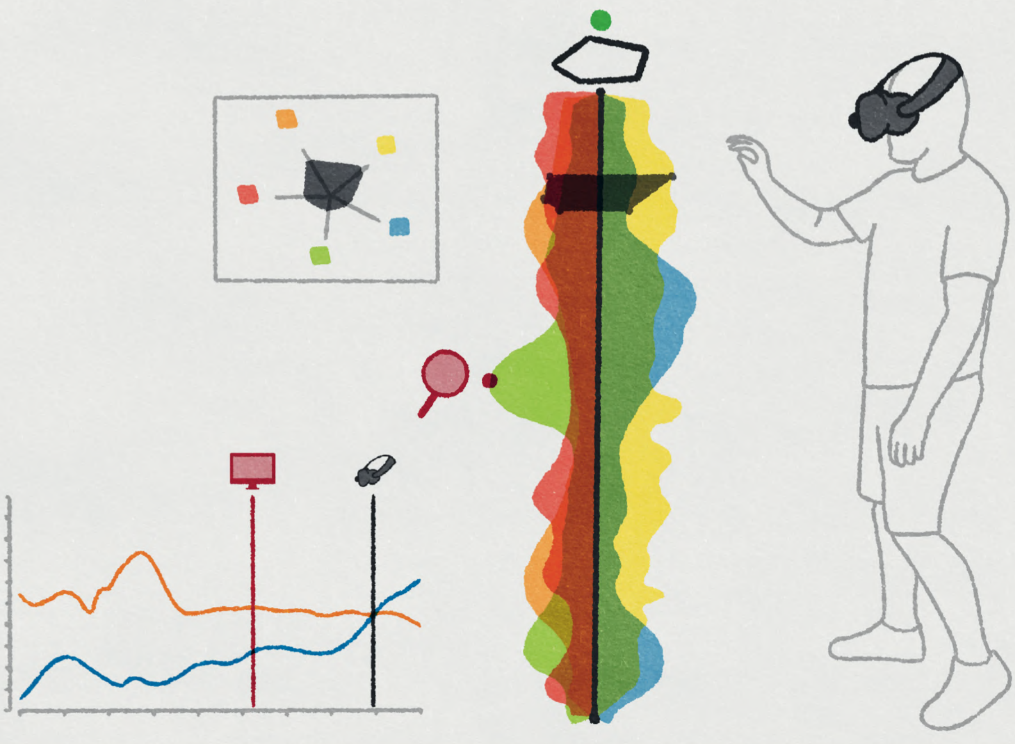
Linnaeus University Dissertations
No 459/2022

Nico Reski

Supporting Data Interaction and Hybrid Asymmetric Collaboration Using Virtual Reality Within the Context of Immersive Analytics



LINNAEUS UNIVERSITY PRESS

Supporting Data Interaction and Hybrid Asymmetric Collaboration Using Virtual Reality Within the Context of Immersive Analytics

Linnaeus University
Dissertations

No 459/2022

Supporting Data Interaction and Hybrid Asymmetric Collaboration Using Virtual Reality Within the Context of Immersive Analytics

Nico Reski

LINNAEUS UNIVERSITY PRESS

Supporting Data Interaction and Hybrid Asymmetric Collaboration Using Virtual Reality Within the Context of Immersive Analytics
Doctoral Dissertation, Department of Computer Science and Media Technology, Linnaeus University, Växjö, 2022

ISBN: 978-91-89709-26-3 (print), 978-91-89709-27-0 (pdf)
Published by: Linnaeus University Press, 351 95 Växjö
Printed by: Exakta Print AB, 2022

Abstract

Reski, Nico (2022). *Supporting Data Interaction and Hybrid Asymmetric Collaboration Using Virtual Reality Within the Context of Immersive Analytics*, Linnaeus University Dissertations No 459/2022,
ISBN: 978-91-89709-26-3 (print), 978-91-89709-27-0 (pdf).

Immersive display and interaction technologies have rapidly evolved in recent years, offering advanced techniques compared to traditional Human-Computer Interaction. Computer-generated Virtual Environments viewed with stereoscopic depth perception and explored using 3D spatial interaction can represent more accurately how humans naturally interact in the real world. Data analysis is a promising area of application for such technologies, holding potential to promote intuitive interaction, user engagement, collaboration, and data curiosity, as well as to foster appropriate contextual visualization. Even when techniques such as Machine Learning and Data Mining assist with the analysis of data, human interpretation, contextualization, and meaning making are still needed. The design of immersive data visualization and interaction is challenging due to the complexity of the involved technologies and human factors, which calls for an interdisciplinary research effort.

The focus of this thesis is to investigate means of exploration, interaction, and collaboration using Virtual Reality and 3D gestural input in immersive environments within the context of spatio-temporal data analysis. Based on existing literature as well as following an applied and interdisciplinary research approach, a design space for this type of Immersive Analytics is defined. The emphasis on spatio-temporal data is relevant across various real-world contexts and scenarios, such as sociolinguistics and climate analysis, given that data collected nowadays commonly feature descriptors of where and when they were captured. An immersive data analysis system has been implemented and evaluated across three virtual environment iterations. Two core themes from a user-centered perspective are interaction and collaboration. The design of useful and engaging 3D gestural interaction techniques support the conduction of typical analytical tasks that aid the data exploration and thus the discovery of insights. Furthermore, data analysis is seldom a solitary activity, but can be conducted in collaboration with multiple analysts, who combine their knowledge to interpret and discuss the discoveries. For this purpose, the concept of Hybrid Asymmetric Collaboration is defined, aiming to facilitate an envisioned broader analytical workflow that assumes a mixture of immersive and non-immersive interfaces (hybrid) as well as distinct user roles (asymmetric). To bridge data analysis across heterogeneous interface types, the design of visual information cues is investigated to support foundational aspects of collaboration, such as awareness, common ground, reference, and deixis.

The conducted research has been empirically evaluated using a combination of standardized and custom methods in a total of six main studies. The outcomes of these studies allow for reflections and the proposal of design guidelines for collaborative data interaction in immersive spaces.

Keywords: 3D gestural input, 3D radar charts, 3D user interfaces, empirical evaluation, head-mounted display, hybrid asymmetric collaboration, immersive analytics, spatio-temporal data interaction, user interface design, virtual reality

To family and friends.

Gratitude and appreciation for all the love, support, inspiration, and simply being there.

Abstract

Immersive display and interaction technologies have rapidly evolved in recent years, offering advanced techniