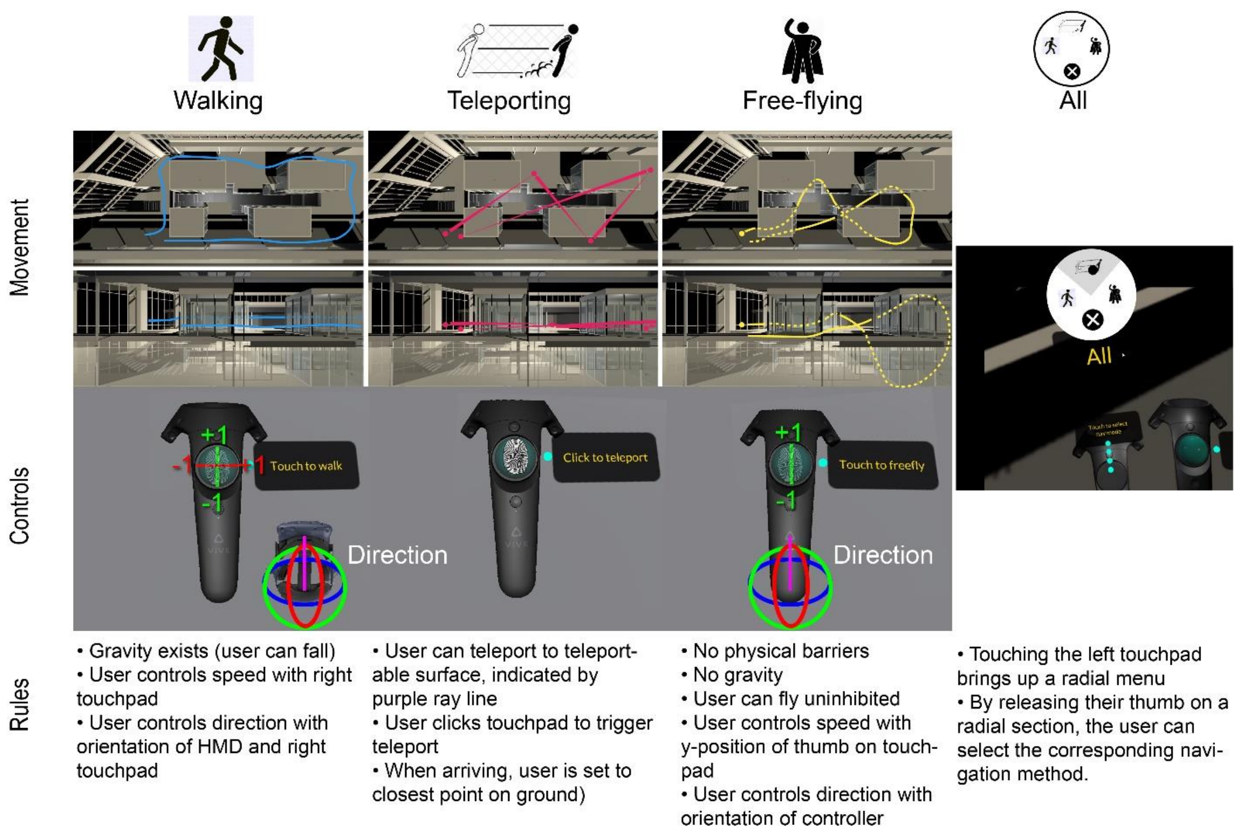

Figure 60. A comparison of movement, controls, and features for all three navigation methods.

Each navigation method was defined by three features: movement, controls, and rules. Controls specified the user's potential inputs via their controllers and HMD; rules were set by the researchers when programming the navigation methods; movement was the result of the previous two and defined how the user could move through the world (see Figure 60). For each navigation method, we created a control scheme the user learned during the tutorial task of each round.

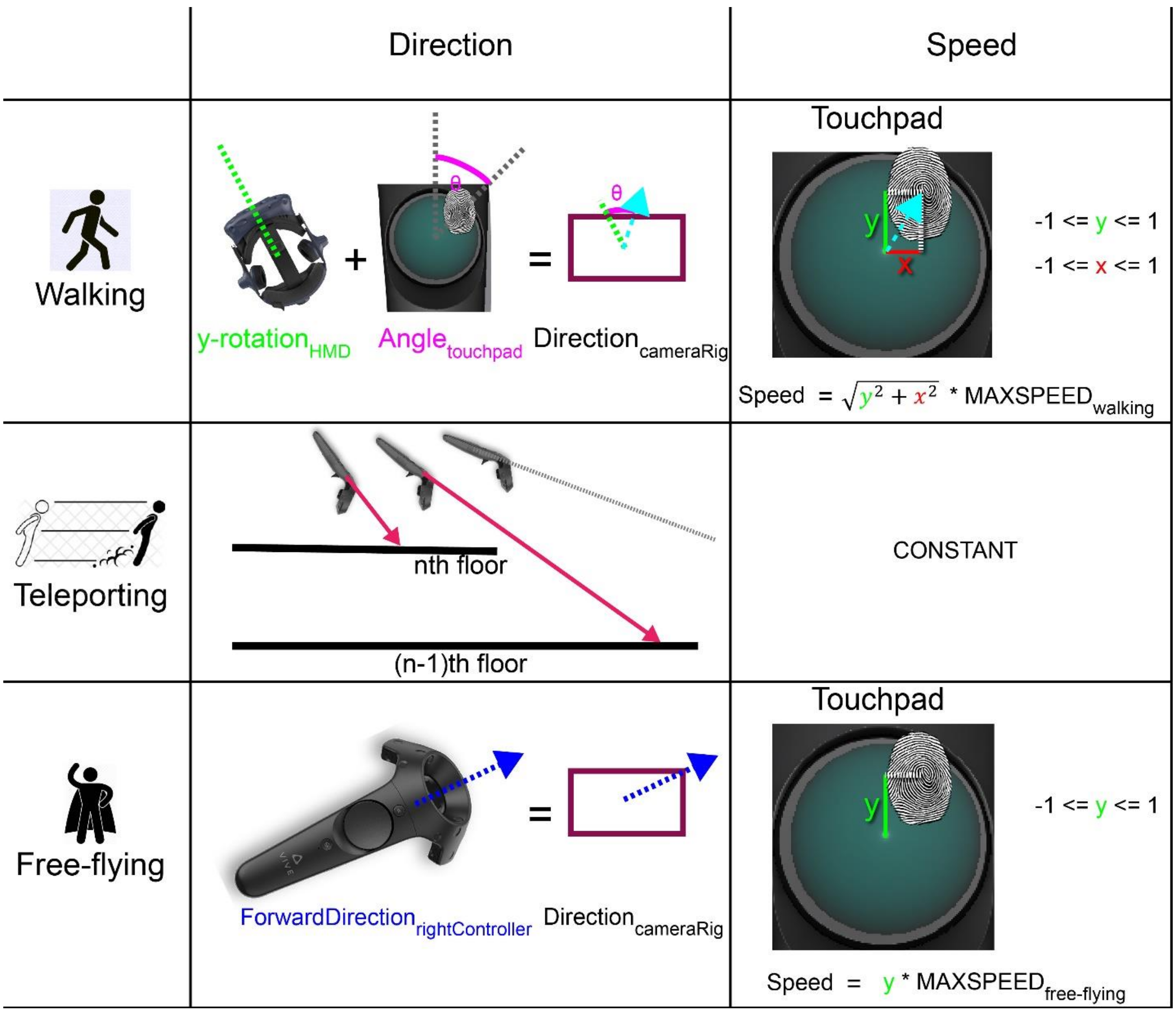

Figure 61. Advanced illustration of how direction and speed are calculated from user input, for each navigation method.

Walking was the navigation method that most closely imitated real-world locomotion on foot. The user was bound by gravity; this meant that when leaping over an edge, they could fall until they hit a surface with a collider. Of course, it was possible to walk up and down the stairs traversing the center of the atrium. We added invisible walls to the outside of the atrium to prevent users from falling to infinity. These walls were only active during sections where the user could only walk. We implemented a simple reset button for the research facilitator in case a user managed to fall out of the building anyway. A short burst of intensive pilot studies before study deployment helped us ensure that this case was very unlikely, even for less experienced study subjects.

The user controlled the speed and direction of their walking with their right controller and their HMD. The thumbpad on the controller is represented as a unit circle in the SteamVR SDK, and touchpad positions are given as (x, y) positions (where -1 <= x <= 1 and -1 <= y <= 1). The walking speed was then determined by the product of the distance of the touch position from the center of the touchpad and the maximum speed possible in the walking mode (2.5 meters per second). The walking direction, on the other hand, was determined by two angles: the y-rotation value of the HMD (i.e., which direction the user was looking, ignoring x- and z-rotations since walking into the air was not possible), and the angular distance between the position of user's thumb and the 12 o'clock line on the touchpad (see Figure 61). This setup allowed the user to walk independently of the direction of their gaze if needed by simple use of their thumb. Modulating the speed with the thumb ensured that faster and slower velocities were possible, and by moving their thumb into the lower quadrants of the thumbpad, the user could also walk backwards if needed.

Teleporting allowed the user to traverse distances within sight with a click of the touchpad on their right controller. In order to teleport, the user had to point their controller towards a suitable surface, and a purple ray would then indicate their target position were they to execute the teleport. Teleporting is a popular navigation method in VR video games.

While the teleport speed was constant, the user set the direction of their teleport with their controller. We embedded a ray caster inside the 3D representation of the controller, which would persistently shoot rays into the scene. If a suitable surface was hit, the user saw a representation of the ray in the form of a straight, purple line, and a small purple sphere marked the hit location, indicating the final destination of the teleport were they to execute it. A teleport was triggered by pressing the button embedded in the right touchpad. When arriving at a teleport destination, the user's virtual camera rig was automatically adjusted such that the user "stood" on the ground. The only surfaces that allowed teleportation were walkable surfaces. The ray always returned a hit location on the first suitable surface it encountered; it was thus only possible to traverse multiple floors at once if the user approached the staircase in the center of the atrium. Teleporting thus only allowed the user to navigate within sight.

Finally, **free-flying** allowed the user to travel through space with increased freedom. Neither gravity nor physical barriers were in their way. Using their right controller, the user could manipulate speed and direction. Speed was defined as the product of the y-value of the user's touch position on the thumbpad (-1 <= y <= 1) and the maximum speed for free-flying, which was set to 3 meters per second. The flying direction corresponded to the forward vector of the right controller, i.e., wherever the user pointed their right hand. Flying backwards was thus also possible.

When **all** navigation methods first became available, the navigation method active previously was deactivated, and we did not give the user a standard method to prevent any biases towards a specific navigation method. The user could switch between all three as they pleased and at any point in time. When putting their thumb onto the left touchpad, a radial menu popped up, with four sections (one for each navigation method, indicated by icons, and one to cancel). The user could then move a dot over the radial menu, using their thumb as a pointer device. When they released their thumb when hovering over a section, the corresponding navigation method was activated (or the menu was simply canceled). The currently active navigation method was displayed as text on a small panel over the left controller.

7.5 Reflective phase

Users in the experiment cohort got to inspect their own data in a mix of 3D and 2D data visualization, all inside VR (see Figure 62 and Figure 63). To that end, we created a 3D trajectory visualization, consisting of a dot density map of user positions over time, and a bar graph on a 2D panel over the user's left controller. This panel also contained a set of checkboxes to turn the data for individual tasks on and off. This setup presents a mix of spatial and abstract data visualization in one comprehensive VR interface for testing if users can utilize these types of data visualizations to improve their performance in VR Trial 2.

The visualization was created at runtime using a custom C# script, reading in data from a CSV file generated while the user completed VR Trial 1. The script iterated through every row in the dataset, instantiated the appropriate graphic symbol depending on the navigation method chosen, and added a data component that could later be used for the interactive legend to turn parts of the data overlay and off depending on user input.

To familiarize themselves with the visualization and the controls, we presented the Reflective phase to the user in two parts: First, we showed them data from a high-performing user in the control cohort; simultaneously, they were listening to a 3:41s audio tutorial introducing them to the base map, the data overlay, and the controls while outlining the goal of the Reflective phase: to identify strategies to improve their own performance while inspecting their own data. Then, after the tutorial was done and the user decided they had had enough practice, we loaded their own data in the visualization.

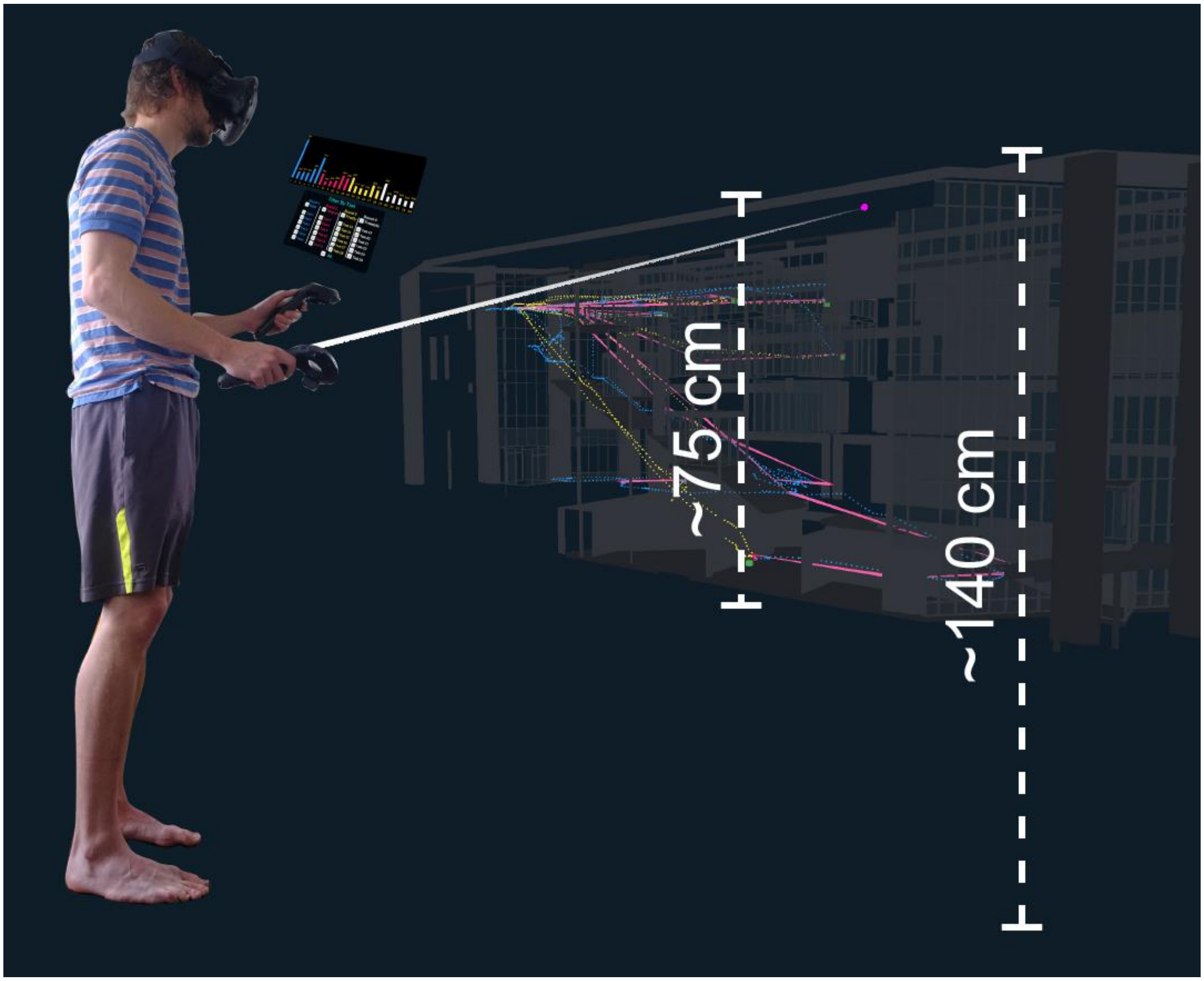

Figure 62. Approximation of a user in relation to the 3D base map and visualization in VR.

Base map and data overlay: The miniature model of Luddy Hall was around 75 cm tall, floating in front of the user such that the roof of the building was around 140 cm above the floor. While in the Reflective phase, the user could free-fly around the model so that they could inspect the model from all angles, independently of their physical height and range of motion. In order to achieve this, the entire visualization needed to be rescaled, also at runtime.

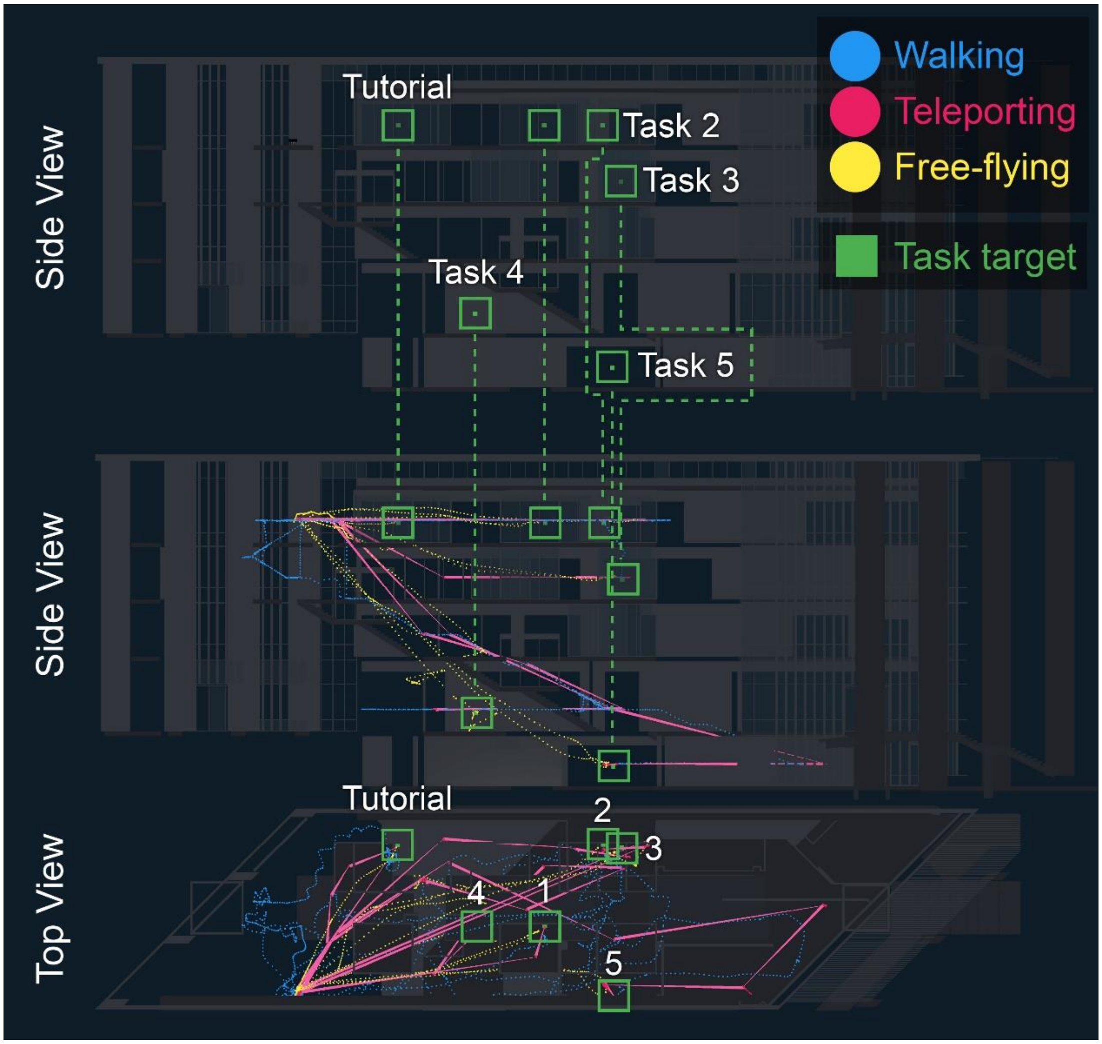

Figure 63. Three views of a user's data in the Reflective phase. Green cubes in the visualization show the location of the task submission sphere for each task. Blue, pink, and yellow visualize walking, teleporting, and free-flying, respectively.

Visual encoding: We chose color hue to represent navigation method, and x, y, z-position of the dot to encode the user's position. For teleport, we needed a way to make links between teleport start and point visually apparent, and encoded each teleport as a pink line, starting with the same width as the dot, and then thinning out evenly towards the teleport target. Figure 64 contains a series of detailed screenshots of the visual encoding across tasks and navigation methods.

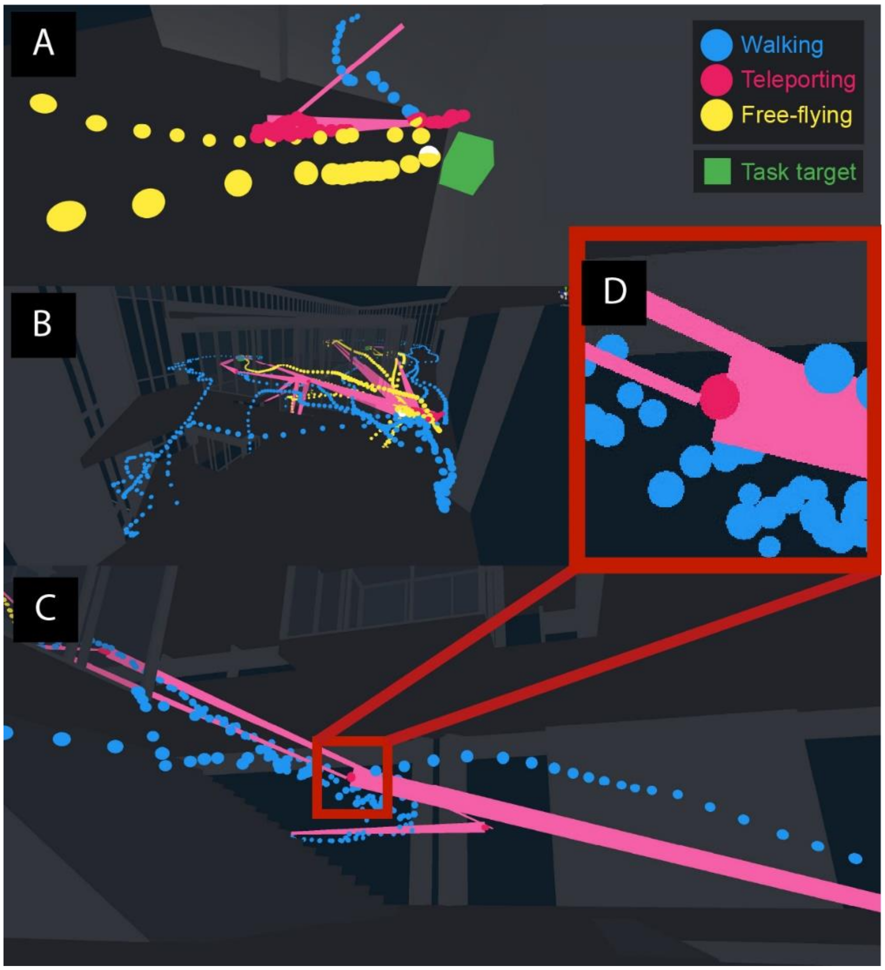

Figure 64. Visual encoding details. ***A****: a user's approaches to task #5. Note that there are two yellow trajectories, because they also used free-fly to finish this task when they had the choice.* ***B****: the start position on the fourth floor and the 24 trajectories leading away from it.* ***C****: a stop during a series of teleports.* ***D****: Note how the line leading to the teleport stop is thin, and the line leading away is wide, signaling the direction of the teleport.*

Bar graph for completion times: In order to give the user insight into their performance, we displayed a 2D bar graph on top of their left controller with completion times for all 24 tasks (see Figure 65). On the y-axis, we showed the completion time in seconds; the color of the bar encoded the navigation method possible during the task; and the x-axis contained the task number. On top of each bar, the completion time in seconds (rounded to one decimal) was displayed.

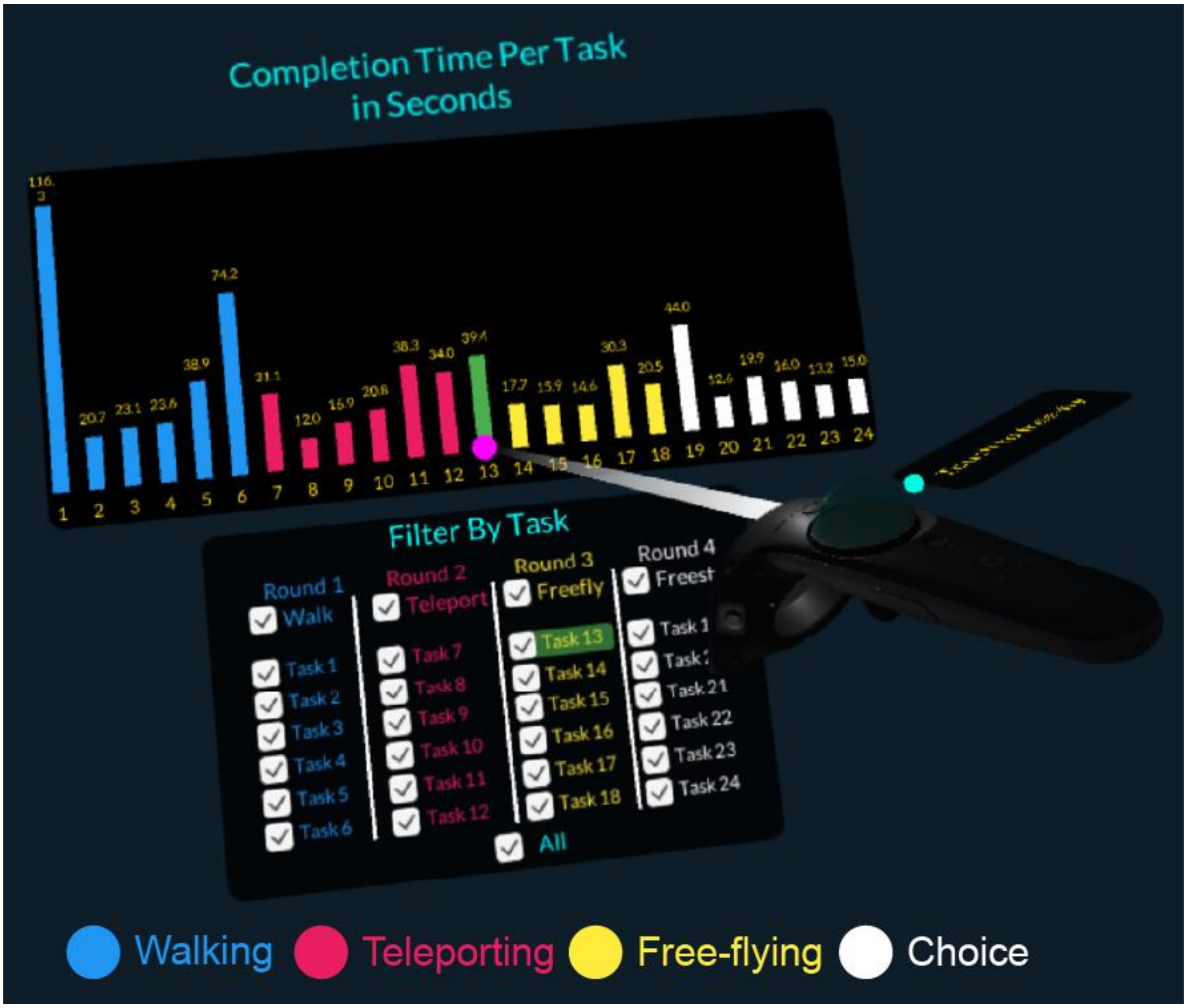

Figure 65. The interactive legend and bar graph visualization for the completion times of all 24 tasks in VR Trial 1. Note the link and brush functionality as the user is hovering over the bar for task #13.

Interactive legend: The bar graph functioned as part of an interactive legend that helped the user turn parts of the data overlay on and off. To help the user understand the connection between task number and completion time, and to allow them to focus on the tasks they wanted to explore, we implemented a link and brush functionality. The subject could use their right controller as a pointer. A purple sphere indicated the current hit point (see Figure 65). When hovering over a bar, both the bar and the corresponding checkbox were highlighted green. This way, the user could identify particularly long or short completion times and then inspect the corresponding trajectories in the Luddy Hall model. A checkbox to show or hide all data was located at the bottom of the interactive legend.

7.6 Mid-Questionnaire

Following the Reflective phase (experiment cohort) and short break (control cohort), all subjects were asked to sit down at the laptop again and answer 10 questions in a mid-questionnaire. In this questionnaire, we asked subjects about the features of the building (number of floors), the tasks they performed (how many tasks total, on how many floors, how many tasks on how many floors etc.), and, finally, we showed them two screenshots of the building (top view and side view), asking them to click where they though the start position was. The goal of the mid-questionnaire was to test whether there was a difference in spatial understanding between the two cohorts.

7.7 Metrics

To describe a user's performance, we captured a variety of metrics, both while in Unity via telemetry as well as derived data (such as completion time) and task scores from the mid-questionnaire.

The main metric to assess performance was **task completion time** in seconds, measured from the moment the user gained control of their movement to the frame where they had touched the virtual submit button for one second (see Figure 58D), at which point the timer was reset.

Another metric for our data analysis was the score users attained in the mid-questionnaire section of the experiment. There was a total of 10 questions for a total of 10 points. Finally, we included a self-reported satisfaction score at the end of the post-questionnaire, where the user had to indicate whether they felt satisfied after the experiment, using a 5-point Likert scale.

7.8 Apparatus

We ran our experiment on an Alienware 17 R4 with a 17.3” display, running Windows 10. We used an HTC Vive VR HMD in a play space of around 9x9 feet (3x3 meters) in a secluded collaborative space in Luddy Hall on the Indiana University campus in Bloomington, IN, USA. We used screen-recording software as well as a Logitech C930e webcam to capture the user’s action with audio and video, both in VR and the physical world.

7.9 Results

We recruited 71 subjects via email lists, social media, word-of-mouth, and from a pool of previous user study subjects over a period of 16 days. We conducted the experiment in a secluded area of Luddy Hall itself, granting privacy for the subjects. While running the experiment, two subjects had to abort their participation during VR Trial 1 due to motion sickness. One more subject had to stop during VR Trial 2 for the same reason. This left us with a total of 68 subjects for data analysis (34 per cohort). Subjects spent an average of **43.5** (**control**) and **60.4 minutes** (**experiment**) in the study, starting with the study information sheet at the beginning of the survey and ending after submitting the post-questionnaire. For all analyses, we omitted the tutorial task, i.e., the first task in every round where users were introduced to the navigation method for that round.

7.9.1 Demographics

30 of our subjects identified as female and 38 as male. The majority of subjects (40) were between 21 and 30 years old; further, 20 were between 18 and 20, six were between 31 and 40, and two were between 51 and 60 years old. The overwhelming majority were English native speakers (54). The largest group of non-native speakers spoke Bengali (five subjects). 65 subjects were right-handed; 3 were left-handed. Around half of the subjects (32) stated that they had no vision impairments; 26 were near-sighted. Participants were allowed to wear glasses under the HMD or contact lenses as needed. 67 subjects indicated that they were not color-blind; one subject preferred not to answer that question.

In terms of prior experience with VR, video games, and 3D applications in general, the overwhelming majority of subjects had used VR before (51); out of these, 37 had used it rarely, nine occasionally, and five often. The HTC Vive, Vive Pro, or Cosmos was the most used VR system, indicated by 21 subjects. 10 did not remember which device they had used. Over two thirds of subjects (43) said that they played video games in the past 12 months, mostly on smartphones or other handheld devices (27). Also, 27 subjects had played first-person shooters. Further, 36 subjects had used 3D software before, such as Rhino (6), AutoCAD (5), and Unity (5). When asked whether they would say that they were familiar with Luddy Hall, 10 subjects strongly agreed, 22 somewhat agreed, 10 neither agreed nor disagreed, nine somewhat agreed, and 17 strongly disagreed. We also asked our subjects to state their familiarity with six basic visualization types (tables, charts, graphs, maps, trees, and networks) and did **not find any correlation** between the self-reported familiarity scores with six visualization types and the correct answers in the mid-questionnaire.

7.9.2 Performance improvement

To answer RQ1, we needed to compare the completion times for VR Trial 2 for both cohorts. We isolated each subject's task completion times (excluding the four tasks that corresponded to the tutorial task in VR Trial 1), leaving us with 20 observations by subject (680 by cohort). To ensure that observations where users needed excessive amounts of time did not obstruct the validity of our analysis, we removed 39 observations from the control cohort (**m = 50.70 seconds**) and 38 observations from the experiment cohort (**m = 47.13 seconds**). We then performed a Welch's Two-Sample t-test with the remaining observations. This showed that there was **a significant difference** in task completion times between the two cohorts during VR Trial 2 (**t = 2.465, p = 0.01383**). We were then interested in determining whether this difference was caused by the Reflective phase or whether subjects in the experiment cohort were naturally more able in VR. We thus compared the completion times for VR Trial 1 after removing 42 and 40 observations from the control and experiment cohorts, respectively. A Welch's Two-Sample t-test then showed **no significant difference** between the two cohorts for completion times in VR Trial 1 (**t = -0.84297, p = 0.3994**). We thus **confirm H1** (the experiment cohort achieves significantly lower completion times than the control cohort during VR Trial 2). Figure 66 contains a collection of boxplots for the completion times per task and navigation method, separated by cohort.

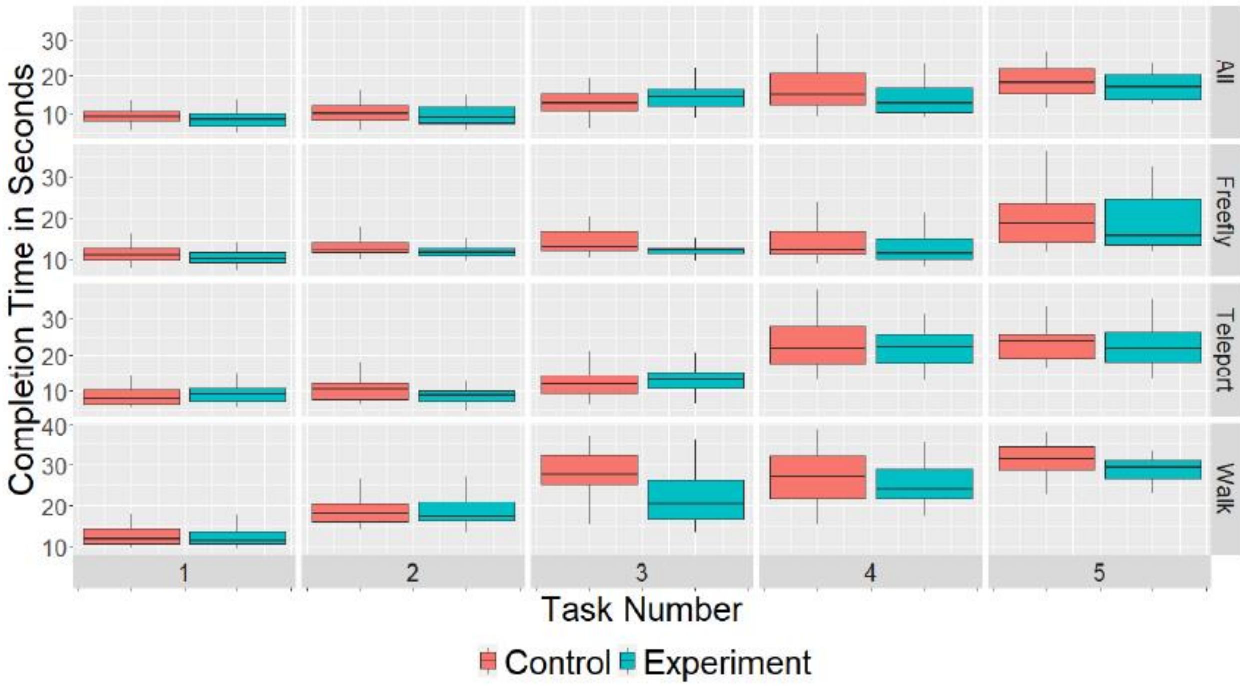

Figure 66. Faceted scatter graphs of completion times by task number (horizontal) per round (vertical).

Further, while it was to be expected that both cohorts would improve their times during VR Trial 2 (due to the learning effect), we wanted to identify the difference in the rate of change for completion times. We found a **significant difference** between the two cohorts ($m_{control}$= -5.38 seconds, $m_{experiment}$ = -7.48 seconds, $\boldsymbol{t = 3.1458}$, $\boldsymbol{p = 0.001693}$). This means that users in the experiment cohort improved by more than 2 seconds compared to users in the control cohort (on average). We thus **confirm H2**.

7.9.3 Mid-Questionnaire score

In addition to checking whether our intervention helped the experiment cohort achieve lower completion times than the control cohort in VR Trial 2, we compared the scores from the mid-questionnaire, where we asked our users to answer questions about the navigation tasks they performed with regards to the spatial layout of Luddy Hall. We found that experiment users performed **significantly better** than control users with mean scores of **5.71** and **4.29**, respectively (**t = -2.7028, p-value = 0.008734**). We thus **confirm H3**. Figure 67 shows violin plots for the total task scores for both cohorts along with jittered points for all 68 scores (and both medians).

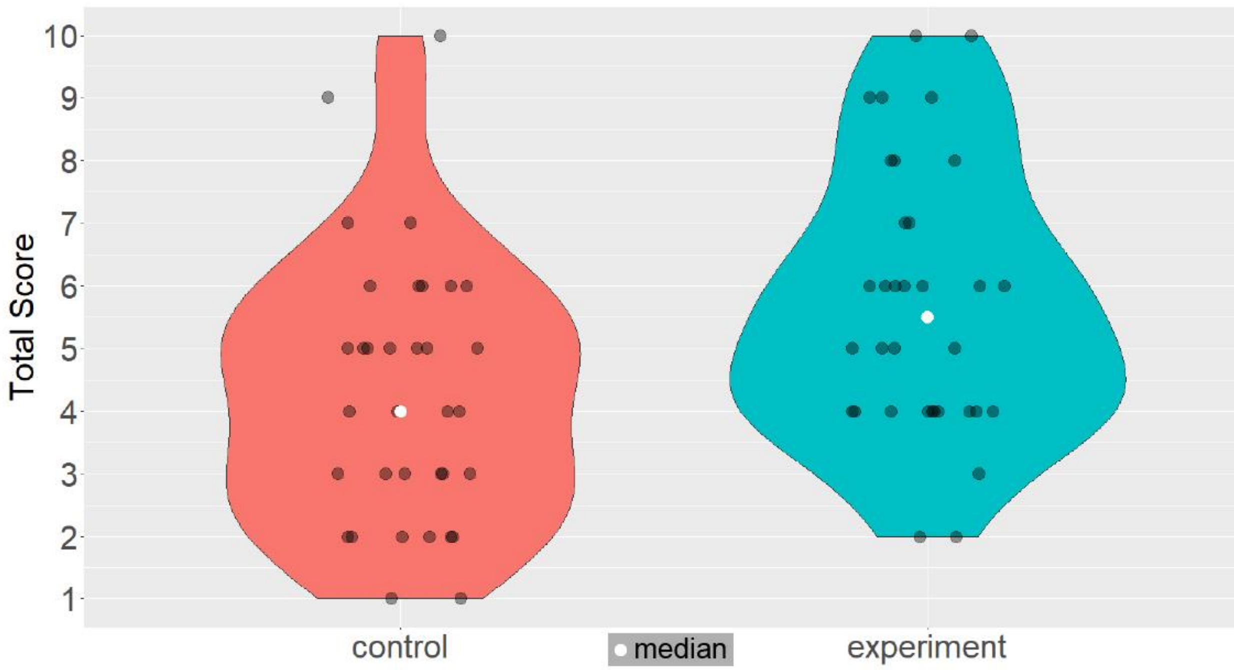

Figure 67. Violin plots of total task scores for bpth cohorts. The white dot in each plot encodes the median score ($median_{Control}$ = 4, $median_{Experiment}$ = 5.5). The black dots symbolize all 68 individual scores. Horizontal jitter has been applied to avoid overplotting.

To further prove the user's understanding of the spatial layout of Luddy Hall, we included a task where we presented them with two screenshots of Luddy Hall, one from above, one from the side. Figure 68 shows a dot density map of user clicks, color-coded for correct vs. incorrect answers, and shape-coded for cohort. We added a 60px margin of tolerance, resulting in an area of ~3.5 m^2 in real-world units.

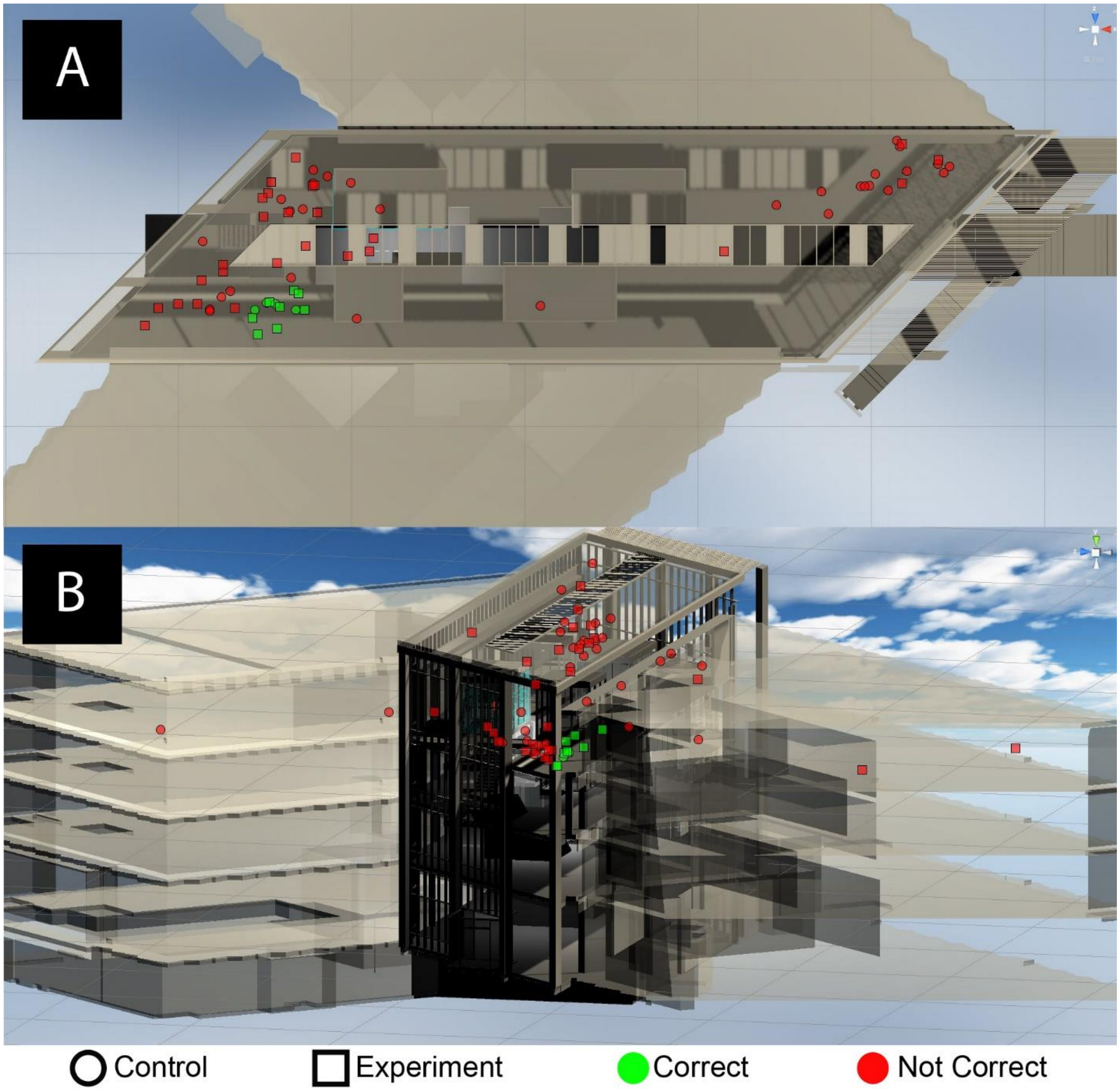

Figure 68. Dot density maps pf click positions where users were asked to indicate where in these screenshots their start position was. ***A****: top view.* ***B****: side view.*

Like the rest of the mid-questionnaire, the visualization shows a stark difference between the cohorts. For the top view, there were 12 correct answers (4 for control, 8 for experiment); for the side view, there were only 10 correct answers (2 for control, 8 for experiment). The difference in correct answers for the side view between the cohorts is significant (***t* = -2.0899, *p* = 0.04158**). This is likely thanks to the Reflective phase, because every trajectory in the Reflective phase visualization marked the user's start position (see Figure 64B), thus giving an advantage to the experiment cohort.

While many users were at least able to determine that they had started somewhere in the atrium of Luddy Hall, a surprisingly large number of subjects mistook the barely modeled wings of the building for the location where they spent the experiment (see Figure 68B). As becomes apparent from Figure 68A, on the other, the top view helped users narrow down the amount of choices, and the majority of subjects picked one of the four corners of the atrium. Note that we found **no significant correlation** between how familiar users were with Luddy Hall as indicated on a 5-point Likert scale and the total score in the mid-questionnaire.

7.9.4 Choice of navigation methods

During the last round of tasks in VR Trials 1 and 2, the subjects could switch between navigation methods at will. In RQ4, we wanted to check what navigation methods subjects would employ when completing these tasks. We had hypothesized that subjects would use teleport to reach targets within sight of the start position (**H4a**), free-fly for targets out of sight of the start position (**H4b**), and walking at the very end as a means to get to floor level (**H4c**).

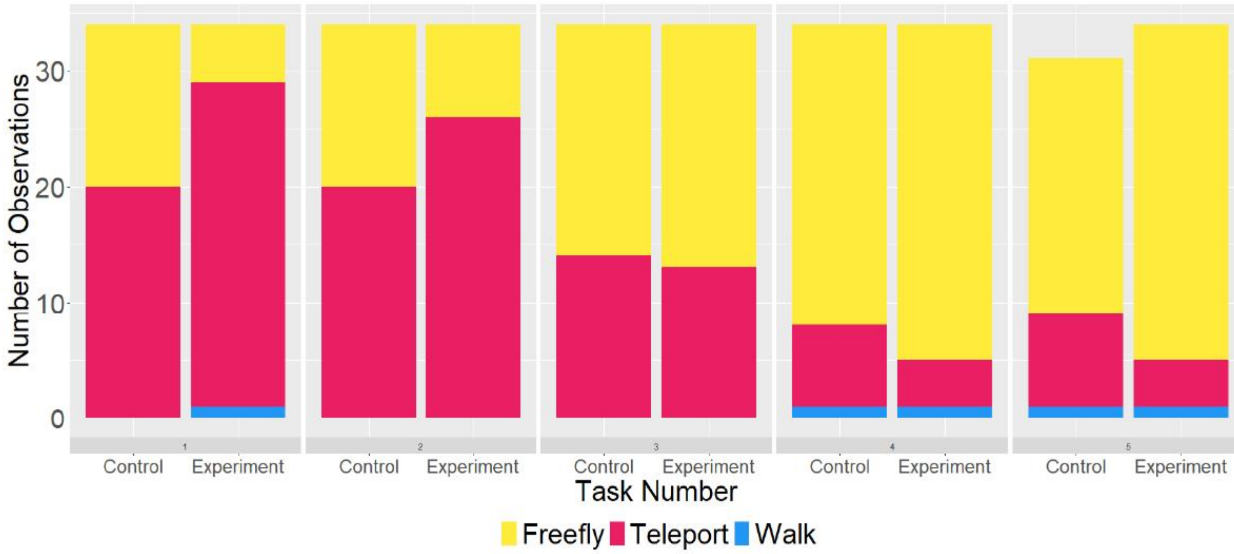

Figure 69. Navigation methods selected by subjects at task submission during VR Trial 2. Three outliers for the control cohort were removed for Task 5.

Figure 69 shows a bar graph of the last logged navigation method for the 5 tasks in the last round of VR Trial 2 for all 68 subjects (minus three outliers in the control cohort), yielding 337 observations. The tendency for all subjects to end tasks 1 and 2 teleporting (both are on the same floor as the start position) becomes apparent; for tasks 3 through 5, subsequently, users preferred free-flying as a way to traverse larger distances quicker (and go through walls thanks to the lifted physical restrictions when using free-flying). While we can **confirm H4a** (teleporting for targets within sight) and **H4b** (free-flying for targets out of sight), only a small minority of users ever ended a task by walking, prompting us to **reject H4c** (walking is preferred for ending tasks quickly).

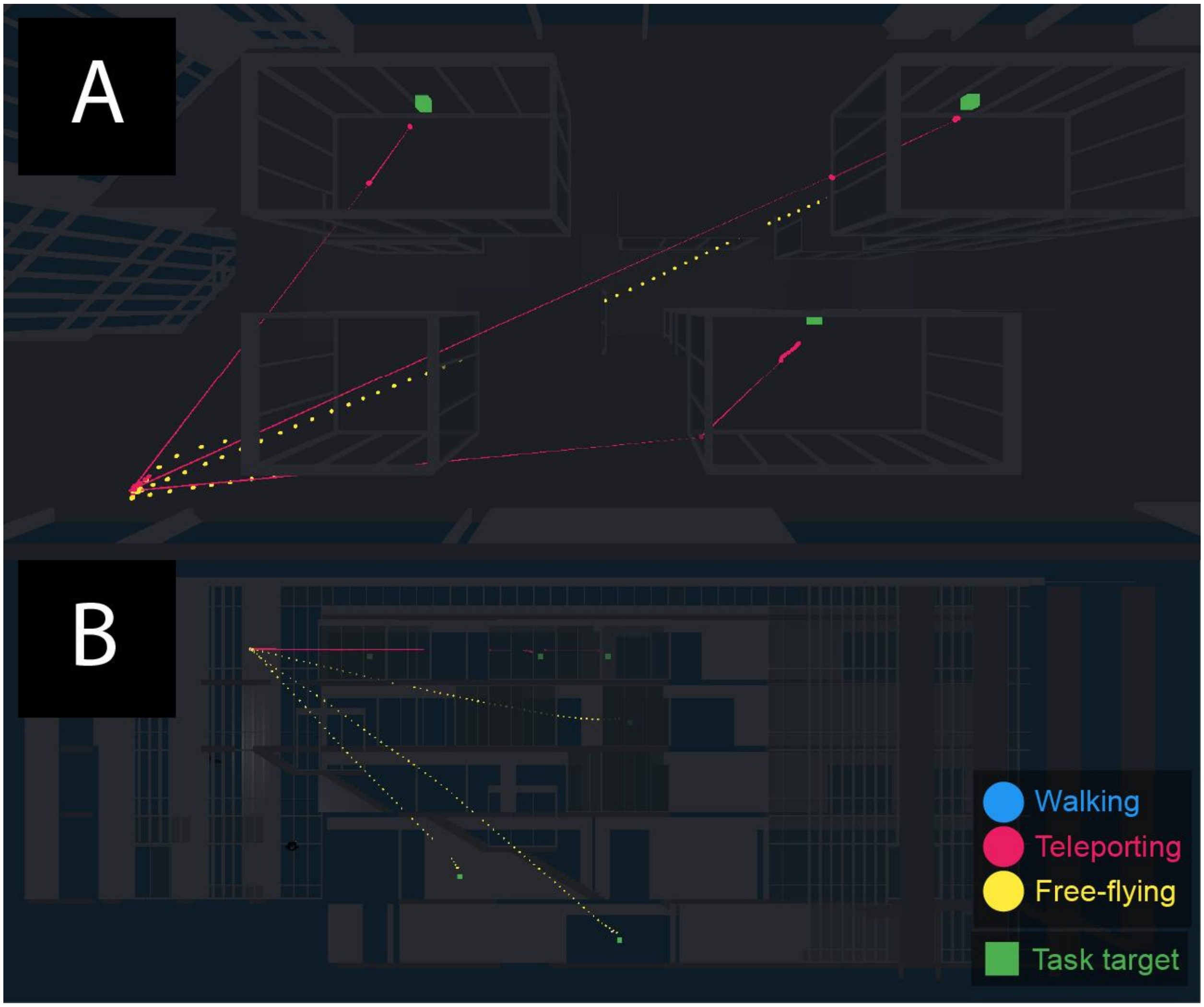

Figure 70. The "winning strategy". A: using teleport for targets on the same floor as the start position. B: using free-fly for all others.

7.9.5 Satisfaction

In terms of satisfaction, we can report that we found **no significant difference** between the cohorts, and thus **confirm H5**. Both had a high mean satisfaction on a 5-point Likert scale (**m = 4.29, SD = 0.84**). Additionally, we found **no significant correlation** between a user's score in the mid-questionnaire and their reported satisfaction. While the overall number of users who reported at least a little motion sickness was quite high (40 subjects, **m = 1.63, SD = 0.57**), it appeared to have **no significant correlation** with the total score. Similarly, we found **no significant difference** in motion sickness between the two cohorts (***t* = -0.63481, *p* = 0.5279**).

7.9.6 Post-Questionnaire results

In our post-questionnaire, we aimed to identify which navigation method was the most popular. To that end, we asked subjects to rank all three navigation methods from most to least favorite. Table 12 contains an overview of different rankings and their frequency

Table 12. Rankings for navigation methods.

Walk	**Teleport**	**Free-fly**	**Frequency**
1	2	3	5
2	1	3	2
2	3	1	7
3	1	2	19
3	2	1	35

The absolute majority of subjects (**n = 35**) preferred **free-fly** to teleport and walk, just over a quarter (**n = 19**) favored teleport over free-fly and walk, and only a minority of subjects liked walk the most (**n = 5**). Note that no single subject preferred walk over free-fly and teleport (in that order. Further, we found that all users, regardless of cohort, liked the VR experience as indicated by high means on a 5-point Likert scale for different aspects: overall (**m = 4.57, SD = 0.61**), hardware (**m = 4.49, SD = 0.78**), and instructions (**m = 4.59, SD = 0.6**).

7.9.7 Reflective phase feedback

Similarly, we aimed to measure how subjects in the experiment cohort perceived the interactive tools at their disposal during the Reflective phase using a 5-point Likert scale, and found that subjects overwhelmingly found them useful: filters and checkboxes (**m = 4.32, SD = 0.81**), color coding (**m = 4.68, SD = 0.77**), the bar graph of completion times (**m = 4.29, SD = 0.97**), and the ability to fly around the miniature base map of Luddy Hall (**m = 4.26, SD = 0.93**). Finally, the majority of users indicated that they were able to identify a better strategy to achieve faster completion times in the next round of tasks (**m = 4.32, SD = 0.59**).

The subjects then indicated how efficiently they thought they had navigated the 3D space while only one navigation method was possible (again using a 5-point Likert scale). These numbers reflect the rankings of the navigation methods discussed earlier: Users overall found that they did not perform efficiently with regards to walking (**m = 2.29, SD = 1.06**) while teleporting (**m = 3.74, SD = 1.05**) and free-fly (**m = 4.12, SD = 1.15**) were ranked higher.

7.10 Discussion

In this chapter and chapter 6 above, we presented two user studies with Reflective phases involving different interactions and base maps (see Table 13).

Table 13. Side-by-side comparison of RUI Reflective and Luddy VR user studies.

	RUI VR Reflective	**Luddy VR**
Setups (VR/Desktop)	3 (2D Desktop, VR Tabletop, VR Standup)	1 (VR)
Cohorts	2 (control, experiment)	2 (control, experiment)
Visualization types	Map (VR setups), graph (2D Desktop)	Map, graph
Scale of reference system	1:1	1:30
Graphic symbols	Volume (VR setups), line (2D Desktop), linguistic/text (2D Desktop)	Volume, line, linguistic/text
Graphic variables	Position (3D, 2D), color hue, color saturation	Position, color hue, size, velocity
Interactions	Filter, navigate, animate/replay	Filter, navigate, link and brush

We designed these Reflective phases by consulting the DVL-FW, specifically its typologies for interactions, graphic symbols, and graphic variables. Notably, for the Luddy VR study, users that experienced the Reflective phase had significantly better performance in terms of task completion time and better scores in the mid-questionnaire. We saw no such improvement for users in the VR setups of the RUI VR study, only for the 2D Desktop users who were presented with line graphs. In this section, we examine the differences between the Reflective phases in these two studies as far as the DVL-FW typology is concerned.

7.10.1 Comparison of Reflective phase implementations

The key differences between the Reflective phases are listed in Table 13. Based on these, we conclude that a variety of factors accounted for our results.

Visualization type and scale of reference system: First, while the RUI VR Reflective phase visualization presented the user's movement and tissue block data on a 1:1 scale, in the Luddy VR study, the reference system for the data was shown at a scale of 1:30 such that the miniature version of Luddy Hall was displayed at a length of approximately 150 cm measuring from the southwest entrance to the northeast entrance (the corresponding side of the long side of the actual building is approximately 45 meters long). We conclude that this difference in scale between the two reference systems provided more value to the Luddy VR users for devising strategies to improve their performance. There is currently no **scale** typology in the DVL-FW that could capture this difference in a meaningful way. Further, Luddy Hall, with its naturally geospatial layout, waypoints, stairs, rooms, and floors, was more easily identified and read as a map, whereas the kidney and buzzer in the RUI VR study provided comparatively sparse orientation.

Second, Luddy VR users were given an additional visualization type in the form of a (bar) graph with their completion times per task, allowing them to quickly identify patterns, trends, and maxima as well as minima in their performance via a more abstract data visualization. In RUI VR, on the other hand, the completion time was not explicitly shown, and could only be inferred from the time stamp on the time slider. This level of abstraction needed for visually non-explicit information probably yielded a disadvantage for VR Tabletop and VR Standup users, while 2D Desktop users could identify the longest and shortest completion time in their intro session to the Reflective phase rather easily (see also Figure 54E and F, where 12 and 13 out of 14 users correctly identified the shortest and longest completion time of the best-performing subject, respectively).

Graphic symbols and graphic variables: Both Reflective phase implementations had a set of graphic symbols and graphic variables in common. Mainly, both used volumes (spheres, cubes) to encode the user's position (and the position of their hands) as well as the location of the tissue block over time (RUI Reflective) and the task destinations (Luddy VR). For the implementation of the more traditional 2D visualizations (line graph and bar graph), lines were used in both studies (though just for 2D Desktop in the RUI VR experiment). On a side note, we consider the bars in bar graphs as lines, because only the length (not the width) is used to encode a data record as the graphic variable size. Both implementations also make use of linguistic symbols, specifically text, to denote task numbers (both), completion times (2D Desktop in RUI VR, all of Luddy VR), and parts of the interactive legends. In terms of graphic variables, there were differences. Both Reflective phases used position (3D for VR visualizations and 2D for line graph and bar graph) and color hue to encode the input device visualized (RUI VR) or the navigation method chosen (Luddy VR). In RUI VR, however, we also used color saturation to indicate the angular difference between the tissue and target blocks (and thus the rotation accuracy), while in Luddy VR, we used the graphic variable of size to encode the user's performance metric (completion time) in the bar graph. Finally, in Luddy VR, we employed velocity in two ways: when visualizing the vector between a user's teleporting start and end point, and to show head and hand movement direction when replaying a dataset in RUI VR. All graphic symbol and graphic variable encodings were received and understood well based on mid- and post-questionnaire data. Generally, we conclude that the 3D trajectories that made up the Reflective phase visualizations illustrated the performance for virtual navigation tasks better, because trajectories through virtual environments are a type of data overlay familiar to most people through navigation systems in cars, on phones, and in video games. The trajectories for the head and hand movements in the RUI Reflective study may have been too abstract for users to derive viable strategies.

Interactions: Over the two studies, we implemented four interaction types: **filter**, **navigate**, **animate**, and **link and brush**. **Filter** was implemented in both studies: The user could turn parts of the data overlay on and off via checkboxes that they could click via a 3D pointer. By default, all checkboxes were checked and all data was visible. The checkboxes were part of the interactive legend. Likewise, users could **navigate** across the visualization by virtue of wearing VR equipment, allowing them to see the data from different angles. **Link and brush** and **animate/replay** were unique to Luddy VR and RUI Reflective, respectively. Luddy VR users could brush over a bar in their graph of completion times and then saw the corresponding menu entry highlighted that let them turn on and off the underlying data for that task. RUI Reflective users could play back their own data at different speeds by using the time slider, effectively creating an animation of their own movement and tissue block manipulation over time. However, it appears that this interaction type was less useful than expected.

It may have been hard for users to identify and select a playback speed that yielded insights about their completion times to them, and offering predetermined playback speeds may have been a better design choice. Additionally, playback for a task worked better when other data was turned off, thus **avoiding clutter**, and the steps needed to hide all data, jump to the time stamp of a task, and then playing it back at an insightful speed, required a series of actions on the user's part that may have been challenging for many users. Also, we assume that the animate/replay interactivity did not help users identify completion times properly, because completion time had to be derived from looking at time stamps, not directly via an auxiliary visualization like in Luddy VR. Link and brush for Luddy VR users, on the other hand, created a visual connection between the bar graph and the dot density map of their trajectories, thus allowing the user to evaluate their performance with this additional, derived data rather than having to infer or compute their performance from time stamps like in RUI VR.

7.10.2 Design implications

After its proposed extension, the DVL-FW interaction typology contains 24 types. While it is not possible to implement and validate this typology in its completeness within one dissertation, we argue that we were able to highlight the advantages and disadvantages of each of the four types for their application for VR visualizations (**filter, link and brush, navigate, animate/replay**). This section contains design implications derived from our implementation of the four types, and section 8.1 provides recommendations for future work on how these specific interactive types can be leveraged to help users improve VR performance.

The filter, link and brush, and navigate interactions proved to be valuable for Luddy VR users, who were able to achieve faster completion times than their counterparts in the control cohort. While we did not find that the Reflective phase helped RUI Reflective experiment users (in the VR setups) perform better than control users in the Plateau phase, we did find significant effects by metrics in the Reflective phase on performance in the Plateau phase and satisfaction regardless. These effects could help identify ways to nudge users to a better performance by tweaking parts of the Reflective phase. Subsequently, we describe which of these insights are actionable for the future design of interventions using the DVL-FW.

We identified one metric that had both a favorable effect on rotation accuracy and satisfaction: **head rotation around the y-axis**. Similarly, we found that many variables that describe the Reflective phase experience made the entire VR experiment less satisfying for subjects, among them total time spent, time slider usage, and time without the base map, which also had a detrimental effect on position and rotation accuracy (see Table 10). What are potential design tweaks to improve performance while maintaining user satisfaction? It appears that a beneficial improvement would be to **bring the data to the user**, e.g., by scaling down the reference system of the visualization (like for Luddy VR), thus minimizing the need for covering a wide area with their entire body or even just their hands while allowing them to turn around on the spot and inspect the data from various perspectives. Given that more body movement (measured by convex hull for the hands as well as the number of clusters for the right hand) had an adverse effect on accuracy measures, minimizing the user's need for movement might not only save them time but also energy and prevent cognitive fatigue. A more refined Reflective phase would thus encourage the user to make use of their head as a "camera" to inspect the data from different angles in a 3D perspective while not moving their entire body around. At the same time, this could also shorten the time spent in this analytical mode.

Additionally, we determined that less aggregation leads to happier users. This means that another good tweak would be to **avoid aggregated data views** where possible, especially at the start of the Reflective phase where all the data was turned on by default. Of course, it is also possible that the visualization type, together with the aggregated data view, was the confusing and less satisfying element for many RUI Reflective users, many of whom may have felt overwhelmed with initial display of data. Had they been presented with a 2D dot density map instead of the 3D + VR version of our experiment, an aggregated view could have revealed more patterns (like the 2D dot density maps in Figure 46).

It would also be possible to design a Reflective phase for RUI VR as a **mix of VR visualizations and traditional, 2D visualizations** such as a bar graph for completion times over trials (like for Luddy VR). The 3D dot density map with some visual encoding is a more advanced visualization type, and even if users extracted insights from their own data, it may not have been supporting their goal to act on these insights accordingly. This would also relieve the user from the need to derive performance variables such as the completion time while trying to memorize more successful strategies.

Yet another, more far-reaching change would be to more closely entangle any reflection about one's own data with the actual task hand rather than outsourcing it to a separate Reflective phase application. **Immediate visual feedback** could be given when the cube-matching task is done to encode accuracy, possible following some of the design goals of fluid interaction, such as "provide immediate visual feedback on interaction" (81). Additionally, because the tasks were performed in VR, **haptic feedback** could be used to indicate position and rotation accuracy during the tasks. Most VR controllers have vibration motors built in so that, e.g., the current accuracy could be mapped to the actuators inside the controller. This would shorten the user's time in the Reflective phase and provide a new environment for visual feedback altogether.

7.10.3 Limitations

We acknowledge several limitations to these studies.

First, while both RUI Reflective and Luddy VR contained a Reflective phase, the overall **study design** was slightly different. In Luddy VR, both cohorts completed the same set of tasks twice, with the intervention in between. In RUI VR, the Ramp-Up phase contained a different set of tasks than the Plateau phase, so it was not possible to compute within-subject improvements for these users.

Second, the **types of task** between the two studies were different (cube-matching vs. movement), which influenced what information users needed to retrieve from the Reflective phase. For example, RUI Reflective users had to identify ways to balance their efforts between position accuracy (with a focus on arm movements) and rotation accuracy (with a focus on hand movements), all without neglecting completion time. Luddy VR users, on the other hand, experienced a more mediated interaction in that they could move their entire body through the virtual space via simple button clicks and touches. As a result, the navigation methods demanded less physical movement and input. Likewise, data from the RUI Reflective cube-matching tasks was richer as users did not only have to take position into account when determining new strategies but also rotation. As a result, improving performance for RUI Reflective users **more challenging**. The data visualization of each user's behavior was not only in VR and in three-dimensional space but also interactive. This would have presented a challenge to any user, regardless of expertise with visualizations and VR, be it due to lack of expertise, spatial ability, confidence in virtual environments, or data visualization literacy. This is highlighted by the fact that users in the Desktop setup (with the simple, static line graph) were able to significantly improve their position and rotation accuracy while feeling more satisfied using a rather simple visualization.

Third, for the experiment cohort in the RUI Reflective study, participation in this study was a lot more **time-consuming** than for the control cohort, potentially yielding a benefit to the latter. From reading the study information sheet at the beginning of the pre-questionnaire to answering the final question of the post-questionnaire, RUI Reflective subjects spent an average of 3627.21 seconds (SD = 789.38 seconds) or ~60.45 minutes on the entire experience versus an average of 1811.67 seconds (SD = 840.1 seconds) or ~30.19 minutes for control subjects. This resulted in an average **difference of slightly over half an hour** between these two cohorts.

This is in stark contrast to Luddy VR, where experiment users needed 3624.82 seconds (SD = 717.5 seconds) or ~60.41 minutes versus 2608.41 seconds (SD = 428.46 seconds) or ~43.47 minutes for the control cohort (on average). This resulted in an average difference of just under 17 minutes, which is only ~57% of the time difference between the cohorts in the RUI Reflective study.

Subjects who went through the Reflective phase in the RUI Reflective study put on and then took off the HMD four different times, switching between HMD and the laptop with the instructions every time. In the future, the research design for similar interventions could be more streamlined.

Finally, while the telemetry data from the HMD and VR controllers allowed us to model a user's behavior for our data analysis, it did not allow us to draw conclusions about what parts of the data visualization in the Reflective phase the user was actually **focused** on. More recent developments in eye-tracking inside the HMD and foveated rendering might lead to the availability of advanced telemetry data in the future so that researchers can derive more meaningful and detailed information about a user's gaze than simplistic head orientation values.

7.10.4 Next steps

In further studies, we aim to explore a variety of potential adjustments for the Reflective phase by testing more interaction types included in the Reflective phase. For example, rather than presenting users with premade visualizations with minimal possible adjustments, users could **create their own visualizations** based on available data records, obtained either through telemetry (head and hand position, rotation) or computed at runtime (completion time, accuracy metrics, and spatial data such as velocity of tissue block placement). We could support this by implementing the new DVL-FW interaction type **visualize/encode**. Additionally, users could **annotate** their data, or compare their own data with someone else's side by side via **arranged and coordinated views**. Additionally, the tasks for these studies are rather abstract and thus might resemble real-world VR training and coaching tasks only superficially. As a result, it could be interesting to design a **real-world user study** with professionals from an application domain (such as the medical or engineering fields). We expand on potential future work in section 8.3.

8. Discussion

In this dissertation, we proposed an extension of the current DVL-FW interaction typology from nine to 24 interaction types. These types were gathered from reviewing literature from a variety of disciplines (among them information visualization, computer science, statistics, geography, and human-computer interaction). We then implemented four interaction types in two user studies. Additionally, we presented research and development efforts for HuBMAP, an effort to create an atlas of the healthy, adult human body at single-cell resolution. Specifically, we validated the design of the RUI as an application running on a 2D interface against two VR implementations, which we also tested in a user study. All three user studies involved a series of performance metrics, gathered via telemetry and questionnaires. In Table 14, Table 15, and Table 16, we listed all the research questions and hypotheses in this dissertation and added marks to denote whether they were confirmed or rejected.

Table 14. Hypotheses and results for RUI VR (HuBMAP) user study (✓ = confirmed, ✗ = rejected)

Hypothesis	**Description**	**Result**
H1a	Users in VR Tabletop and VR Standup achieve significantly higher position accuracy than users in 2D Desktop.	✗
H1b	Users in VR Tabletop and VR Standup achieve significantly higher rotation accuracy than users in 2D Desktop.	✓
H1c	Users in VR Tabletop and VR Standup have significantly lower completion times than users in 2D Desktop.	✓
H2a	We do not expect any major bias for any setup in any dimension.	✓
H2b	We expect the error to be greatest for the 2D Desktop setup due to its restricted input devices and limited viewing positions.	✓
H3a	More complex tasks lead to lower position accuracy for all setups.	✗
H3b	More complex tasks lead to lower rotation accuracy for all setups.	✓
H3c	More complex tasks lead to higher completion times for all setups.	✗
H4	VR users need a lower number of tasks to plateau than 2D Desktop users.	✓
H5	In all setups, the more time users spend on a task, the higher position and rotation accuracy they achieve.	✗
H6a	Users in both VR setups are more satisfied with their performance than 2D Desktop users.	✓
H6b	There is no significant difference in user satisfaction between VR Standup and VR Tabletop users.	✓

Table 15. Hypotheses and results for RUI Reflective user study (✓ = confirmed, ✗ = rejected)

Hypothesis	Description	Result
H1	H1: There will be a significant difference in completion time and accuracy for Ramp-up and Plateau phases between control (without Reflective phase) and experiment group (with Reflective phase). However, this will only occur for the VR Standup and VR Tabletop setups; the Desktop users will not be able to gain significant gains in accuracy and completion time over their peers.	✗
H2	H2: Users in the experiment group that previously spread out across a larger area in the Ramp-Up phase will use less space and concentrate on an overall smaller work area in the Plateau phase. They will also use less space in the Plateau phase on average than the control group.	✗
H3a	H3a: Most users will use the time slider to scroll through around 1000% (=10 times) the time span of their dataset.	✗
H3b	H3b: The most selected location for the play head of the slider will be towards the very end of the timecode in the dataset.	✗
H3c	H3c: Users will spend the majority of time with the kidney turned on as the presence of a reference organ is highly useful to understand the data overlay. Many users may not be able to conceptualize the potential value of having the kidney turned off while inspecting their data to remove clutter. The kidney is turned on by default.	✓
H4a	H4a: More distance traveled has a negative effect completion times in the Plateau phase.	✗
H4b	H4b: More distance traveled has a negative effect on distance (higher position accuracy) in the Plateau phase.	✓
H4c	H4c: More head rotations have a negative effect on completion times in the Plateau phase.	✗
H4d	H4d: More head rotations have a negative effect on distance (higher position accuracy) in the Plateau phase. This may be due to high-performing users feeling more comfortable in 3D environments in general, and VR specifically, enabling them to move around their own data more fluently in the first place.	✗

Hypothesis	Description	Result
H5a	H5a: There is a significant negative correlation between task score in the mid-questionnaire and position accuracy in the Plateau phase in terms of distance.	✓
H5b	H5b(1): There is a significant correlation between task score in the mid-questionnaire and position accuracy in the Plateau phase in terms of median x-error.	✗
	H5b(2): There is a significant correlation between task score in the mid-questionnaire and position accuracy in the Plateau phase in terms of median y-error.	✗
	H5b(3): There is a significant correlation between task score in the mid-questionnaire and position accuracy in the Plateau phase in terms of median z-error.	✗
	H5b(4): There is a significant correlation between task score in the mid-questionnaire and position accuracy in the Plateau phase in terms of median bias.	✓
H5c	H5c: There is a significant negative correlation between task score in the mid-questionnaire and rotation accuracy in the Plateau phase.	✗
H5d	H5d: There is no significant negative correlation between task score in the mid-questionnaire and completion time in the Plateau phase.	✓
H5e	H5e: The majority of users agree or strongly agree that the subject shown to them in the Reflective phase was highly fast and accurate.	✓

Table 16. Hypotheses and results for Luddy VR user study (✓ = confirmed, ✗ = rejected)

Hypothesis	Description	Result
H1	The experiment cohort achieves significantly lower completion times than the control cohort during VR Trial 2.	✓
H2	The experiment cohort achieves significantly larger changes in completion times between trial 1 and trial 2.	✓
H3	The experiment cohort achieves higher scores in the mid-questionnaire than the control cohort.	✓
H4a	Subjects prefer teleporting when finalizing a task within sight of the start position.	✓
H4b	Subjects prefer free-flying when finalizing a task out of sight of the start position	✓
H4c	Subjects prefer walking just as they finalize a task.	✗
H5	There will be no significant difference in satisfaction between the cohorts.	✓

In user study 1 (HuBMAP), we found that subjects in the two VR implementations outperformed 2D Desktop users in terms completion times, rotation accuracy, and satisfaction; however, we found no significant difference for position accuracy, see Table 14 (beyond a significant error in the x-dimension, likely due to the predefined virtual camera locations through which the user can view the reference organ from different perspectives). We were further able to validate the 2D Desktop implementation of the RUI by identifying the median position accuracy of the 2D Desktop users as **1.3 mm** given the kidney height on the laptop display after an average of **8 identical tasks** in a sequence. Likewise, with an average completion time of **22.6 seconds** and a median rotation accuracy of **5.89 degrees**, we met performance benchmarks set during the initial conceptualization of the RUI (one minute per registration, 1-2 mm accuracy). The next step could be to collect interaction logs from users in the HuBMAP consortium who use the recently deployed RUI version 1.5 (65). Performing user studies with this version that is currently being used to TMCs across the country and the globe could not only provide "in the wild" numbers but also identify frequently used features, allowing us to recommend more developer time for improving other key aspects of the interface, such as metadata entry and sematic annotation via collision detection, in addition to the 3D interaction we tested in our user study.

The two user studies with the Reflective phases yielded mixed results. For the RUI VR Reflective study, see Table 15, while we did not find significant differences in performance between the control and experiment cohorts for the two VR setups, we found a significant difference in rotation accuracy between the cohorts for 2D Desktop setup after users in the experiment cohort were shown a line graph of their performance so far. Further, we identified the total degrees of head rotations around the y-axis to have a favorable effect on satisfaction and rotation accuracy, in addition to a range of other behavioral metrics that had significant effects on the post-intervention performance. For the Luddy Reflective phase study, on the other hand, we found most of our hypotheses confirmed (see Table 16) while identifying significant differences not only in completion time between the cohorts during the second trial but also for the scores in the mid-questionnaire. In this questionnaire, we tested how well users remembered the topology of the virtual building as well as the number and destination of tasks.

All 152 subjects across the three studies used the same VR equipment consisting of an HTC Vive HMD and VR controllers. The VR applications were made in Unity 2019.4 (running on an Alienware 17 R4 using Windows 10), and selected scripts were made available alongside videos detailing the study procedures and all parts of the VR experiences. These materials have been collected in the Supporting Information and can also be viewed on GitHub (https://github.com/andreasbueckle/bueckle-dissertation-supporting-information).

8.1 Recommendations

Based on our findings, we conclude that one of the strongest benefits of VR is the high satisfaction for users. In the RUI VR user study, users in both VR setups were more satisfied with their performance than 2D Desktop users (Table 14, H6a), and there was no significant difference between those VR users who were standing and those who were sitting (Table 14, H6b). Likewise, in the Luddy VR user study, we found that users reported high satisfaction (m = 4.29 on a 5-point Likert scale) regardless of whether they experienced a Reflective phase (Table 16, H5). Designers of VR visualizations can thus expect a lot of **good will and curiosity** from experienced and novice VR users alike.

Further, for VR visualizations, we recommend to **bring the data to the user** based on the different success rates of the Reflective phases of the RUI Reflective and Luddy VR studies. The scale of the Reflective phase visualization was a major difference between the RUI Reflective and the Luddy VR study, where the RUI Reflective dot density map was at a 1:1 scale, while the Luddy VR dot density map was contained in a miniature version of the building at a 1:30 scale, allowing the user to gain an overview of the dataset much easier while still being able to inspect various subsets in an organic way simply by moving their head around and inside the virtual building.

Likewise, **less aggregation makes happier users**. In RUI reflective, we found that high time slider values (indicating that the user inspected data towards the end of the dataset with many tasks visible at once) had a negative effect on satisfaction. Because the time slider showed data based on the position of the play head, the graphic symbols for later tasks were shown on top of earlier tasks unless they were turned off first. Encouraging the user to focus on a smaller number of tasks from the beginning (by not showing all the data at once) seems like a good strategy to not overwhelm the user. This is an assumption that would need to be confirmed in further studies.

Lastly, VR offers a unique opportunity to mix 2D and 3D visualizations in a natural interface that the user can interact with using their hands. In Luddy VR, we showed that a 2D bar graph of completion times helped the user identify their fastest and slowest performances and then inspect trajectories accordingly to derive new strategies for the second round of tasks. This setup shows the user **abstract data via a traditional, static data visualization** on the one hand, **and spatial data via VR** on the other hand. This way, designers can utilize the most effective graphic variables such as position and length (59, 101) for abstract variables (such as, in our case, completion time) while still allowing the user to see, in 3D, the geospatial data that the abstract data is based on while making use of the interaction afforded by VR.

A final note has to be written about the **superb accuracy** of the mouse as an input device. While 2D Desktop users lagged behind VR users in all performance metrics plus satisfaction, there were no significant differences in terms of position accuracy.

8.2 Major challenges

The application of VR for data visualization (as well as practical applications in a domain, such as tissue registration like we presented in chapter 5) poses a range of challenges.

First, setting up and using VR equipment still requires a considerable amount of time. High-end VR HMDs with external tracking (such as the HTC Vive used here) need tracking stations. And even if these HMDs have inside-out tracking, i.e., they have tracking sensors built in and do not rely on external stations, still have to be connected to a computer with a capable graphics cards to run VR applications at the desired resolution and frame rate. This results in a setup time and cost that far exceeds what is typically needed for a 2D setup. There is an increasing number of consumer-grade VR standalone HMDs without such hardware and setup requirements, e.g., the Oculus Quest 2 (https://www.oculus.com/quest-2/), see also our listing of Modern VR hardware in the Supporting Information). These devices trade in setup time for performance, but the user's transition between the physical and the virtual environment still provides a bottleneck for time that is not present in 2D Desktop applications. This is especially salient regarding the integration of VR into rigid existing workflows. For instance, in section 5.1.1, we presented a range of SOPs that capture steps needed to indicate the spatial origin of tissue samples. Implementing a VR version of the RUI into these procedures would necessitate increased time and training investment into an already rigid workflow.

Second, VR, by its very nature, requires space. In order to function as an input device, VR HMDs and controllers need to be moved. While the mouse as a 2D input device can be used on a fraction of the surface of a table (think about how small a mousepad is), VR spaces are typically around three by three meters (around nine by nine feet) large. While VR platforms typically provide visual indicators to warn users that come too close to the edge of their play space, a human facilitator is needed in research settings to ensure the physical safety of the user. This would be especially important in an environment that may not always be suited for wearing a VR HMD (such as wet bench labs with expensive equipment for anatomical procedures and medical imaging).

Third, another challenge for VR is cost, although there exists an ever-increasing amount of hardware for many budgets. The aforementioned Oculus Quest 2, at a price of $299 for the 64 GB version, is the latest addition to the growing roster of standalone VR HMDs that could accelerate the transition of VR to a mass medium. In section 8.3, we elaborate on future planned studies involving the Oculus Quest 2 and planned improvements to our user VR study methodology.

With regards to using data visualizations to improve performance in VR tasks, there are limitations to what be meaningfully visualized in VR. Both the cube-matching and movement tasks in this dissertation yielded spatial data by design, thus allowing us to create the Reflective phases as VR visualizations without major computation of derived variables in the first place. However, tasks that are executed on 2D screens, cognitive tests on questionnaires, or similarly abstract data such as infection rates for a disease, stock market values, and health data would require a spatial component via mapping to then be visualized to VR. Our research does not cover these cases.

Additionally, while our user studies focused on measuring whether the user was able to derive strategies for better performance in the second set of tasks, optimization may not always be the main goal of such interventions. For example, a user going through VR training for a machine assembly task may not be interested in performing their completion time or accuracy but rather improve memorization by repetition. For these cases, inspecting one's own data in an interactive VR visualization yields limited results. While the data we collected in this dissertation was quantitative, in future studies, it could be illuminating to capture insights based on spoken user feedback, e.g., via think-aloud methods or semi-structured interviews.

Lastly, from a practical standpoint, using VR requires more physical involvement than using 2D interfaces, due to the fact that more senses are engaged and more body parts are used to generate input. At this point, it is hard to imagine how an 8-hour work day for a data analyst would like were they only to use VR for creating and reading visualization.

8.3 Future Work

We presented design recommendations and challenges for implementing VR visualizations to improve user performance based on the user studies in this dissertation. The research and development of these experiments was guided by the type of user to be tested eventually: members of the general population, without major concentrations in terms of data visualization literacy or expertise in data-related fields. One of the major paths we are envisioning for future work are thus **domain-specific applications**. For example, we deployed the RUI user study in chapter 5 mostly to students at Indiana University and residents of Bloomington, IN, and not the wet bench staff, medical imaging specialists, and biologists that would usually register tissue blocks with the RUI. This would present an ideal real-world use case for a study with a smaller sample size and a strong focus on acquiring qualitative as well as quantitative data, which could help us better gauge the advantages and disadvantages of VR for this specific task across a variety of labs.

Further, for improving VR performance through data visualization, it could be enlightening to design VR training with a Reflective phase component **in cooperation with an organization** such as the Indiana University Office of Capital Planning and Facilities (CPF, https://cpf.iu.edu/index.html) or the Naval Surface Warfare Center, Crane Division (https://www.navsea.navy.mil/Home/Warfare-Centers/NSWC-Crane/Who-We-Are/). Working with these entities could allow us to design tasks that are closer to what real-world workers need to learn, perform, and optimize, and could yield more insight needs of practical nature. For example, in our user studies, subjects had to derive strategies to improve accuracy and completion time for simple matching tasks, which are unlikely to occur in high-stakes training scenarios. For instance, when training an enlisted navy sailor for using a radar device to perform reconnaissance of an area of operation, a VR version of the sailor's environment could be built to practice a range of scenarios. A mix of 2D and 3D visualizations in VR could then help the sailor assess their performance post-session. However, developing these kinds of tasks with useful results necessitates insights into how these trainings are designed, run, and evaluated, necessitating a cooperation with an organization outside of the university lab holding mutual research and development interests.

Another application domain for VR visualizations could be **architecture, engineering, and construction**, where stakeholders such as university officials, facility managers, and other administrative staff may be interested in seeing different layers of data. In fact, the Luddy VR study in this dissertation was inspired by initial conversations with university staff about the use of data visualization to optimize energy usage in IU buildings with advanced sensors. The basic layer, like in our Luddy VR user study, would be the building as the base map with a dot density map or 3D heatmap of occupancy, visualized through the DVL-FW. This would provide palpable advantages over 2D visualizations as occupancy data could be seen in its spatial context. Next, sensor data for electricity and temperature could be added as an additional layer. This would allow the stakeholders to derive conclusions about the interplay between the number of users in the building, their trajectories, energy consumption, and, last but not least, cost. Of course, a mix of 2D and 3D visualizations where appropriate could optimize the stakeholder's gain of actionable insights into building management while utilizing the known strength of bar graphs, line graphs, and scatter graphs. For example, an automation engineer may be interested to learn which rooms are the least used, at what time of day, and at what occupancy levels in the rest of the building. A fire marshal may want to know the most used trajectories across a building to optimize safety inspections. Lastly, a third and final layer could be simulations. If, for example, a building is to be substantially renovated or expanded, an architect may not only want to view the 3D model with the planned modifications but also a projection of future usage. They could then optimize the location of power outlets, staircases, even windows for optimal distribution of light across the new or renewed parts of the building. Naturally, building and testing applications like this requires lots of data and telemetry already in place, as well as buy-in from stakeholders. Those types of research and development for real-world usage scenarios in increasingly data-driven fields hold a lot of potential, and could be especially promising for **mixing 2D and 3D visualizations inside a VR application** for an “information-rich virtual environment” (38).

Another potential stream of future projects could be dedicated to **innovating VR user study methodology**. All our user studies were executed in a lab setting under supervision of a researcher. Due to the low penetration rate of VR across the consumer market, many users do not currently have VR gear in their homes that would allow researchers to easily deploy VR user studies to subjects remotely. Further, because VR requires specialized hardware and dedicated physical space, the subject’s home environment plays a role in determining the integrity of how the experiment is executed across many users in as many living conditions. While there only is limited research on remote VR data collection (144), we envision that with the growing presence of standalone VR HMDs on the market, user studies “in the wild” (162) could become an increasingly promising field, scaling up the number of subjects while reducing the amount of hours spent on face-to-face interaction by researchers and collecting data from within a variety of living conditions. Further, the ever-larger presence of web-based VR frameworks, such as A-Frame (192) and WebXR (210), could greatly simplify the pipeline through which users can access VR content. Ideally, accessing any VR application should be as easy as clicking on a link in a web browser. Integrating these new VR deployment frameworks into research designs could yield valuable contributions to how VR user studies are planned, run, and evaluated.

Finally, a future goal is to **further advance the DVL-FW interaction typology** by implementing more interaction types for VR. For example, the newly added **visualize/encode** or **annotate** types could be prime example of useful interactions in VR visualizations. In our user studies, the visual encoding was predetermined during development, and subjects were unable to add pictorial or linguistic symbols to mark important features in the data. However, with the structure already in place, e.g., the 2D panel for the filter menus, adding an additional panel to set visual encodings would be a feasible feature. Likewise, users could place annotation marks into 3D space simply by placing them there with their controllers. Visualize/encode and annotate are substantial interaction types for data visualization and analysis, because they allow the user to customize the data display to suit the insight need currently in focus, and, in the case of annotate, to outsource insights from their memory to an external representation. It would then be interesting to compare implementations of these interaction types between setups, e.g., 2D screens and VR.

Finally, **other typologies** of the DVL-FW, especially those for **graphic symbols and graphic variables**, provide fertile ground for more user studies. The four-page spread of graphic symbol and variable pairings put together by Börner (31) (p.36-39) presents a range of combinations that are not applicable for visualizations that are static, 2D, or both. For example, the time-based, quantitative graphic variables of rhythm, velocity, and speed could be implemented in VR in a way that is not possible (or reasonable) in a scatter graph. Additionally, some graphic symbol and variable pairings in Börner's book are blank, e.g., the combination curvature and volume or angle and surface. Adding a third dimension makes these pairings feasible, and having three axes for rotation in VR would allow for a richer design space specifically for angle and curvature.

In conclusion, there is a large space of potential research avenues following the studies presented in this dissertation. One of the greatest challenges and, at the same time, most promising features of VR visualizations is their novelty. This novelty can result in an urge to visualize data in VR that would best be represented with more traditional means. Harmonizing the aesthetic appeal of colorful, complex data visualizations with the cognitive limits of the human mind and the physical strain of VR on its users is a paramount research challenge for the foreseeable future. At the same time, it will be fascinating to observe what kinds of user studies become possible with an ever-increasing amount of hardware options, used by ever-more diverse sets of users, in real-world scenarios as more and more organizations and individuals embrace this engaging medium.

References

1. 3Dconnexion. SpaceMouse; 2021 [cited 2021 Jul 15]. Available from: https://3dconnexion.com/uk/spacemouse/.

2. Abraham A, Su W. XV-1 Refugee Flow; 2019 [cited 2021 Jul 15]. Available from: http://www.scimaps.org/mapdetail/refugee_flow_233.

3. Ackerman MJ. The Visible Human project. Proceedings of the IEEE. 1998;86(3):504-511. doi: 10.1109/5.662875

4. Adsay NV, Basturk O, Saka B, Bagci P, Ozdemir D, Balci S, et al. Whipple made simple for surgical pathologists. The American Journal of Surgical Pathology. 2014;38(4):480-493. doi: 10.1097/pas.0000000000000165

5. Agarwal M, Srinivasan A, Stasko J. VisWall: Visual data exploration using direct combination on large touch displays. 2019 IEEE Visualization Conference (VIS); Vancouver, BC, Canada. Washington, DC: IEEE; 2019. p. 26-30. doi: 10.1109/VISUAL.2019.8933673

6. Ahlberg C, Shneiderman B. Visual information seeking: Tight coupling of dynamic query filters with starfield displays. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; Boston, MA. New York City, NY: ACM; 1994. p. 313-317. doi: 10.1145/191666.191775

7. Albrecht K, Barabási A-L. XIII-1 The Cosmic Web; 2017 [cited 2021 Jul 15]. Available from: http://www.scimaps.org/mapdetail/the_cosmic_web_208.

8. Allen J, Brewer M, Neumann E, Gutierrez D, deCaestecker M, Spraggins J. Freezing fresh tissue v.2 2020 [cited 2021 Jul 15]. Available from: https://www.protocols.io/view/freezing-fresh-tissue-6wghfbw.

9. Alper B, Riche NH, Chevalier F, Boy J, Sezgin M. Visualization literacy at elementary school. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems; Denver, CO. New York City, NY: ACM; 2017. p. 5485-5497. doi: 10.1145/3025453.3025877

10. Alphabet Inc. Google Maps: Transit & Food; 2021 [cited 2021 Jul 15]. Available from: https://apps.apple.com/us/app/google-maps-transit-food/id585027354.

11. Amar R, Eagan J, Stasko J. Low-level components of analytic activity in information visualization. IEEE Symposium on Information Visualization; Minneapolis, MN. Washington, DC: IEEE; 2005. p. 111-117. doi: 10.1109/INFVIS.2005.1532136

12. Amar R, Stasko J. A knowledge task-based framework for design and evaluation of information visualizations. IEEE Symposium on Information Visualization; 2004; Austin, TX. Washington, DC: IEEE; 2004. p. 143-49. doi: 10.1109/INFVIS.2004.10

13. Anderson A, Weng Z. VRDD: Applying virtual reality visualization to protein docking and design. Journal of Molecular Graphics and Modelling. 1999;17(3-4):180-186. doi: 10.1016/S1093-3263(99)00029-7

14. Anderson D, Neumann E, Allen J, Brewer M, Gutierrez D, Spraggins J. Cryostat sectioning of tissues for 3D multimodal molecular imaging; 2020 [cited 2021 Jul 15]. Available from: https://www.protocols.io/view/cryostat-sectioning-of-tissues-for-3d-multimodal-m-7ethjen.

15. Association of Directors of Anatomic and Surgical Pathology. Recommendations for the reporting of surgically resected thymic epithelial tumors. American Journal of Clinical Pathology. 2009;132(1):10-15. doi: 10.1309/AJCPHII8GRIG3JHV

16. Ausburn LJ, Ausburn FB. Visual literacy: Background, theory and practice. Programmed Learning and Educational Technology. 1978;15(4):291-297. doi: 10.1080/0033039780150405

17. Autodesk Inc. Maya 2020; 2021 [cited 2021 Jul 15]. Available from: https://www.autodesk.com/products/maya/overview.

18. Avgerinou MD. Towards a visual literacy index. Journal of Visual Literacy. 2007;27(1):29-46. doi: 10.1080/23796529.2007.11674644

19. Bach B, Dachselt R, Carpendale S, Dwyer T, Collins C, Lee B. Immersive analytics: Exploring future interaction and visualization technologies for data analytics. ACM International Conference on Interactive Surfaces and Spaces; Niagara Falls, ON, Canada. New York City, NY: ACM; 2016. p. 529-533. doi: 10.1145/2992154.2996365

20. Bach B, Sicat R, Beyer J, Cordeil M, Pfister H. The hologram in my hand: How effective is interactive exploration of 3D visualizations in immersive tangible augmented reality? IEEE Transactions on Visualization and Computer Graphics. 2017;24(1):457-467. doi: 10.1109/TVCG.2017.2745941

21. Bailey JH, Witmer BG. Learning and transfer of spatial knowledge in a virtual environment. Proceedings of the Human Factors and Ergonomics Society Annual Meeting; Nashville, TN. Los Angeles, LA: Sage Publication; 1994. p. 1158-1162. doi: 10.1177/154193129403801803

22. Balakrishnan R, Baudel T, Kurtenbach G, Fitzmaurice G. The Rockin'Mouse: Itegral 3D manipulation on a plane. Proceedings of the ACM SIGCHI Conference on Human factors in Computing Systems; Atlanta, GA, USA. New York City, NY, USA: ACM; 1997. p. 311–318. doi: 10.1145/258549.258778

23. Batch A, Cunningham A, Cordeil M, Elmqvist N, Dwyer T, Thomas BH, et al. There is no spoon: Evaluating performance, space use, and presence with expert domain users in immersive analytics. IEEE Transactions on Visualization and Computer Graphics. 2019;26(1):536-546. doi: 10.1109/TVCG.2019.2934803

24. Beaudouin-Lafon M. Designing interaction, not interfaces. Proceedings of the Working Conference on Advanced Visual Interfaces; Palermo, Italy. New York City, NY: ACM; 2004. p. 15-22. doi: 10.1145/989863.989865

25. Beaudouin-Lafon M. Instrumental interaction: An interaction model for designing post-WIMP user interfaces. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; The Hague, Netherlands. New York City, NY: ACM; 2000. p. 446-453. doi: 10.1145/332040.332473

26. Beheshti E, Kim D, Ecanow G, Horn MS. Looking inside the wires: Understanding museum visitor learning with an augmented circuit exhibit. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems; Denver, CO. New York City, NY: ACM; 2017. p. 1583–1594. doi: 10.1145/3025453.3025479

27. Benoît G. Introduction to information visualization: Transforming data into meaningful information. London, UK: Rowman & Littlefield; 2019.

28. Bertin J. Semiology of graphics: Diagrams, networks, maps. Seattle, WA: Esri Press; 1983.

29. Billinghurst M, Cordeil M, Bezerianos A, Margolis T. Collaborative Immersive Analytics. In: Marriott K, Schreiber F, Dwyer T, Klein K, Riche NH, Itoh T, et al., editors. Immersive Analytics. Cham, Switzerland: Springer Nature; 2018. p. 221-257.

30. Blender Foundation. Blender 2.91; 2021 [cited 2021 Jul 15]. Available from: https://www.blender.org/.

31. Börner K. Atlas of knowledge: Anyone can map. Cambridge, MA: MIT Press; 2015 2015.

32. Börner K, Bueckle A, Ginda M. Data Visualization Literacy: Definitions, conceptual frameworks, exercises, and assessments. Proceedings of the National Academy of Sciences. 2019;116(6):1857-1864. doi: 10.1073/pnas.1807180116

33. Börner K, Heimlich JE, Peppler K, Keune A, Kennedy B, Uzzo S. Make-A-Vis: Advancing public understanding of scientific data. PLoS ONE. Forthcoming 2021.

34. Börner K, Maltese A, Balliet RN, Heimlich J. Investigating aspects of data visualization literacy using 20 information visualizations and 273 science museum visitors. Information Visualization. 2016;15(3):198-213. doi: 10.1177/1473871615594652

35. Börner K, Polley DE. Visual insights: A practical guide to making sense of data. Cambridge, MA: MIT Press; 2014.

36. Börner K, Simpson AH, Bueckle A, Goldstone RL. Science map metaphors: a comparison of network versus hexmap-based visualizations. Scientometrics. 2018;114(2):409-426. doi: 10.1007/s11192-017-2596-3

37. Bostock M, Ogievetsky V, Heer J. D³ data-driven documents. IEEE Transactions on Visualization and Computer Graphics. 2011;17(12):2301-2309. doi: 10.1109/TVCG.2011.185

38. Bowman DA, North C, Chen J, Polys NF, Pyla PS, Yilmaz U. Information-rich virtual environments: Theory, tools, and research agenda. Proceedings of the ACM Symposium on Virtual Reality Software and Technology; Osaka, Japan. New York City, NY: ACM; 2003. p. 81-90. doi: 10.1145/1008653.1008669

39. Boy J, Rensink RA, Bertini E, Fekete J-D. A principled way of assessing visualization literacy. IEEE Transactions on Visualization and Computer Graphics. 2014;20(12):1963-1972. doi: 10.1109/TVCG.2014.2346984

40. Brehmer M, Munzner T. A multi-level typology of abstract visualization tasks. IEEE Transactions on Visualization and Computer Graphics. 2013;19(12):2376-2385. doi: 10.1109/TVCG.2013.124

41. Brewer M, Allen J, Neumann E, Fogo A, Harris R, Gutierrez D, et al. Collection and post-surgical excision of human kidney tissue through the cooperative human tissue network; 2020 [cited 2021 Jul 15]. Available from: https://www.protocols.io/view/collection-and-post-surgical-excision-of-human-kid-7gehjte.

42. Bryson S. Virtual reality in scientific visualization. Communications of the ACM. 1996;39(5):62–71. doi: 10.1145/229459.229467

43. Bueckle A, Buehling K, Shih PC, Börner K. Comparing completion time, accuracy, and satisfaction in virtual reality vs. desktop implementation of the Common Coordinate Framework Registration User Interface (CCF RUI); 2021 [cited 2021 Jul 15]. Available from: https://arxiv.org/abs/2102.12030.

44. Buja A, Cook D, Swayne DF. Interactive high-dimensional data visualization. Journal of Computational and Graphical Statistics. 1996;5(1):78-99. doi: 10.1080/10618600.1996.10474696

45. Buja A, McDonald JA, Michalak J, Stuetzle W. Interactive data visualization using focusing and linking. Proceedings of the 2nd Conference on Visualization; San Diego, CA. Washington, DC: IEEE; 1991. p. 156-163. doi: 10.1080/10618600.1996.10474696

46. Bukowski RW, Séquin CH. Object associations: a simple and practical approach to virtual 3D manipulation. Proceedings of the 1995 Symposium on Interactive 3D graphics; Monterey, CA, USA. New York City, NY, USA: ACM; 1995. p. 131–ff. doi: 10.1145/199404.199427

47. Büschel W, Chen J, Dachselt R, Drucker S, Dwyer T, Görg C, et al. Interaction for Immersive Analytics. In: Marriott K, Schreiber F, Dwyer T, Klein K, Riche NH, Itoh T, et al., editors. Immersive Analytics. Cham, Switzerland: Springer Nature; 2018. p. 95-138.

48. Buxton W. Human skills in interface design. New York City, NY: Wiley; 1994.

49. Buxton W. Lexical and pragmatic considerations of input structures. Computer Graphics. 1983;17(1):31-37. doi: 10.1145/988584.988586

50. Buxton W. There's more to interaction than meets the eye: Some issues in manual input. In: Norman DA, Draper SW, editors. User-Centered System Design: New Perspectives on Human-Computer Interaction. Hillsdale, NJ: Lawrence Erlbaum Associates; 1986. p. 122–137.

51. Buxton W, Myers B. A study in two-handed input. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; Boston, MA. New York City, NY: ACM; 1986. p. 321-326. doi: 10.1145/22339.22390

52. Card SK, Mackinlay JD, Robertson GG. The design space of input devices. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; Seattle, WA. New York City, NY: ACM; 1990. p. 117-124. doi: 10.1145/97243.97263

53. Card SK, Mackinlay JD, Robertson GG. A morphological analysis of the design space of input devices. ACM Transactions on Information Systems. 1991;9(2):99-122. doi: 10.1145/123078.128726

54. Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). COVID-19 Dashboard; 2020 [cited 2021 Jul 15]. Available from: https://gisanddata.maps.arcgis.com/apps/opsdashboard/index.html.

55. Chandler T, Cordeil M, Czauderna T, Dwyer T, Glowacki J, Goncu C, et al. Immersive analytics. 2015 Big Data Visual Analytics; Hobart, TAS, Australia: IEEE; 2015. p. 1-8. doi: 10.1109/BDVA.2015.7314296

56. Chuah MC, Roth SF. On the semantics of interactive visualizations. IEEE Symposium on Information Visualization. Washington, DC: IEEE; 1996. p. 29-36. doi: 10.1109/INFVIS.1996.559213

57. Chuah MC, Roth SF, Mattis J, Kolojejchick J. SDM: Selective dynamic manipulation of visualizations. ACM Symposium on User Interface Software and Technology; Pittsburgh, PA. New York City, NY: ACM; 1995. p. 61-70. doi: 10.1145/215585.215654

58. Cikajlo I, Potisk KP. Advantages of using 3D virtual reality based training in persons with Parkinson's disease: A parallel study. Journal of NeuroEngineering and Rehabilitation. 2019;16(1):1-14. doi: 10.1186/s12984-019-0601-1

59. Cleveland WS, McGill R. Graphical perception: Theory, experimentation, and application to the development of graphical methods. Journal of the American Statistical Association. 1984;79(387):531-554. doi: 10.1080/01621459.1984.10478080

60. Cordeil M, Cunningham A, Bach B, Hurter C, Thomas BH, Marriott K, et al. IATK: An immersive analytics toolkit. 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR); Osaka, Japan. Washington, DC: IEEE; 2019. p. 200-209. doi: 10.1109/VR.2019.8797978

61. Cordeil M, Cunningham A, Dwyer T, Thomas BH, Marriott K. ImAxes: Immersive axes as embodied affordances for interactive multivariate data visualisation. Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology. New York City, NY: ACM; 2017. p. 71-83. doi: 10.1145/3126594.3126613

62. Cordeil M, Dwyer T, Klein K, Laha B, Marriott K, Thomas BH. Immersive Collaborative Analysis of Network Connectivity: CAVE-style or Head-Mounted Display? IEEE Transactions on Visualization and Computer Graphics. 2017;23(1):441-450. doi: 10.1109/TVCG.2016.2599107

63. Coutrix C, Nigay L. Mixed reality: A model of mixed interaction. Proceedings of the Working Conference on Advanced Visual Interfaces; Venezia, Italy. New York City, NY: ACM; 2006. p. 43-50. doi: 10.1145/1133265.1133274

64. Crystal Dynamics, Feral Interactive, Cameron Suey, Eidos Montréal. Rise of the Tomb Raider [software]. Square Enix, Microsoft Studis, Feral Interactive. 2015 [cited 2021 Jul 15]. Available from: https://square-enix-games.com/en_US/games/rise-of-the-tomb-raider-20-year-celebration.

65. Cyberinfrastructure for Network Science Center. Common Coordinate Framework Registration User Interface 1.5 (RUI); 2021 [cited 2021 Jul 15]. Available from: https://hubmapconsortium.github.io/ccf-ui/rui/.

66. Cyberinfrastructure for Network Science Center. Exhibitions; 2020 [cited 2021 Jul 15]. Available from: http://www.scimaps.org/exhibitions.

67. Cyberinfrastructure for Network Science Center. Make-a-Vis: An online tool for creating visualizations; 2021 [cited 2021 Jul 15]. Available from: https://cns-iu.github.io/make-a-vis/.

68. Cyberinfrastructure for Network Science Center. Places & Spaces: Mapping science; 2019 [cited 2021 Jul 15]. Available from: http://www.scimaps.org.

69. Cyberinfrastructure for Network Science Center. Visual Analytics Certificate; 2021 [cited 2021 Jul 15]. Available from: https://visanalytics.cns.iu.edu/.

70. Czauderna T, Haga J, Kim J, Klapperstück M, Klein K, Kuhlen T, et al. Immersive Analytics applications in life and health sciences. In: Marriott K, Schreiber F, Dwyer T, Klein K, Riche NH, Itoh T, et al., editors. Immersive Analytics. Cham, Switzerland: Springer Nature; 2018. p. 289-330.

71. Detterbeck FC, Moran C, Huang J, Suster S, Walsh G, Kaiser L, et al. Which way is up? Policies and procedures for surgeons and pathologists regarding resection specimens of thymic malignancy. Journal of Thoracic Oncology. 2011;6(7):S1730-S1738. doi: 10.1097/JTO.0b013e31821ea567

72. Dix A, Ellis G. By chance: Enhancing interaction with large data sets through statistical sampling. Proceedings of the Working Conference on Advanced Visual Interfaces;

Trento, Italy. New York City, NY: ACM; 2002. p. 167-176. doi: 10.1145/1556262.1556289

73. Dix A, Ellis G. Starting simple: Adding value to static visualisation through simple interaction. Proceedings of the Working Conference on Advanced Visual Interfaces; L'Aquila, Italy. New York City, NY: ACM; 1998. p. 124-134. doi: 10.1145/948496.948514

74. Djorgovski S, Hut P, Knop R, Longo G, McMillan S, Vesperini E, et al. The MICA experiment: Astrophysics in virtual worlds; 2013 [cited 2021 Jul 15]. Available from: https://arxiv.org/abs/1301.6808.

75. Donalek C, Djorgovski SG, Davidoff S, Cioc A, Wang A, Longo G, et al. Immersive and collaborative data visualization using virtual reality platforms. IEEE International Conference on Big Data; Washington, DC2014. p. 609-614. doi: 10.1109/BigData.2014.7004282

76. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. The Lancet Infectious Diseases. 2020;20(5):533-534. doi: 10.1016/S1473-3099(20)30120-1

77. Dontnod Entertainment, Deck Nine, Feral Interactive. Life is Strange [software]. Square Enix. 2015 [cited 2021 Jul 15]. Available from: https://lifeisstrange.square-enix-games.com/en-us.

78. Doug Engelbart Institute. 1968 "Mother of All Demos" with Doug Engelbart & Team (1/3); 2017 [cited 2021 Jul 15]. Available from: https://www.youtube.com/watch?v=M5PgQS3ZBWA.

79. Dwyer T, Marriott K, Isenberg T, Klein K, Riche N, Schreiber F, et al. Immersive Analytics: An introduction. In: Marriott K, Schreiber F, Dwyer T, Klein K, Riche NH, Itoh T, et al., editors. Immersive Analytics. Cham, Switzerland: Springer Nature; 2018. p. 1-23.

80. EA Redwood Shores. Dead Space [software]. Electronic Arts. 2008 [cited 2021 Jul 15]. Available from: https://www.ea.com/games/dead-space.

81. Elmqvist N, Moere AV, Jetter H-C, Cernea D, Reiterer H, Jankun-Kelly T. Fluid interaction for information visualization. Information Visualization. 2011;10(4):327-340. doi: 10.1177/1473871611413180

82. Emmons SR, Light RP, Börner K. MOOC visual analytics: Empowering students, teachers, researchers, and platform developers of massively open online courses. Journal of the Association for Information Science and Technology. 2017;68(10):2350-2363. doi: 10.1002/asi.23852

83. Entertainment O, Virtuos. The Outer Worlds [software]. Private Division. 2019 [cited 2021 Jul 15]. Available from: https://outerworlds.obsidian.net/en.

84. Epic Games. Fortnite [software]. Epic Games, Warner Bros. Interactive Entertainment. 2017 [cited 2021 Jul 15]. Available from: https://www.epicgames.com/fortnite/en-US/home.

85. Epic Games Inc. Unreal Engine; 2021 [cited 2021 Jul 15]. Available from: https://www.unrealengine.com/en-US.

86. Febretti A, Nishimoto A, Thigpen T, Talandis J, Long L, Pirtle J, et al. CAVE2: A hybrid reality environment for immersive simulation and information analysis. The Engineering Reality of Virtual Reality 2013. Bellingham, WA: International Society for Optics and Photonics; 2013. p. 1-12. doi: 10.1117/12.2005484

87. Few S. Show me the numbers: Designing tables and graphs to enlighten. Oakland, CA: Analytics Press; 2012.

88. Firaxis Games, Aspyr. Civilization VI [software]. 2K Games. 2016 [cited 2021 Jul 15]. Available from: https://civilization.com/.

89. Fisher D, Meyer M. Making data visual: A practical guide to using visualization for insight. Sebastopol, CA: O'Reilly Media, Inc.; 2017.

90. Fishkin KP, Moran TP, Harrison BL. Embodied user interfaces: Towards invisible user interfaces. IFIP International Conference on Engineering for Human-Computer Interaction; Heraklion, Crete, Greece. Berlin, Heidelberg, Germany: Springer; 1998. p. 1-18. doi: 10.1007/978-0-387-35349-4_22

91. Foley JD, Van FD, Van Dam A, Feiner SK, Hughes JF, Hughes J, et al. Computer graphics: Principles and practice. Boston, MA: Addison-Wesley Professional; 1996.

92. Foley JD, Wallace VL, Chan P. The human factors of computer graphics interaction techniques. IEEE Computer Graphics and Applications. 1984;4(11):13-48. doi: 10.1109/MCG.1984.6429355

93. Fransecky RB, Debes JL. Visual literacy: A way to learn--a way to teach. Washington DC: Association for Educational Communications and Technology; 1972.

94. Friedman JH, Stuetzle W. John W. Tukey's work on interactive graphics. The Annals of Statistics. 2002;30(6):1629-1639. doi: 10.1214/aos/1043351250

95. Froehlich B, Hochstrate J, Skuk V, Huckauf A. The GlobeFish and the GlobeMouse: two new six degree of freedom input devices for graphics applications. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; Montréal, Québec, Canada. New York City, NY, USA: ACM; 2006. p. 191–199. doi: 10.1145/1124772.1124802

96. Frontier Developments. Elite Dangerous [software]. Frontier Developments pls. 2014 [cited 2021 Jul 15]. Available from: https://www.elitedangerous.com/.

97. Google. Tilt Brush [software]. Google. 2016 [cited 2021 Jul 15]. Available from: https://www.tiltbrush.com/.

98. Halushka M. Human heart tissue harvesting; 2021 [cited 2021 Jul 15]. Available from: https://www.protocols.io/view/human-heart-tissue-harvesting-brsvm6e6.

99. Hardawar D. 2020 is VR's make-or-break year; 2019 [cited 2021 Jul 15]. Available from: https://sg.news.yahoo.com/2019-12-24-vr-in-2020-oculus-quest-half-life-alyx.html.

100. Hattwig D, Bussert K, Medaille A, Burgess J. Visual literacy standards in higher education: New opportunities for libraries and student learning Libraries and the Academy. 2013;13(1):61–89. doi: 10.1353/pla.2013.0008

101. Heer J, Bostock M. Crowdsourcing graphical perception: Using Mechanical Turk to assess visualization design. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; Atlanta, GA. New York City, NY: ACM; 2010. p. 203-212. doi: 10.1145/1753326.1753357

102. Heer J, Bostock M, Ogievetsky V. A Tour through the visualization zoo. Communications of the ACM. 2010;8(5):20–30. doi: 10.1145/1743546.1743567

103. Heer J, Shneiderman B. Interactive dynamics for visual analysis. Communications of the ACM. 2012;55(4):45-54. doi: 10.1145/2133806.2133821

104. Henderson Jr DA, Card S. Rooms: The use of multiple virtual workspaces to reduce space contention in a window-based graphical user interface. ACM Transactions on Graphics. 1986;5(3):211-243. doi: 10.1145/24054.24056

105. Hornecker E, Buur J. Getting a grip on tangible interaction: A framework on physical space and social interaction. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; Montréal, Québec, Canada. New York City, NY: ACM; 2006. p. 437-446. doi: 10.1145/1124772.1124838

106. Huang B, Jiang B, Li H. An integration of GIS, virtual reality and the Internet for visualization, analysis and exploration of spatial data. International Journal of Geographical Information Science. 2001;15(5):439-456. doi: 10.1080/13658810110046574

107. Hurter C, Riche NH, Drucker SM, Cordeil M, Alligier R, Vuillemot R. FiberClay: Sculpting three dimensional trajectories to reveal structural insights. IEEE Transactions on Visualization and Computer Graphics. 2018;25(1):704-714. doi: 10.1109/TVCG.2018.2865191

108. Hurtienne J, Israel JH. Image schemas and their metaphorical extensions: Intuitive patterns for tangible interaction. Proceedings of the 1st International Conference on Tangible and Embedded Interaction; Baton Rouge, LA. New York City, NY: ACM; 2007. p. 127-134. doi: 10.1145/1226969.1226996

109. Hutchins EL, Hollan JD, Norman DA. Direct manipulation interfaces. Human-Computer Interaction. 1985;1(4):311-338. doi: 10.1207/s15327051hci0104_2

110. Ibrahim Z, Money AG. Computer mediated reality technologies: A conceptual framework and survey of the state of the art in healthcare intervention systems. Journal of Biomedical Informatics. 2019;90:103102. doi: 10.1016/j.jbi.2019.103102

111. id Software LLC. Doom [software]. id Software LLC. 1993 [cited 2021 Jul 15]. Available from: https://bethesda.net/en/store/product/DO1GNGPCBG01.

112. Jackson RL, Fagan E. Collaboration and learning within immersive virtual reality. Proceedings of the Third International Conference on Collaborative Virtual Environments; San Francisco, CA. New York City, NY: ACM; 2000. p. 83-92. doi: 10.1145/351006.351018

113. Jacob RJ, Girouard A, Hirshfield LM, Horn MS, Shaer O, Solovey ET, et al. Reality-based interaction: a framework for post-WIMP interfaces. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; Florence, Italy. New York City, NY: ACM; 2008. p. 201-210. doi: 10.1145/1357054.1357089

114. Jorgensen M. Case processing SOP (spleen) v.3; 2020 [cited 2021 Jul 15]. Available from: https://www.protocols.io/view/case-processing-sop-spleen-bc3kiykw.

115. Jorgensen M, Nick J. Case processing SOP for thymus; 2020 [cited 2021 Jul 15]. Available from: https://www.protocols.io/view/case-processing-sop-for-thymus-bbgmiju6.

116. Kaiser P, Vasak P, Suorineni F, Thibodeau D. New dimensions in seismic data interpretation with 3-D virtual reality visualization for burst-prone mines. Proceedings of the Sixth International Symposium on Rockburst and Seismicity in Mines Proceedings; Perth, WA, Australia. Perth, WA, Australia: Australian Centre for Geomechanics; 2005. p. 33-45. doi: 10.36487/ACG_repo/574_0.3

117. Keim DA. Information visualization and visual data mining. IEEE Transactions on Visualization and Computer Graphics. 2002;8(1):1-8. doi: 10.1109/2945.981847

118. Keim DA. Visual exploration of large data sets. Communications of the ACM. 2001;44(8):38-44. doi: 10.1145/381641.381656

119. Kim Y, Heer J. Assessing effects of task and data distribution on the effectiveness of visual encodings. Computer Graphics Forum. 2018;37(3):157-167. doi: 10.1111/cgf.13409

120. Klemmer SR, Hartmann B, Takayama L. How bodies matter: Five themes for interaction design. Proceedings of the 6th Conference on Designing Interactive systems; University Park, PA. New York City, NY: ACM; 2006. p. 140-149. doi: 10.1145/1142405.1142429

121. Kosslyn SM, Ball TM, Reiser BJ. Visual images preserve metric spatial information: Evidence from studies of image scanning. Journal of Experimental Psychology: Human Perception and Performance. 1978;4(1):47-60. doi: 10.1037//0096-1523.4.1.47

122. Kwon O, Muelder C, Lee K, Ma K. A Study of layout, rendering, and interaction methods for immersive graph visualization. IEEE Transactions on Visualization and Computer Graphics. 2016;22(7):1802-1815. doi: 10.1109/TVCG.2016.2520921

123. Lanier J. You are not a gadget: A manifesto. New York City, NY, USA: Vintage Books; 2011.

124. Lee B, Isenberg P, Riche NH, Carpendale S. Beyond mouse and keyboard: Expanding design considerations for information visualization interactions. IEEE Transactions on Visualization and Computer Graphics. 2012;18(12):2689-2698. doi: 10.1109/TVCG.2012.204

125. Lee KK, Leung MC, Wong KH, Chang MY. A hand-held 3D display system that facilitates direct manipulation of 3D virtual objects. Proceedings of the 8th International Conference on Virtual Reality Continuum and its Applications in Industry; Yokohama, Japan. New York City, NY, USA: ACM; 2009. p. 65–70. doi: 10.1145/1670252.1670268

126. Lee S, Kim S-H, Hung Y-H, Lam H, Kang Y-a, Yi JS. How do people make sense of unfamiliar visualizations? A grounded model of novice's information visualization sensemaking. IEEE Transactions on Visualization and Computer Graphics. 2016;22(1):499-508. doi: 10.1109/TVCG.2015.2467195

127. Lee S, Kim S-H, Kwon BC. VLAT: Development of a visualization literacy assessment test. IEEE Transactions on Visualization and Computer Graphics. 2017;23(1):551-560. doi: 10.1109/TVCG.2016.2598920

128. Leung YK, Apperley MD. A review and taxonomy of distortion-oriented presentation techniques. ACM Transactions on Computer-Human Interaction. 1994;1(2):126-160. doi: 10.1145/180171.180173

129. Mackinlay J. Automating the design of graphical presentations of relational information. ACM Transactions on Graphics. 1986;5(2):110-141. doi: 10.1145/22949.22950

130. Mackinlay J, Card SK, Robertson GG. A semantic analysis of the design space of input devices. Human-Computer Interaction. 1990;5(2):145-190. doi: 10.1080/07370024.1990.9667153

131. Maltese AV, Harsh JA, Svetina D. Data visualization literacy: Investigating data interpretation along the novice—expert continuum. Journal of College Science Teaching. 2015;45(1):84-90.

132. Mann S. Mediated reality with implementations for everyday life; 2002 [cited 2021 Jul 15]. Available from: http://wearcam.org/presence_connect/.

133. Marchevsky AM, Hammond MEH, Moran C, Suster S. Protocol for the examination of specimens from patients with thymic epithelial tumors located in any area of the mediastinum. Archives of Pathology & Laboratory Medicine. 2003;127(10):1298-1303. doi: 10.5858/2003-127-1298-PFTEOS

134. Marriott K, Schreiber F, Dwyer T, Klein K, Riche NH, Itoh T, et al. Immersive Analytics. Cham, Switzerland: Springer Nature; 2018.

135. Masliah MR, Milgram P. Measuring the allocation of control in a 6 degree-of-freedom docking experiment. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; The Hague, The Netherlands. New York City, NY, USA: ACM; 2000. p. 25–32. doi: 10.1145/332040.332403

136. McCormack J, Roberts JC, Bach B, Freitas CDS, Itoh T, Hurter C, et al. Multisensory Immersive Analytics. In: Marriott K, Schreiber F, Dwyer T, Klein K, Riche NH, Itoh T, et al., editors. Immersive Analytics. Cham, Switzerland: Springer Nature; 2018. p. 57-94.

137. McKight PE, Najab J. Kruskal-Wallis Test. 2010. The Corsini Encyclopedia of Psychology. Hoboken, NJ, USA: John Wiley & Sons. doi: 10.1002/9780470479216.corpsy0491

138. Mendes D, Sousa M, Lorena R, Ferreira A, Jorge J. Using custom transformation axes for mid-air manipulation of 3D virtual objects. Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology; Gothenburg, Sweden. New York City, NY, USA: ACM; 2017. p. 1-8. doi: 10.1145/3139131.3139157

139. Milgram P, Kishino F. A taxonomy of mixed reality visual displays. IEICE Transactions on Information and Systems. 1994;77(12):1321-1329.

140. Millais P, Jones SL, Kelly R. Exploring data in virtual reality: Comparisons with 2D data visualizations. Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems; Montréal, Québec, Canada. New York City, NY: ACM; 2018. p. LBW007, 1-6. doi: 10.1145/3170427.3188537

141. Mizell D, Jones S, Jackson P, Pickett D. Is VR better than a workstation? A report on human performance experiments in progress. Selected Papers of the Eurographics Workshops in Barcelona, Spain, 1993, and Monte Carlo, Monaco, 1995. Vienna, Austria: Springer Vienna; 1995. p. 1-7. doi: 10.1007/978-3-7091-9433-1_1

142. Mobius Digital. Outer Wilds [software]. Annapurna Interactive. 2019 [cited 2021 Jul 15]. Available from: https://www.mobiusdigitalgames.com/outer-wilds.html.

143. Molina MA, Snell S, Franceschi D, Jorda M, Gomez C, Moffat FL, et al. Breast specimen orientation. Annals of Surgical Oncology. 2009;16(2):285-288. doi: 10.1245/s10434-008-0245-z

144. Mottelson A, Hornbæk K. Virtual reality studies outside the laboratory. Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology; Gothenburg, Sweden. New York City, NY, USA: ACM; 2017. p. 1-10. doi: 10.1145/3139131.3139141

145. Mullins PM, Treu S. A task-based cognitive model for user-network interaction: Defining a task taxonomy to guide the interface designer. Interacting with Computers. 1993;5(2):139-166. doi: 10.1016/0953-5438(93)90016-M

146. Munzner T. Visualization analysis and design. Boca Raton, FL: CRC Press; 2014.

147. Newman M. Maps of the 2012 US presidential election results; 2012 [cited 2021 Jul 15]. Available from: http://www-personal.umich.edu/~mejn/election/2012/.

148. Newman M. Maps of the 2016 US presidential election results; 2016 [cited 2021 Jul 15]. Available from: http://www-personal.umich.edu/~mejn/election/2016/.

149. Nielsen J. Noncommand user interfaces. Communications of the ACM. 1993;36(4):82-100. doi: 10.1145/255950.153582

150. OECD. PISA 2015 draft mathematics framework. Paris, France: OECD Publishing; 2013.

151. OECD. PISA 2015 draft reading literacy framework. Paris, France: OECD Publishing; 2013.

152. Pfeuffer K, Geiger MJ, Prange S, Mecke L, Buschek D, Alt F. Behavioural biometrics in VR: Identifying people from body motion and relations in virtual reality. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems; Glasgow, UK. New York City, NY: ACM; 2019. p. 1-12. doi: 10.1145/3290605.3300340

153. Pinker S. A theory of graph comprehension. In: Freedle R, editor. Artificial intelligence and the future of testing. Hillsdale, N.J: L. Erlbaum Associates; 1990. p. 73-126.

154. Prabhat, Forsberg A, Katzourin M, Wharton K, Slater M. A comparative study of desktop, fishtank, and cave systems for the exploration of volume rendered confocal data sets. IEEE Transactions on Visualization and Computer Graphics. 2008;14(3):551-563. doi: 10.1109/TVCG.2007.70433

155. Quercia D, Schifanella R, Aiello LM. XII-1 Smelly Maps; 2016 [cited 2021 Jul 15]. Available from: http://www.scimaps.org/mapdetail/smelly_maps_189.

156. Radonic T, Dickhoff C, Mino-Kenudson M, Lely R, Paul R, Thunnissen E. Gross handling of pulmonary resection specimen: maintaining the 3-dimensional orientation. Journal of Thoracic Disease. 2018;11(1):S37-S44. doi: 10.21037/jtd.2018.12.36

157. Rao R, Card SK. The table lens: Merging graphical and symbolic representations in an interactive focus+context visualization for tabular information. Proceedings of the SIGCHI Conference on Human factors in Computing Systems; Boston, MMA. New York City, NY: ACM; 1994. p. 318-322. doi: 10.1145/191666.191776

158. Reyna VF, Nelson WL, Han PK, Dieckmann NF. How numeracy influences risk comprehension and medical decision making. Psychological Bulletin. 2009;135(6):943-973. doi: 10.1037/a0017327

159. Roberts JC. Chapter 8: Exploratory visualization with multiple linked views. In: Dykes J, MacEachren AM, Kraak M-J, editors. Exploring Geovisualization. Oxford, UK: Elsevier; 2005. p. 159-180.

160. Robertson G, Card SK, Mackinlay JD. The cognitive coprocessor architecture for interactive user interfaces. Proceedings of the 2nd Annual ACM SIGGRAPH Symposium on User Interface Software and Technology; Williamsburg, VA. New York City, NY: ACM; 1989. p. 10-18. doi: 10.1145/73660.73662

161. Robertson GG, Card SK, Mackinlay JD. Information visualization using 3D interactive animation. Communications of the ACM. 1993;36(4):56-72. doi: 10.1145/108844.109003

162. Rogers Y, Marshall P. Research in the wild. San Rafael, CA: Morgan & Claypool Publishers; 2017.

163. Roth R. An empirically-derived taxonomy of interaction primitives for interactive cartography and geovisualization. IEEE Transactions on Visualization and Computer Graphics. 2013;19(12):2356-2365. doi: 10.1109/TVCG.2013.130

164. Salamon D. Deceased donor kidney and ureter tissue retrieval protocol; 2021 [cited 2021 Jul 15]. Available from: https://www.protocols.io/view/deceased-donor-kidney-and-ureter-tissue-retrieval-bhgtj3wn.

165. Salamon D. Surgical nephrectomy kidney and ureter tissue retrieval protocol; 2021 [cited 2021 Jul 15]. Available from: https://protocols.io/view/surgical-nephrectomy-kidney-and-ureter-tissue-retr-bhgzj3x6.

166. Satyanarayan A, Moritz D, Wongsuphasawat K, Heer J. Vega-lite: A grammar of interactive graphics. IEEE Transactions on Visualization and Computer Graphics. 2017;23(1):341-350. doi: 10.1109/TVCG.2016.2599030

167. Satyanarayan A, Russell R, Hoffswell J, Heer J. Reactive Vega: A streaming dataflow architecture for declarative interactive visualization. IEEE Transactions on Visualization and Computer Graphics. 2015;22(1):659-668. doi: 10.1109/TVCG.2015.2467091

168. Schmidt R, Singh K, Balakrishnan R. Sketching and composing widgets for 3D manipulation. Computer Graphics Forum. 2008;27(2):301-310. doi: 10.1111/j.1467-8659.2008.01127.x

169. Sci2 Team. Science of Science (Sci2) Tool [software]. Indiana University and SciTech Strategies. 2009 [cited 2021 Jul 15]. Available from: http://sci2.cns.iu.edu.

170. Seay AF, Krum DM, Hodges L, Ribarsky W. Simulator sickness and presence in a high FOV virtual environment. IEEE Virtual Reality. Washington, DC: IEEE; 2001. p. 299-300. doi: 10.1109/VR.2001.913806

171. Shah P. A model of the cognitive and perceptual processes in graphical display comprehension. In: Anderson M, editor. Reasoning with Diagrammatic Representations. Menlo Park, CA: AAAI Press; 1997. p. 94-101.

172. Shedroff N, Noessel C. Make it so: Interaction design lessons from science fiction. Brooklyn, NY: Rosenfeld Media; 2012.

173. Sherman WR, Craig AB. Understanding virtual reality: Interface, application, and design. Cambridge, MA: Morgan Kaufmann; 2018.

174. Shiryaev D. [4k, 60 fps] Arrival of a Train at La Ciotat (The Lumière Brothers, 1896); 2006 [cited 2021 Jul 15]. Available from: https://www.youtube.com/watch?v=3RYNThid23g.

175. Shneiderman B. Direct manipulation: A step beyond programming languages. Computer. 1983;16(8):57-69. doi: 10.1145/800276.810991

176. Shneiderman B. The eyes have it: A task by data type taxonomy for information visualizations. IEEE Symposium on Visual Languages; Boulder, CO. Washington, DC: IEEE; 1996. p. 336-343. doi: 10.1109/VL.1996.545307

177. Shneiderman B. Why not make interfaces better than 3D reality? IEEE Computer Graphics and Applications. 2003;23(6):12-15. doi: 10.1109/MCG.2003.1242376

178. Shneiderman B, Plaisant C. Strategies for evaluating information visualization tools: Multi-dimensional in-depth long-term case studies. Proceedings of the 2006 AVI Workshop on BEyond Time and Errors: Novel Evaluation Methods for Information Visualization; Venice, Italy. New York City, NY: ACM; 2006. p. 1-7. doi: 10.1145/1168149.1168158

179. Simpson M, Wallgrün JO, Klippel A, Yang L, Garner G, Keller K, et al. Immersive analytics for multi-objective dynamic integrated climate-economy (DICE) models. Proceedings of the 2016 ACM Companion on Interactive Surfaces and Spaces; Niagara Falls, ON, Canada. New York City, NY: ACM; 2016. p. 99-105. doi: 10.1145/3009939.3009955

180. Slater M, Linakis V, Usoh M, Kooper R. Immersion, presence and performance in virtual environments: An experiment with tri-dimensional chess. Proceedings of the ACM Symposium on Virtual Reality Software and Technology; Hong Kong. New York City, NY: ACM; 1996. p. 163-172. doi: 10.1145/3304181.3304216

181. Slater M, Usoh M, Steed A. Depth of presence in virtual environments. Presence: Teleoperators & Virtual Environments. 1994;3(2):130-144. doi: 10.1162/pres.1994.3.2.130

182. Slater M, Usoh M, Steed A. Taking steps: The influence of a walking technique on presence in virtual reality. ACM Transactions on Computer-Human Interaction. 1995;2(3):201-219. doi: 10.1145/210079.210084

183. Smets G, Overbeeke C, Stappers PJ. Designing in virtual reality: Perception-action coupling and form semantics. INTERACT'93 and CHI'93 Conference Companion on

Human Factors in Computing Systems; Amsterdam, Netherlands. New York City, NY: ACM; 1993. p. 11-12. doi: 10.1145/259964.259975

184. Snyder MP, Lin S, Posgai A, Atkinson M, Regev A, Rood J, et al. The human body at cellular resolution: the NIH Human Biomolecular Atlas Program. Nature. 2019;574(7777):187-192. doi: 10.1038/s41586-019-1629-x

185. Soh J, Choi Y, Park Y, Yang HS. User-friendly 3D object manipulation gesture using Kinect. Proceedings of the 12th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry; Hong Kong, Hong Kong. New York City, NY, USA: ACM; 2013. p. 231–234. doi: 10.1145/2534329.2534338

186. Spence R. Information visualization: Design for interaction. Upper Saddle River, NJ: Prentice-Hall, Inc.; 2007.

187. Spitzer V, Ackerman MJ, Scherzinger AL, Whitlock D. The visible human male: a technical report. Journal of the American Medical Informatics Association. 1996;3(2):118-130. doi: 10.1136/jamia.1996.96236280

188. Springmeyer RR, Blattner MM, Max NL. A characterization of the scientific data analysis process. Proceedings of the 3rd Conference on Visualization; Boston, MA. Washington, DC: IEEE; 1992. p. 235-242. doi: 10.1109/VISUAL.1992.235203

189. Steed A, Frlston S, Lopez MM, Drummond J, Pan Y, Swapp D. An ‘In the Wild’ experiment on presence and embodiment using consumer virtual reality equipment. IEEE Transactions on Visualization and Computer Graphics. 2016;22(4):1406-1414. doi: 10.1109/TVCG.2016.2518135

190. Steuer J. Defining virtual reality: Dimensions determining telepresence. Journal of Communication. 1992;42(4):73-93. doi: 10.1111/j.1460-2466.1992.tb00812.x

191. Stolte C, Tang D, Hanrahan P. Polaris: A system for query, analysis, and visualization of multidimensional relational databases. IEEE Transactions on Visualization and Computer Graphics. 2002;8(1):52-65. doi: 10.1109/2945.981851

192. Supermedium. A-Frame; 2021 [cited 2021 Jul 15]. Available from: https://aframe.io/.

193. Tanriverdi V, Jacob RJK. VRID: A design model and methodology for developing virtual reality interfaces. Proceedings of the ACM Symposium on Virtual Reality Software and Technology; Banff, AB, Canada. New York City, NY: ACM; 2001. p. 175–182. doi: 10.1145/505008.505042

194. Trickett SB, Trafton JG. Toward a comprehensive model of graph comprehension: Making the case for spatial cognition. International Conference on Theory and Application of Diagrams. Berlin, Heidelberg, Germany: Springer; 2006. p. 286-300. doi: 10.1007/11783183_38

195. Tufte ER. The visual display of quantitative information. Cheshire, CT: Graphics Press; 2001.

196. Tufte ER, Goeler NH, Benson R. Envisioning information. Cheshire, CT: Graphics Press 1990.

197. Tweedie L. Characterizing interactive externalizations. Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems; Atlanta, GA. New York City, NY: ACM; 1997. p. 375-382. doi: 10.1145/258549.258803

198. Uber vis.gl Visualization Team. Kepler.gl; 2021 [cited 2021 Jul 15]. Available from: https://kepler.gl/.

199. Ubisoft Montréal, Ubisoft Toronto. Far Cry 5 [software]. Ubisoft. 2018 [cited 2021 Jul 15]. Available from: https://www.ubisoft.com/en-us/game/far-cry/far-cry-5.

200. Unity Technologies. Unity real-time development platform; 2021 [cited 2021 Jul 15]. Available from: https://unity.com/.

201. Unwin A. Requirements for interactive graphics software for exploratory data analysis. Computational Statistics. 1999;14(1):7-22. doi: 10.1007/PL00022706

202. Usoh M, Arthur K, Whitton MC, Bastos R, Steed A, Slater M, et al. Walking>walking-in-place>flying, in virtual environments. Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques. New York City, NY: ACM; 1999. p. 359-364. doi: 10.1145/311535.311589

203. Valve Corporation. Half Life: Alyx [software]. Valve Corporation. 2020 [cited 2021 Jul 15]. Available from: https://www.half-life.com/en/alyx/.

204. Valve Corporation. openvr (GitHub code repository); 2021 [cited 2021 Jul 15]. Available from: https://github.com/ValveSoftware/openvr.

205. Valve Corporation. SteamVR; 2021 [cited 2021 Jul 15]. Available from: https://store.steampowered.com/app/250820/SteamVR/.

206. Van Dam A. Post-WIMP user interfaces. Communications of the ACM. 1997;40(2):63-67. doi: 10.1145/253671.253708

207. Van Dam A, Forsberg AS, Laidlaw DH, LaViola JJ, Simpson RM. Immersive VR for scientific visualization: A progress report. IEEE Computer Graphics and Applications. 2000;20(6):26-52. doi: 10.1109/38.888006

208. van Dam A, Laidlaw DH, Simpson RM. Experiments in immersive virtual reality for scientific visualization. Computers & Graphics. 2002;26(4):535-555. doi: 10.1016/S0097-8493(02)00113-9

209. Villar N, Izadi S, Rosenfeld D, Benko H, Helmes J, Westhues J, et al. Mouse 2.0: multi-touch meets the mouse. Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology; Victoria, BC, Canada. New York City, NY, USA: ACM; 2009. p. 33–42. doi: 10.1145/1622176.1622184

210. W3C Immersive Web Working Group. WebXR; 2021 [cited 2021 Jul 15]. Available from: https://immersiveweb.dev/.

211. Ward MO, Grinstein G, Keim D. Interactive data visualization: Foundations, techniques, and applications. Natick, MA: A. K. Peters Ltd.; 2015.

212. Ward MO, Yang J. Interaction spaces in data and information visualization. Sixth Joint Eurographics - IEEE TCVG Conference on Visualization; Konstanz, Germany. Goslar, Germany: Eurographics Association; 2004. p. 137-146.

213. Ware C. Visual thinking for design. Burlington, MA: Morgan Kaufmann; 2008.

214. Wickham H. A layered grammar of graphics. Journal of Computational and Graphical Statistics. 2010;19(1):3-28. doi: 10.1198/jcgs.2009.07098

215. Wilkinson L. The Grammar of Graphics (statistics and computing). New York City, NY: Springer Science+Business Media, Inc.; 2005.

216. Witmer BG, Bailey JH, Knerr BW, Parsons KC. Virtual spaces and real world places: Transfer of route knowledge. International Journal of Human-Computer Studies. 1996;45(4):413-428. doi: 10.1006/ijhc.1996.0060

217. Witmer BG, Singer MJ. Measuring presence in virtual environments: A presence questionnaire. Presence. 1998;7(3):225-240. doi: 10.1162/105474698565686

218. Yi JS, ah Kang Y, Stasko J, Jacko JA. Toward a deeper understanding of the role of interaction in information visualization. IEEE Transactions on Visualization and Computer Graphics. 2007;13(6):1224-1231. doi: 10.1109/TVCG.2007.70515

219. Yiing Lin, Lin S. HuBMAP tissue preservation protocol v2; 2020 [cited 2021 Jul 15]. Available from: https://www.protocols.io/view/hubmap-tissue-preservation-protocol-v2-bqm5mu86.

220. Zhou MX, Feiner SK. Visual task characterization for automated visual discourse synthesis. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; Los Angeles, CA. New York City, NY: ACM; 1998. p. 392-399. doi: 10.1145/274644.274698

Glossary

6 degrees of freedom (6-DoF): Indicates freedom of movement of a body in 3D space in position and rotation. Usually, each movement along or around the x, y, and z axes of a virtual environment has a name; right (along x axis towards +∞), left (along x axis towards -∞), up (along y axis towards +∞, down (along y axis towards -∞), forward (along z axis towards +∞), back ((along z axis towards -∞). For rotation: pitch (around x axis), yaw (around y axis), roll (around z axis).

Architecture, engineering, and construction (AEC): Umbrella term for creating, maintaining, and improving the built environment through the collaboration of three industries.

Data records: observations in a dataset.

Data Visualization Literacy Framework (DVL-FW): A structure of data visualization with a range of types (such as insight needs, graphic encodings, and interactions) to interpret, create, and teach data visualizations.

Field of view (FOV): The amount of 3D space in front of a camera captured on the screen, expressed as an angle in degrees. Commonly used to describe hardware properties of head-mounted displays (HMDs) or other display media.

Graphic symbols are of geometric, linguistic, or pictorial nature, and are used to visualize data records and variables.

Graphic variables, e.g., size, color, position, are properties of graphic symbols and can be used to encode data records and variables visually.

Head-mounted display (HMD)" A piece of hardware consisting of glasses with two built-in screens to present a stereoscopic view of a virtual environment. Usually attached to a user's head with a strap.

Level of detail (LOD): Indicates the complexity of a given 3D model depending on a multitude of factors (such as distance to the user, graphics settings, and graphics capability), usually for performance optimization in computer graphics.

Reference system: the basemap of a visualization, e.g., an x-y coordinate plane or a geographic map.

Software development kit (SDK): A collection of tools to allow individual creators to build high-level applications on top of a more low-level underlying software architecture.

User interface (UI): A collection of (graphical) elements with which a user can enact change or retrieve information from on the application, often in the form of buttons, sliders, text fields, etc.

Virtual reality (VR): A collection of hardware and software that allows for immersing a user in virtual environments through the use of a tracked head-mounted display (HMD), often with the user of controllers.

Visualization type: refers to tables, charts, graphs, maps, trees, and networks.

VR controller: A handheld device allowing user input through buttons and, depending on the model, a touchpad or an analog stick.

Windows, icons, menus, pointers (WIMP): A user interface paradigm popularized in the 1980s, optimized for 2D screens, using windows, icons, menus, and pointers.

Image Sources

If images in figures were taken from scholarly papers, citations were added in the captions or the text surrounding the figure in question. Otherwise, sources were listed below.

Figure 6

- A: https://cdn.landfallnavigation.com/media/catalog/product/cache/1/image/9df78eab33525d08d6e5fb8d27136e95/4/0/400_.jpg
- B: https://cdn.landfallnavigation.com/media/catalog/product/cache/1/image/9df78eab33525d08d6e5fb8d27136e95/2/5/25664_.jpg

Figure 9

- A: https://cdn.arstechnica.net/wp-content/uploads/2015/04/Engelbart-68-demo_0-2-640x426.jpg
- B: https://upload.wikimedia.org/wikipedia/commons/2/2b/Douglas_Engelbart%27s_prototype_mouse_-_Computer_History_Museum.jpg
- C: https://invention.si.edu/sites/default/files/blog-hintz-eric-2018-12-10-chord-keyset-mouse.jpg

Figure 10

- https://help.irisvr.com/hc/article_attachments/360067694094/Viewpoints_In_VR_IrisVR.png

Figure 11

- https://www.youtube.com/watch?v=3mRI1hu9Y3w

Figure 12

- A: https://store-images.s-microsoft.com/image/apps.43055.14303655336501012.2beb08d9-395e-453b-b5e3-0d4ac24d9d15.1208c626-fca0-4653-8be7-78aacd968001?w=672&h=378&q=80&mode=letterbox&background=%23FFE4E4E4&format=jpg
- B: https://cdn.mos.cms.futurecdn.net/C9oP55pgEzHCJegQAhyL3R.jpg

Figure 13

- https://www.youtube.com/watch?v=jxZe2r3opKY

Figure 14

- https://i.ytimg.com/vi/iqXxqV7haoU/maxresdefault.jpg

Figure 16

- A: https://i.ytimg.com/vi/tpv3hEfmB34/maxresdefault.jpg

- B:
  https://i.pinimg.com/originals/fb/c2/28/fbc228cdfeb606d7ff7a56d7b809e735.jpg

Figure 17

- https://cdn.arstechnica.net/wp-content/uploads/2016/04/Screenshot-99.png

Figure 18

- https://www.youtube.com/watch?v=F4S-YmzLfsE

Figure 19

- https://razinghel.com/wp-content/uploads/2016/11/production-1280x768.jpg

Figure 21

- https://www.youtube.com/watch?v=nIfZu1clbRg

Supporting Information

User study 1: RUI (IRB #1910331127)

Study information sheet

INDIANA UNIVERSITY STUDY INFORMATION SHEET FOR

Virtual Reality vs. Desktop Registration User Interface (IRB # 1910331127)

You are invited to participate in a research study of virtual reality (VR) vs. a more traditional 2D ("Desktop") interface. You were selected as a possible subject because you are 18+ years old. Please read this form and ask any questions you may have before agreeing to be in the study.

The study is being conducted by Dr. Katy Borner (katy@indiana.edu) and Andreas Bueckle (abueckle@indiana.edu) from the Luddy School of Informatics, Computing, and Engineering at Indiana University, and Kilian Buehling (kilian.buehling@tu-dresden.de) from the Technical University of Dresden in Germany. It is funded by the National Institutes of Health under OT2OD026671.

STUDY PURPOSE

The purpose of this study is to explore how users interact with and align 3D objects with each other. We want to know if there are differences in task completion time, accuracy, and user satisfaction between three conditions: a traditional "Desktop" interface, a VR interface where the user is standing and walking around ("VR Standup"), and a VR interface where the user is sitting at a desk ("VR Tabletop"). To that end, we are collecting data on timing and task accuracy alongside behavioral metrics (such as hand and head positions in VR as well as mouse position in Desktop) and user inputs such as button presses. We will also ask questions about the usability of the tools used across the three conditions. Please note that you have to be 18+ years old. People with an epilepsy diagnosis are not eligible.

NUMBER OF PEOPLE TAKING PART IN THE STUDY

If you agree to participate, you will be one of ~60 subjects who will be participating in this research.

PROCEDURES FOR THE STUDY

If you agree to be in the study, you will come to our research site during a previously agreed-upon timeslot. Then you will complete a pre-questionnaire to gather basic demographic information as well as information about your current usage and comfort with data visualizations, VR, and 3D environments. Subsequently, you will be assigned to one of our three conditions as per the researcher's discretion: Desktop (computer screen), VR Standup, or VR Tabletop. You will then be given instructions on how to use your tool, and then presented with a set of tasks. Finally, you will be given a post-questionnaire where you can share ideas for improvement. The study will take approximately 30-45 minutes. You will be recorded with audio and video, and we will log your actions in the physical world and in the virtual space for later analysis.

RISKS AND BENEFITS OF TAKING PART IN THE STUDY

The risks of participating in this research are discomfort answering questions about

unfamiliar visualizations. Further, some users can experience discomfort from using VR. Some users of VR headsets report motion sickness. Please be aware that you can terminate your participation in the study at any time. You may also tell the investigator if you need to take a break.

CONFIDENTIALITY

Efforts will be made to keep your personal information confidential. We cannot guarantee absolute confidentiality. Your personal information may be disclosed if required by law. Your identity will be held in confidence in reports in which the study may be published and databases in which results may be stored.

Organizations that may inspect and/or copy your research records for quality assurance and data analysis include groups such as the study investigator and his/her research associates, the Indiana University Institutional Review Board or its designees, the study sponsor, and (as allowed by law) state or federal agencies, specifically the Office for Human Research Protections (OHRP), the National Institutes of Health (NIH), etc., who may need to access your research records.

All research funded by the NIH is automatically granted a Certificate of Confidentiality. Information on these protections are described in the following paragraphs. Some of the details may sound odd in the context of this user study. However, we still want to fully inform you about these protections.

For the protection of your privacy, this research is covered by a Certificate of Confidentiality from the National Institutes of Health. The researchers may not disclose or use any information, documents, or specimens that could identify you in any civil, criminal, administrative, legislative, or other legal proceeding, unless you consent to it. Information, documents, or specimens protected by this Certificate may be disclosed to someone who is not connected with the research:

- if there is a federal, state, or local law that requires disclosure (such as to report child abuse or communicable diseases);
- if you consent to the disclosure, including for your medical treatment;
- if it is used for other scientific research in a way that is allowed by the federal regulations that protect research subjects;
- for the purpose of auditing or program evaluation by the government or funding agency;

A Certificate of Confidentiality does not prevent you from voluntarily releasing information about yourself. If you want your research information released to an insurer, medical care provider, or any other person not connected with the research, you must provide consent to allow the researchers to release it.

FUTURE USE

Information collected from you for this study may be used for future research studies or shared with other researchers for future research. If this happens, information which could identify you will be removed before any information or specimens are shared. Since identifying information will be removed, we will not ask for your additional consent.

PAYMENT

Upon completion of your participation in the study, you will receive a $5 Amazon.com gift card.

CONTACTS FOR QUESTIONS OR PROBLEMS

For questions about the study, please contact researcher Andreas Bueckle at abueckle@indiana.edu. For questions about your rights as a research participant or to discuss problems, complaints or concerns about a research study, or to obtain information, or offer input, contact the IU Human Subjects Office at 812-856-4242 or irb@iu.edu.

VOLUNTARY NATURE OF THIS STUDY

Taking part in this study is voluntary. You may choose not to take part or may leave the study at any time. Leaving the study will not result in any penalty or loss of benefits to which you are entitled. Your decision whether or not to participate in this study will not affect your current or future relations with the Luddy School of Informatics, Computing, and Engineering.

Data collection instruments

Qualtrics survey: https://github.com/cns-iu/rui-tissue-registration

Recruitment materials

Email

Subject line: Virtual Reality User Study Needs Participants - $5 Amazon Gift Card

Hello!

For a user study investigating 3D alignment capability in **virtual reality** and **2D screens**, **we would like to ask for your participation in a user study in Bloomington**. Your participation should take you about 30-45 minutes. You will receive a $5 Amazon gift card upon completion. For more information, please see the attached study information sheet (SIS).

We encourage you to share this survey with anyone who you may think could be interested in participating. You can sign up for the study under this link:

[INSERT LINK TO (IF NEEDED) + SHORT DESCRIPTION OF SCHEDULING INTERFACE]

About this study: This research is funded by the National Institutes of Health (NIH) grant #OT2OD026671, and has been approved under #1910331127 by the Institutional Review Board at Indiana University Bloomington. For questions about the study, please contact researcher Andreas Bueckle at abueckle@indiana.edu. For questions about your rights as a research participant or to discuss problems, complaints or concerns about a research study, or to obtain information, or offer input, contact the IU Human Subjects Office at 812-856-4242.
We thank you in advance for your time and contribution to our study.

Best wishes,
Andreas Bueckle, Katy Börner, Kilian Bühling (research team)

[ATTACH STUDY INFORMATION SHEET]

Social media

Facebook (lab account):

We are looking for participants in a user study on 3D alignment in virtual reality (VR) and 2D. The entire study takes place in Bloomington. Payment: $5 Amazon gift card for ~30-45 minutes of your time. If you are interested, please contact Andreas Bueckle at abueckle@indiana.edu. Please note that you have to be 18+. People with an epilepsy diagnosis are not eligible.

This study has approved under IRB #1910331127 by IU's Human Subjects Office. Please see this study information sheet for more info:

[INSERT LINK + SHORT DESCRIPTION TO SCHEDULING INTERFACE]

[INSERT LINK TO SIS HERE]

Facebook (personal account):

Hello Indiana friends! I am looking for participants in a user study about virtual reality and 2D (for my dissertation). $5 Amazon gift card for ~30-45 minutes of your time in Bloomington. DM me for details.

[INSERT LINK + SHORT DESCRIPTION TO SCHEDULING INTERFACE]

[INSERT LINK TO SIS HERE]

[INSERT PHOTO OF VR GEAR]

Twitter: (lab account):

Participants wanted for study on virtual reality (VR) in Bloomington. $5 Amazon gift card. Takes 30-45 mins. Email or DM Andreas at abueckle@indiana.edu. Study information sheet:

[INSERT LINK + SHORT DESCRIPTION TO SCHEDULING INTERFACE]

[INSERT LINK TO SIS HERE]

[INSERT PHOTO OF VR GEAR]

Snapchat (personal account):

Looking for VR research participants for my dissertation. $5 Amazon gift card for 30-45 mins of your time. Message me for details.

[(IF NEEDED) INSERT LINK + SHORT DESCRIPTION TO SCHEDULING INTERFACE]

[INSERT PHOTO OF VR GEAR]

Instagram (personal account):

Looking for VR research participants for my dissertation. $5 Amazon gift card for 30-45 mins of your time. Message me for details.

[INSERT LINK + SHORT DESCRIPTION TO SCHEDULING INTERFACE]

[INSERT PHOTO OF VR GEAR]

Generic social media post:

Looking for VR research participants for my dissertation. $5 Amazon gift card for 30-45 mins of your time. Message [INSERT CORRECT SOCIAL MEDIA HANDLE] for details.

[(IF NEEDED) INSERT LINK + SHORT DESCRIPTION TO SCHEDULING INTERFACE]

[INSERT PHOTO OF VR GEAR]

Verbal script

We are looking for participants in a user study on 3D alignment in virtual reality (VR) and 2D screens. Payment: $5 Amazon gift card for ~30-45 minutes of your time. This study has approved under IRB #1910331127 by IU's Human Subjects Office. If you want, I can send you more information.

[INSERT LINK + SHORT DESCRIPTION TO SCHEDULING INTERFACE + SIS if desired]

User study 2: RUI Reflective (IRB #1910331127)

Study information sheet

INDIANA UNIVERSITY STUDY INFORMATION SHEET FOR

Virtual Reality vs. Desktop Registration User Interface with Reflective Phase VR Intervention (IRB # 1910331127, Amendment 004)

You are invited to participate in a research study of virtual reality (VR) vs. a more traditional 2D ("Desktop") interface. You were selected as a possible subject because you are 18+ years old, and because you have not participated in this study previously**.** Please read this form and ask any questions you may have before agreeing to be in the study.

The study is being conducted by Dr. Katy Borner (katy@indiana.edu) and Andreas Bueckle (abueckle@indiana.edu) from the Luddy School of Informatics, Computing, and Engineering at Indiana University, and Kilian Buehling (kilian.buehling@tu-dresden.de) from the Technical University of Dresden in Germany. It is funded by the National Institutes of Health under OT2OD026671.

STUDY PURPOSE

The purpose of this study is to explore how users manipulate 3D objects and then optimize their behavior based on visualizations of their own data in VR. We want to know if there are differences in task completion time, accuracy, and user satisfaction between two cohorts: a control cohort that performs all the tasks in one go, and an experiment cohort that gets to inspect data of their own actions in VR before completing the second round of tasks ("Reflective Phase"). In this call for participants, we aim to recruit subjects for the experiment cohort only.

Additionally, we assign our subjects to one out of three conditions: a traditional "Desktop" interface, a VR interface where the user is standing and walking around ("VR Standup"), and a VR interface where the user is sitting at a desk ("VR Tabletop"). To that end, we are collecting data on timing and task accuracy alongside behavioral metrics (such as hand and head positions in VR as well as mouse position in Desktop) and user inputs such as button presses. We will also ask questions about the usability of the tools used across the three conditions. Please note that you have to be 18+ years old. People with an epilepsy diagnosis are not eligible.

PAYMENT

Upon completion of your participation in the study, you will receive $20 in Amazon.com gift cards.

NUMBER OF PEOPLE TAKING PART IN THE STUDY

If you agree to participate, you will be one of ~42 subjects who will be participating in this research.

PROCEDURES FOR THE STUDY

If you agree to be in the study, you will be handed a surgical mask upon arrival at the research site as needed and asked to wash your hands before the experiment. Further safety precautions may need to be implemented as needed, pending policy changes from IU, the Luddy School of Informatics, Computing, and Engineering, or other entities.

If you agree to be in the study, you will come to our research site during a previously agreed-upon timeslot. Then you will complete a pre-questionnaire to gather basic demographic information as well as information about your current usage and comfort with data visualizations, VR, and 3D environments. Subsequently, you will be assigned to one of our three conditions as per the researcher's discretion: Desktop (computer screen), VR Standup, or VR Tabletop. You will then be given instructions on how to use your tool, and then be presented with a set of tasks plus a brief intervention ("Reflective Phase") in VR. Finally, you will be given a post-questionnaire where you can share ideas for improvement. The study will take approximately 45 to 75 minutes. You will be recorded with audio and video, and we will log your actions in the physical world and in the virtual space for later analysis.

RISKS AND BENEFITS OF TAKING PART IN THE STUDY

The risks of participating in this research involve discomfort answering questions about unfamiliar visualizations. Further, some users can experience discomfort from using VR. Some users of VR headsets report motion sickness. Please be aware that you can terminate your participation in the study at any time. You may also tell the investigator if you need to take a break.

CONFIDENTIALITY

Efforts will be made to keep your personal information confidential. We cannot guarantee absolute confidentiality. Your personal information may be disclosed if required by law. Your identity will be held in confidence in reports in which the study may be published and databases in which results may be stored.

Organizations that may inspect and/or copy your research records for quality assurance and data analysis include groups such as the study investigator and his/her research associates, the Indiana University Institutional Review Board or its designees, the study sponsor, and (as allowed by law) state or federal agencies, specifically the Office for Human Research Protections (OHRP), the National Institutes of Health (NIH), etc., who may need to access your research records.

All research funded by the NIH is automatically granted a Certificate of Confidentiality. Information on these protections are described in the following paragraphs. Some of the details may sound odd in the context of this user study. However, we still want to fully inform you about these protections.

For the protection of your privacy, this research is covered by a Certificate of Confidentiality from the National Institutes of Health. The researchers may not disclose or use any information, documents, or specimens that could identify you in any civil, criminal, administrative, legislative, or other legal proceeding, unless you consent to it. Information, documents, or specimens protected by this Certificate may be disclosed to someone who is not connected with the research:

- if there is a federal, state, or local law that requires disclosure (such as to report child abuse or communicable diseases);
- if you consent to the disclosure, including for your medical treatment;
- if it is used for other scientific research in a way that is allowed by the federal regulations that protect research subjects;
- for the purpose of auditing or program evaluation by the government or funding agency.

A Certificate of Confidentiality does not prevent you from voluntarily releasing information about yourself. If you want your research information released to an insurer, medical care provider, or any other person not connected with the research, you must provide consent to allow the researchers to release it.

FUTURE USE

Information collected from you for this study may be used for future research studies or shared with other researchers for future research. If this happens, information which could identify you will be removed before any information or specimens are shared. Since identifying information will be removed, we will not ask for your additional consent.

CONTACTS FOR QUESTIONS OR PROBLEMS

For questions about the study, please contact researcher Andreas Bueckle at abueckle@indiana.edu. For questions about your rights as a research participant or to discuss problems, complaints or concerns about a research study, or to obtain information, or offer input, contact the IU Human Subjects Office at 812-856-4242 or irb@iu.edu.

VOLUNTARY NATURE OF THIS STUDY

Taking part in this study is voluntary. You may choose not to take part or may leave the study at any time. Leaving the study will not result in any penalty or loss of benefits to which you are entitled. Your decision whether or not to participate in this study will not affect your current or future relations with the Luddy School of Informatics, Computing, and Engineering.

Data collection instruments

Qualtrics survey: https://github.com/andreasbueckle/bueckle-dissertation-supporting-information/tree/main/rui_vr_reflective

Recruitment materials

Email

Subject line: Seeking Participants for a New Virtual Reality Study - $20 Amazon Gift Card

Hello!

We are looking for participants in a user study about interactive data visualization in virtual reality (VR). You will manipulate 3D objects in VR or on a laptop and then see if you can improve your behavior for the second round of tasks by inspecting your own data using a VR headset.

Compensation:

You will receive a $20 Amazon.com gift card upon completion.

Location:

Collaboration Space #4026 in Luddy Hall on the IU Bloomington campus.

How to sign up:

Please follow this link to enter your information:

https://iu.co1.qualtrics.com/jfe/form/SV_0eXoWGhB9Una6sR

Duration:

Your participation should take you ~45-75 minutes.

Eligibility:

You must be 18+.

People with an epilepsy diagnosis are not eligible.

People who participated in our recent user study on 3D alignment abilities in VR and on laptops are not eligible.

For direct questions, contact Andreas Bueckle (abueckle@indiana.edu).

Please share this survey with anyone who could be interested. Precautions will be taken to create a safe and sanitary research environment. For more information, please see the attached study information sheet (SIS).

Thank you in advance for your time and contribution to our study!

Andreas Bueckle, Katy Börner, Kilian Bühling

[ATTACH STUDY INFORMATION SHEET]

Social media

Personal accounts

Facebook:

Hello Indiana friends! I am looking for participants in another user study about virtual reality and 2D (for my dissertation). $20 in Amazon.com gift cards for ~45-75 minutes of your time in Bloomington.

If you are interested, please follow this link for more information:

[INSERT LINK TO SCHEDULING INTERFACE]

Of course, precautions will be taken to create a safe and sanitary research environment.

You can also DM me for details. Please note that if you participated in my most recent VR study (you know who you are), you'll have to sit this one out.

See you soon!

Snapchat:

Looking for VR research participants for my dissertation. $20 in Amazon.com gift cards for ~45-75 mins of your time. Message me for details.

[(IF NEEDED) INSERT LINK + SHORT DESCRIPTION TO SCHEDULING INTERFACE]

[INSERT PHOTO OF VR GEAR]

[ADD AS NEEDED: Precautions will be taken to create a safe and sanitary research environment. We will provide more information to you once you have signed up for a time slot. More information can also be found in the study information sheet.]

Instagram:

Looking for VR research participants for my dissertation$20 in Amazon.com gift cards for ~45-75 mins of your time. Message me for details.

[INSERT LINK + SHORT DESCRIPTION TO SCHEDULING INTERFACE]

[INSERT PHOTO OF VR GEAR]

[ADD AS NEEDED: Precautions will be taken to create a safe and sanitary research environment. We will provide more information to you once you have signed up for a time slot. More information can also be found in the study information sheet.]

Lab accounts

Facebook:

We are looking for participants in a user study on 3D alignment in virtual reality (VR) and 2D. Payment: $20 in Amazon.com gift cards for ~45-75 minutes of your time in Luddy Hall in Bloomington, IN. If you are interested, please follow this link to provide your information:

[INSERT LINK TO SCHEDULING INTERFACE]

Please note that if you participated in our recent study, you'll have to sit this one out.

For direct questions, email Andreas Bueckle (abueckle@indiana.edu).

Precautions will be taken to create a safe and sanitary research environment. Please note that you have to be 18+. People with an epilepsy diagnosis are not eligible.

This study has been approved under IRB #1910331127 by IU's Human Subjects Office.

Twitter:

Participants wanted for study on virtual reality (VR) in Bloomington. $20 in Amazon.com gift cards. Takes ~45-75 mins. If you are interested, please follow this link to provide your information:

[INSERT LINK TO SCHEDULING INTERFACE]

For questions, email or DM Andreas at abueckle@indiana.edu.

Instagram:

Participants wanted for study on virtual reality (VR) in Bloomington. $20 in Amazon.com gift cards. Takes ~45-75 mins. If you are interested, please follow this link to provide your information:

[INSERT LINK TO SCHEDULING INTERFACE]

For questions, email or DM Andreas at abueckle@indiana.edu.

Generic social media post:

[Statement of intent for user study involving VR]

[Payment, duration, location]

[scheduling interface + link to SIS as needed]

[Request to reach out to social media handle or email address for further questions]

[insert photo as needed]

[add as needed: Precautions will be taken to create a safe and sanitary research environment. We will provide more information to you once you have signed up for a time slot. More information can also be found in the study information sheet.]

Verbal script

We are looking for participants in a user study on 3D alignment in virtual reality (VR) and 2D screens with a VR intervention called “Reflective Phase”. Payment: $20 in Amazon.com gift cards for ~45-75 minutes of your time. If you want, I can send you more information.

[INSERT LINK + SHORT DESCRIPTION TO SCHEDULING INTERFACE + SIS if desired]

[ADD AS NEEDED: Precautions will be taken to create a safe and sanitary research environment. We will provide more information to you once you have signed up for a time slot. More information can also be found in the study information sheet. This study has been approved under IRB #1910331127 by IU’s Human Subjects Office.]

User study 3: Luddy Hall VR (IRB #1911941428)

Study information sheet

INDIANA UNIVERSITY STUDY INFORMATION SHEET FOR

Luddy Hall VR Navigation User Study (IRB #1911941428)

You are invited to participate in a research study of navigation in virtual reality (VR). You were selected as a possible subject because you are 18+ years old. Please note that subjects with an epilepsy diagnosis are not eligible. Please read this form and ask any questions you may have before agreeing to be in the study.

The study is being conducted by Andreas Bueckle (abueckle@indiana.edu), Dr. Patrick Shih (patshih@indiana.edu), and Dr. Katy Borner (katy@indiana.edu) from the Luddy School of Informatics, Computing, and Engineering at Indiana University.

STUDY PURPOSE

The purpose of this study is to explore how users navigate through a three-dimensional model of Luddy Hall on IU campus. We want to know if there are differences in completion time, task accuracy, and user satisfaction between a group of subjects who repeat the same set of navigation tasks twice (control) and a group of subjects that does so with an interventional treatment in-between (experiment). Additionally, we aim to understand what kinds of data visualizations can help users improve their own navigational performance in VR. To that end, we are collecting data on timing and task accuracy while also asking questions about the usability of the navigation methods presented in this user study.

NUMBER OF PEOPLE TAKING PART IN THE STUDY

If you agree to participate, you will be one of ~68 subjects who will participate in this research.

PROCEDURES FOR THE STUDY

If you agree to be in the study, you will be handed a surgical mask upon arrival at the research site (if needed) and asked to wash your hands before the experiment. Further safety precautions may need to be implemented, pending policy changes from IU, the Luddy School of Informatics, Computing, and Engineering, or other entities.

You will come to our research site during a previously agreed upon timeslot. Then, you will complete a pre-questionnaire to gather basic demographic data as well as information about your current usage and comfort with data visualizations, 3D applications, and VR. Subsequently, you will be given a VR headset and controllers. You will then receive instructions on how to use your tools and be presented with a set of tasks. If we select you for the control cohort, you will perform these tasks twice with a brief break in-between. If you are part of the experiment cohort, after the first part, we will show you a selection of data visualizations you generated with your movements in VR, and ask you questions about them while recording your answers, and then you will repeat the tasks. Finally, you will be presented with a post-questionnaire where you can share ideas for improvement. The study will take approximately 30-60 minutes of your time. You will be recorded in audio and video for the remainder of the experiment, starting the moment you enter the VR experiment, and we will record your actions in the physical world and in the virtual space with video and audio.

PAYMENT

Upon completion of your participation in the study, you will receive $20 in Amazon.com gift cards. In order to receive this payment, you need to complete the survey in its entirety.

RISKS AND BENEFITS OF TAKING PART IN THE STUDY

The risks of participating in this research are discomfort answering questions about unfamiliar visualizations. Further, some users can experience discomfort from using VR (such as motion sickness). You may also tell the investigator if you need to take a break.

CONFIDENTIALITY

Efforts will be made to keep your personal information confidential. We cannot guarantee absolute confidentiality. Your personal information may be disclosed if required by law. Your identity will be held in confidence in reports in which the study may be published and databases in which results may be stored.

Organizations that may inspect and/or copy your research records for quality assurance and data analysis include groups such as the study investigator and his/her research associates, the Indiana University Institutional Review Board or its designees, the study sponsor, and (as allowed by law) state or federal agencies, specifically the Office for Human Research Protections (OHRP), etc., who may need to access your research records.

FUTURE USE

Information collected from you for this study may be used for future research studies or shared with other researchers for future research. If this happens, information which could identify you will be removed before any information or specimens are shared. Since identifying information will be removed, we will not ask for your additional consent.

CONTACTS FOR QUESTIONS OR PROBLEMS

For questions about the study, please contact researcher Andreas Bueckle at abueckle@indiana.edu. For questions about your rights as a research participant or to discuss problems, complaints or concerns about a research study, or to obtain information, or to offer input, contact the IU Human Subjects Office at 812-856-4242 or irb@iu.edu.

VOLUNTARY NATURE OF THIS STUDY

Taking part in this study is voluntary. You may choose not to take part or may leave the study at any time. Your decision whether or not to participate in this study will not affect your current or future relations with the Luddy School of Informatics, Computing, and Engineering.

Data collection instruments

Qualtrics survey: https://github.com/andreasbueckle/bueckle-dissertation-supporting-information/tree/main/luddy_vr_reflective

Recruitment materials

Email

Subject line: Seeking Participants for a Luddy Hall Virtual Reality Navigation Study - $20 Amazon Gift Card

Hello!

We are again looking for participants in a virtual reality (VR) user study, this time about navigating buildings in VR. You traverse a virtual model of Luddy Hall on IU campus and perform various visualization-related tasks using a VR headset and controllers.

Compensation:

You will receive a $20 Amazon.com gift card upon completion.

Location:

Collaboration Space #4026 in Luddy Hall on the IU Bloomington campus.

How to sign up:

Please follow this link to enter your information:

[INSERT LINK TO SCHEDULING INTERFACE]

Duration:

Your participation should take you ~30-60 minutes.

Eligibility:

You must be 18+.

People with an epilepsy diagnosis are not eligible.

For direct questions, contact Andreas Bueckle (abueckle@indiana.edu).

Please share this survey with anyone who could be interested. Precautions will be taken to create a safe and sanitary research environment. For more information, please see the attached study information sheet (SIS).

Thank you in advance for your time and contribution to our study!

Andreas Bueckle, Patrick Shih, Katy Börner

[ATTACH STUDY INFORMATION SHEET]

Email to previous participants

This email is intended for subjects who participated in user study #1910331127 and who indicated that they would like to be contacted for future VR user studies.

Subject line: Seeking Participants for a Luddy Hall Virtual Reality Navigation Study - $20 Amazon Gift Card

Hello!

I am contacting you, because you were a participant in my last virtual reality (VR) user study and indicated that you would be interested in future VR user studies as well. That moment is here!

We are again looking for participants in a VR user study, this time about navigating buildings in VR. You traverse a virtual model of Luddy Hall on IU campus and perform various visualization-related tasks using a VR headset and controllers.

Compensation:

You will again receive a $20 Amazon.com gift card upon completion.

Location:

Collaboration Space #4026 in Luddy Hall on the IU Bloomington campus (same location as last time).

How to sign up:

Please follow this link to enter your information:

[INSERT LINK TO SCHEDULING INTERFACE]

Duration:

Your participation should take you ~30-60 minutes.

Eligibility:

You must be 18+.

People with an epilepsy diagnosis are not eligible.

For direct questions, contact Andreas Bueckle (abueckle@indiana.edu).

Please share this survey with anyone who could be interested. Precautions will be taken to create a safe and sanitary research environment. For more information, please see the attached study information sheet (SIS).

Thank you in advance for your time and contribution to our study!|

Andreas Bueckle, Patrick Shih, Katy Börner

[ATTACH STUDY INFORMATION SHEET]

Email to individuals who indicated interest in a previous study but did not participate

This email is intended for subjects who signed up to participate in user study #1910331127 and who did not get to participate due to time constraints, or because the subject slots were already full by the time they signed up. We thus offer them participation in a user study with

- *The same risk level (minimal)*
- *The same payment*
- *Around the same time investment*

Subject line: Seeking Participants for a Luddy Hall Virtual Reality Navigation Study - $20 Amazon Gift Card

Hello!

I am contacting you, because you indicated interest in my last virtual reality (VR) user study (IRB #1910331127) and did not get into the subject pool as we were filled up before you could participate. Now, there is another chance for you to be a research subject for us, with the same payment and about the same time investment, see below.

We are again looking for participants in a VR user study, this time about navigating buildings in VR. You traverse a virtual model of Luddy Hall on IU campus and perform various visualization-related tasks using a VR headset and controllers.

Compensation:

You will again receive a $20 Amazon.com gift card upon completion.

Location:

Collaboration Space #4026 in Luddy Hall on the IU Bloomington campus [ADD AS NEEDED: (same location as last time)].

How to sign up:

Please follow this link to enter your information:

[INSERT LINK TO SCHEDULING INTERFACE]

Duration:

Your participation should take you ~30-60 minutes.

Eligibility:

You must be 18+.

People with an epilepsy diagnosis are not eligible.

For direct questions, contact Andreas Bueckle (abueckle@indiana.edu).

Please share this survey with anyone who could be interested. Precautions will be taken to create a safe and sanitary research environment. For more information, please see the attached study information sheet (SIS).

Thank you in advance for your time and contribution to our study!

Andreas Bueckle, Patrick Shih, Katy Börner

[ATTACH STUDY INFORMATION SHEET]

Social media

Personal accounts

Facebook (personal account):

Hello! One last time, I am looking for participants in yet another user study about virtual reality for my dissertation; this time it's all about walking and flying through virtual buildings. $20 in Amazon.com gift cards for ~30-60 minutes of your time in Bloomington.

If you are interested, please follow this link for more information:

[INSERT LINK TO SCHEDULING INTERFACE]

[AS NEEDED: INSERT PHOTO OF VR GEAR]

Of course, precautions will be taken to create a safe and sanitary research environment.

You can also DM me for details (or email me at abueckle@indiana.edu). Everyone is eligible unless they are not 18 or have had an epilepsy diagnosis.

See you soon!

Snapchat:

Looking for VR research participants for my dissertation. $20 in Amazon.com gift cards for ~30-60 mins of your time. Message me for details.

[INSERT LINK TO SCHEDULING INTERFACE]

[AS NEEDED: INSERT PHOTO OF VR GEAR]

[ADD AS NEEDED: Precautions will be taken to create a safe and sanitary research environment. We will provide more information to you once you have signed up for a time slot. More information can also be found in the study information sheet.]

Instagram:

Looking for VR research participants for my dissertation (it's all about walking and flying through virtual buildings). $20 in Amazon.com gift cards for ~30-60 mins of your time. Message me for details.

[INSERT LINK TO SCHEDULING INTERFACE]

[AS NEEDED: INSERT PHOTO OF VR GEAR]

[ADD AS NEEDED: Precautions will be taken to create a safe and sanitary research environment. We will provide more information to you once you have signed up for a time slot. More information can also be found in the study information sheet.]

Lab accounts

Facebook:

We are looking for participants for a new user study on walking and flying through virtual buildings. Payment: $20 in Amazon.com gift cards for ~30-60 minutes of your time in Luddy Hall in Bloomington, IN. If you are interested, please follow this link to provide your information:

[INSERT LINK TO SCHEDULING INTERFACE]

Everyone is eligible unless they are not 18 or have had an epilepsy diagnosis.

For direct questions, email Andreas Bueckle (abueckle@indiana.edu).

Precautions will be taken to create a safe and sanitary research environment. Please note that you have to be 18+. People with an epilepsy diagnosis are not eligible.

This study has been approved under IRB #1910331127 by IU's Human Subjects Office.

Twitter [WILL BE ADJUSTED AS NEEDED TO CONFORM TO CHARACTER LIMITATIONS]:

Participants wanted for a new study about walking and flying through buildings in virtual reality (VR) in Bloomington. $20 in Amazon.com gift cards. Takes ~30-60 mins. If you are interested, please follow this link to provide your information:

[INSERT LINK TO SCHEDULING INTERFACE]

For questions, email or DM Andreas at abueckle@indiana.edu.

Instagram:

Participants wanted for a new study about walking and flying through buildings in virtual reality (VR) in Bloomington. $20 in Amazon.com gift cards. Takes ~30-60 mins. If you are interested, please follow this link to provide your information:

[INSERT LINK TO SCHEDULING INTERFACE]

For questions, email or DM Andreas at abueckle@indiana.edu.

Generic social media post:

[Statement of intent for user study involving VR]

[Payment, duration, location]

[scheduling interface + link to SIS as needed]

[Request to reach out to social media handle or email address for further questions]

[insert photo as needed]

[add as needed: Precautions will be taken to create a safe and sanitary research environment. We will provide more information to you once you have signed up for a time slot. More information can also be found in the study information sheet.]

Luddy Hall monitors

[INSERT GRAPHIC/HEADER]

Virtual Reality Study Needs Participants - $20 Amazon.com Gift Card

We are looking for participants in a user study on navigation in virtual reality (VR).

Payment: $20 in Amazon.com gift cards for ~30-60 minutes of your time here in Luddy Hall.

If interested, go to [INSERT LINK TO SCHEDULING INTERFACE] or scan the QR code below. Precautions will be taken to create a safe and sanitary research environment.

Eligibility: You must be 18+. People with an epilepsy diagnosis are not eligible.

If you have questions about your eligibility, please contact us at abueckle@indiana.edu.

[INSERT QR CODE]

[INSERT MORE GRAPHICS AS NEEDED]

Verbal script

We are looking for participants in a user study on navigation in 3D environments in virtual reality (VR). The entire study takes place in Luddy Hall on IU campus.
Payment: $20 Amazon gift card for ~30-60 minutes of your time.

[ADD AS NEEDED: This study has approved under IRB #1911941428 by IU's Human Subjects Office.]

If you want, I can send you more information.

[INSERT LINK TO SCHEDULING INTERFACE + SIS as needed]

[ADD AS NEEDED: Precautions will be taken to create a safe and sanitary research environment. We will provide more information to you once you have signed up for a time slot. More information can also be found in the study information sheet.]

Technical walkthroughs and code for RUI and Luddy Hall VR user studies

Video demos: https://github.com/andreasbueckle/bueckle-dissertation-supporting-information#video-demos-of-the-three-setups-used-in-the-luddy-vr-study-user-study-3

Reflective phase:

- RUI: https://github.com/andreasbueckle/bueckle-dissertation-supporting-information#technical-walkthrough-for-rui-reflective-study
- Luddy: https://github.com/andreasbueckle/bueckle-dissertation-supporting-information#technical-walkthrough-for-luddy-vr-reflective-study

Video demos of all VR applications used in this dissertation

For chapter 5: https://github.com/cns-iu/rui-tissue-registration#video-demos-of-the-three-setups

For chapter 6: https://github.com/andreasbueckle/bueckle-dissertation-supporting-information#video-demos-of-the-three-setups-used-in-the-rui-reflective-study-user-study-2

For chapter 7: https://github.com/andreasbueckle/bueckle-dissertation-supporting-information#video-demos-of-the-three-setups-used-in-the-luddy-vr-study-user-study-3

Modern VR hardware

Table 17. Overview of selected current consumer-grade VR HMDs with specs and supported interaction input as of 7/16/2021. The system used in this dissertation is ***bold****.*

Device	**HTC Vive**	HTC Vive Pro	Valve Index	HTC Vive Cosmos	Oculus Rift CV1
Field of View [degrees]	**110**	110	130	110	110 diagonal, 90
Resolution over both eyes [pixels]	**2160 x 1200**	2880 x 1600	2880 x 1600	2880 x 1700	2160 x 1200
Refresh rate [Hz]	**90**	90	120-144	90	90
Rotation	**Yes**	Yes	Yes	Yes	Yes
Positional movement	**Yes**	Yes	Yes	Yes	Yes
Controller	**Yes**	Yes	Yes	Yes	Yes
Release year	**2016**	2018	2019	2019	2016
Platform	**SteamVR/PC**	SteamVR/PC	SteamVR/PC	SteamVR/PC	SteamVR/Oculus Store/PC
Price [USD]	**499**	799	999	699	399

Device	Oculus Go	Oculus Quest 2	Oculus Quest	Oculus Rift S	Google Cardboard	Samsung Odyssey
Field of View [degrees]	101	N/A	~100	115	Depends on smartphon	110
Resolution over both eyes	2560 x 1440	2664 x 1920	2880 x 1600	2560 x 1440	Depends on smartphon	2880 x 1600
Refresh rate [Hz]	60-72	72	72	80	Depends on smartphon	90
Rotation	Yes	Yes	Yes	Yes	Yes	Yes
Positional movement	No	Yes	Yes	Yes	No	Yes
Controller	Yes	Yes	Yes	Yes	No	Yes
Release year	2018	2020	2019	2019		2017
Platform	Oculus Platform SDK	Oculus Platform SDK	Oculus Platform SDK	SteamVR/ Oculus Store/PC	Google VR	Windows MR/PC/SteamVR
Price [USD]	199	299-399			Depends on smartphon	399

Resume/Curriculum Vitae

Andreas Bueckle

*Note: See **www.andreas-bueckle.com** for a more comprehensive overview of papers and talks.*

---

abueckle@iu.edu | www.andreas-bueckle.com

Research Gate | LinkedIn | Google Scholar | ORCID

Education

Indiana University, Bloomington, IN (USA), 08/2015-07/2021

Ph.D. in Information Science (minor: Informatics)

Title: "Optimizing Performance and Satisfaction in Virtual Reality Environments with the Data Visualization Literacy Framework"

Indiana University, Bloomington, IN (USA), 2014-2015

Non-degree graduate student in Information Science on a scholarship from the German Academic Exchange Service

Berlin University of the Arts, Berlin (Germany), 2012-2014

M.A. in Communications

Master thesis: "A theoretical framework for game design"

California State University, Chico, CA (USA), 2010-2011

Undergraduate exchange student

Eberhard Karls Universitaet, Tuebingen (Germany), 2008-2011

B.A. in Media Studies, minor: French

Bachelor thesis: "'Point-of-View' – a film exploring unreliable narration and convoluted montage"

Research Experience

Papers

K. Börner, S. A. Teichmann, E. M. Quardokus, J. Gee, K. Browne, D. Osumi-Sutherland, B. W. Herr II, A. Bueckle, et al. (2021). "Anatomical Structures, Cell Types, and Biomarkers Tables Plus 3D Reference Organs in Support of a Human Reference Atlas," bioRxiv preprint, 2021. [Online]. Available: https://www.biorxiv.org/content/10.1101/2021.05.31.446440v1. doi: 10.1101/2021.05.31.446440

A. Bueckle, K. Buehling, P. C. Shih, and K. Börner, "Comparing Completion Time, Accuracy, and Satisfaction in Virtual Reality vs. Desktop Implementation of the Common Coordinate Framework Registration User Interface (CCF RUI)," arXiv preprint arXiv: 2102.12030, 2021. [Online]. Available: https://arxiv.org/abs/2102.12030.

K. Börner, E. M. Quardokus, B. W. Herr II, L. E. Cross, E. G. Record, Y. Ju, A. Bueckle, J. P. Sluka, J. C. Silverstein, K. M. Browne, S. Jain, C. H. Wasserfall, M. L. Jorgensen, J. M. Spraggins, N. H. Patterson, M. A. Musen, and G. M. Weber, "Construction and Usage of a Human Body Common Coordinate Framework Comprising Clinical, Semantic, and Spatial Ontologies," arXiv preprint arXiv:2007.14474, 2020. [Online]. Available: https://arxiv.org/abs/2007.14474.

K. Börner, A. Bueckle, and M. Ginda, "Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments," Proceedings of the National

Academy of Sciences, vol. 116, no. 6, pp. 1857-1864, 2019, doi: 10.1073/pnas.1807180116.

K. Börner, A. H. Simpson, A. Bueckle, and R. L. Goldstone, “Science map metaphors: a comparison of network versus hexmap-based visualizations,” Scientometrics, vol. 114, no. 2, pp. 409-426, 2018, doi: 10.1007/s11192-017-2596-3.

K. Börner, A. Bueckle, P. Beesley, and M. Spremulli, “Lifting the Veil: Visualizing Sentient Architecture,” presented at the Future Technologies Conference, Vancouver, BC, Canada, 2017. [Online]. Available: https://cns.iu.edu/docs/publications/2017-Lifting-the-Veil.pdf.

Talks (selection)

A. Bueckle, "Optimizing Performance and Satisfaction in Matching and Movement Tasks in Virtual Reality with Interventions Using the Data Visualization Literacy Framework", Doctoral Dissertation Defense, 2021. [video]

A. Bueckle, "The CCF Tissue Registration User Interface (RUI) and Exploration User Interface (EUI)", Common Fund Data Ecosystem (CFDE) Workshop on Introduction to 3D Tissue Mapping, 2021.

A. Bueckle, "The CCF Tissue Registration User Interface (RUI)", Spatial Biology Europe (Online), 2021.

A. Bueckle, "Science Map Metaphors: A Comparison of Network Versus Hexmap-Based Visualization", Dagstuhl Seminar on Multi-Level Graph Representation for Big Data Arising in Science Mapping, 2021. [video] [paper]

A. Bueckle, Kilian Buehling, "Using VR Data Visualizations to Improve Time, Accuracy, Satisfaction in 2D Desktop vs. VR RUI", CNS Research Showcase, 2021. [video]

A. Bueckle, Katy Börner, "HuBMAP Reference Atlas: ASCT+B Tables & 3D Reference Library Registration User Interface & Visible Human Massive Open Online Course", HuBMAP SciTech Webinar, 2021.

A. Bueckle, "'I feel like Superman (in Luddy Hall)': Optimizing Performance in Virtual Reality Navigation Tasks with Interventions Using the Data Visualization Literacy Framework", Cyberinfrastructure for Network Science Center Create & Innovate Talk Series, 2021. [video]

Y. Jain, Andreas Bueckle, Ellen M. Quardokus, Yingnan Ju, Seth Winfree, Hrishikesh Paul, Bruce W. Herr II, Nick Tustison, James Gee, and Katy Börner, "R2UI: Adding ANTs to the Registration User Interface (RUI) to Support Spatial Registration of Tissue Blocks at Single-Cell Level" (Poster), HuBMAP December Demo Day, 2020.

A. Bueckle, Kilian Buehling, "Comparing Time, Accuracy, Satisfaction, and Accomplishment in Virtual Reality vs. Desktop Implementation of the Common Coordinate Framework Registration User Interface (CCF RUI)", Cyberinfrastructure for Network Science Center Research Showcase, 2020. [video]

A. Bueckle, "Investigating Your Traces: Optimizing Completion Times and Accuracy for Manipulation Tasks in Virtual Reality using Data Visualization", The Campus Alliance for Advanced Visualization Conference (CAAVCon), 2020.

A. Bueckle, Ellen M. Quardokus, Yingnan Ju, Sumeet Sarode, Yashvardhan Jain, Seth Winfree, Hrishikesh Paul, Bruce W. Herr II, Nick Tustison, James Gee, and Katy

Börner, “R2UI: Stage 2 Registration User Interface for HuBMAP for Accurate Spatial Registration of Tissue Blocks and Associated Cell Types and Biomarkers Within 3D Organ Models” (Poster), Human Biomolecular Atlas HIVE Virtual Demo Day, 2020.

A. Bueckle, “The Making of the Visible Human Massive Open Online Course: Empowering HuBMAP members with the Force of Video”, Human Biomolecular Atlas HIVE Virtual Demo Day, 2020.

A. Bueckle and Kilian Buehling, “The Common Coordinate Framework Registration User Interface (CCF RUI): A Comparison of Three Modalities”, Luddy School of Informatics, Computing, and Engineering, Indiana University Bloomington, IN, USA, 2019.

A. Bueckle, Leonard Cross, and Katy Börner, “HuBMAP CCF Registration User Interface (RUI)”, HuBMAP Inaugural Sci/Tech Webinar Series, Indiana University Bloomington, IN, USA, 2019.

A. Bueckle, “Data Visualization Literacy Gone Virtual”, The Campus Alliance for Advanced Visualization conference (CAAV), Indiana University Bloomington, IN, USA, 2019. Video recording

A. Bueckle, “Make-a-Vis & Tavola”, VISUALISE conference, Exploratorium, San Francisco, CA, USA, 2019. [video] [proceedings]

A. Bueckle and K. Börner, “Envisioning the Internet of Things”, Living Architecture Systems Group Symposium, Toronto, ON, Canada, 2019. [video] [proceedings](Keynote)

A. Bueckle, “Mapping Science, Technology, and Expertise on a Global Scale”, Pacific Operational Science & Technology Conference, Honolulu, HI, USA, 2018.

A. Bueckle, “IoT Visualization & Amatria”, Scientific Visualization Workshop Series, Indiana University Bloomington, IN, USA, 2018.

A. Bueckle, “Lifting the Veil: Visualizing Sentient Architecture”, Future Technologies Conference, Vancouver, BC, Canada, 2017. [proceedings] (Best Presentation Award)

A. Bueckle and Y. Ju, “Breathing Life into Things: Sentient Architecture & Augmented Reality”, Fashion Tech Week NY, New York City, NY, USA, 2017.

A. Bueckle, “Sentient Architecture: Visualizing Signal Flow in Intelligent Systems”, Intelligent & Interactive Systems Talk Series, Luddy School of Informatics, Computing, and Engineering, Indiana University Bloomington, IN, USA, 2017.

A. Bueckle, “Sentient Architecture: Visualizing Signal Flow in Intelligent Systems”, 1st Annual Graduate Conference at the Media School, Indiana University Bloomington, IN, USA, 2017.

A. Bueckle, “Science Forecast”, Modeling Science, Technology & Innovation Conference, Washington D.C., USA, 2016.

A. Bueckle, “Enhancing Sentient Architecture with Augmented Reality”, Living Architecture Systems Group Symposium, Toronto, ON, Canada, 2016. [white papers]

A. Bueckle, “Communicating and Visualizing the Internet of Things: Designing Augmented-Reality Data Visualizations to Communicate the Inner Workings of Living

Architecture Sculptures to Enhance Computational Literacy”, ILS Doctoral Research Forum, Indiana University Bloomington, IN, USA, 2016.

Guest Lectures

“Visual Insights Studio Tour”, Bridge to Informatics Program for DePauw University Undergraduate Students, Instructor: Dr. Patrick Shih, Indiana University, Bloomington, IN, 1/2018

“Visual Insights Studio Tour”, Welcoming a Delegation from Naval Surface Warfare Center Crane Division, Indiana University, Bloomington, IN, 1/2018

“Dendrites: Building Sentient Architecture”, Introduction to Intelligent Systems Engineering, Instructor: Dr. Katie Siek, Indiana University, Bloomington, IN, 12/2017

“Re-Creating Virtual Worlds with Unity 3D”, Introduction to Art History, Instructor: Martin Horn, Indiana University, Bloomington, IN, 10/2017

“Introduction to Information Visualization”, Introduction to Information Science, Instructor: Dr. Howard Rosenbaum, Indiana University, Bloomington, IN, 5/2016

Demos (selection)

“Dendrite Field Array Visualization”, TEDx Bloomington, Bloomington, IN, 8/2018

“Lifting the Veil: Visualizing Sentient Architecture”, Complex Networks and Systems Demo Series, Indiana University, Bloomington, IN, 4/2018

“Lifting the Veil: Visualizing Sentient Architecture”, Makevention, Bloomington, IN, 8/2018

Teaching Experience

Associate Instructor, Department of Computer Science, Indiana University

"Augmented Reality Visualizations of IoT Data", Hello Research Workshop for Undergraduate Women in Computer Science, *Indiana University, Bloomington, IN, 10/2018*

Associate Instructor & Video Content Creator, Department of Intelligent Systems Engineering, Indiana University

Visual Analytics Certificate (https://visanalytics.cns.iu.edu/), IU Expand, 2019-present

Z-637 Information Visualization, 2017-present

Information Visualization Massive Open Online Course (IVMOOC, https://ivmooc.cns.iu.edu/), 2017-present

Associate Instructor, Department of Information & Library Science, Indiana University

Z-637 Information Visualization, Spring 2017

Introduction to Search, Instructor: Dr. Xiaozhong Liu, Fall 2017

Introduction to Game Programming with Construct 2 and Panda,js, Instructor: Chabane Maidi, Spring 2016-Fall 2016

Evaluating Information and Intelligence, Instructor: Dr. Carol Choksy, Fall 2015

Service

Reviewer for CHI, 2021

Reviewer for Scientometrics, Spring 2017

Student member of Undergraduate Curriculum Sub-committee of the Department of Library and Information Science, Fall 2016

Awards

"Best Presentation Award", Future Technologies Conference, Vancouver, BC, 11/2017. Title: "Lifting the Veil: Visualizing Sentient Architecture"

"Department of Information & Library Science Ph.D. Travel Award", Indiana University, Fall 2017

"Information Science Leadership Award", financial support including full tuition coverage, Indiana University 2015-2020

"German Academic Exchange Service Scholarship for Graduate Students", 2014-2015

"Baden-Wuerttemberg Foundation Scholarship", Stipend for exchange semester at California State University Chico, CA (USA), 2010-2011

Skills

Unity 3D, SteamVR, HTC Vive, Microsoft HoloLens, Oculus Quest 1 + 2, Blender (basic), Maxon Cinema 4D (basic), SketchUp (basic)

Programming languages: C#, Python, Processing, JavaScript

Visualization: ggplot2, Tableau, R, Gephi

Web development: Angular 12

Office products: Microsoft Word, Excel, PowerPoint

Adobe Creative Cloud: Premiere Pro (video editing), After Effects (motion graphics/animations), Lightroom (photo editing), XD (UI prototyping), Illustrator

Photography: documenting social issues, interpersonal communication/events, and nature/landscape (portfolio)

Videography: image films, interviews, pitch videos, news (portfolio)

Languages: German (native), English (proficient), French (intermediate), Latin, Ancient Greek, Ancient Hebrew

Extracurricular

Coach, Sailing Club at Indiana University, Indiana University, Bloomington, IN, 2020-presend

Captain, Sailing Club at Indiana University, Indiana University, Bloomington, IN, 2016-2020

Marketing Chair, Sailing Club at Indiana University, Indiana University, Bloomington, IN, 2014-2016

US Sailing Certification, since 2018

American Red Cross CPR/AED/SFA certified, since 2015

German Coastal Sailing License (550 nautical miles on coastal and offshore waters as co-skipper to date), since 2014

German Inland Waters Sailing License, since 2014

Short Range Radio Communication License, since 2014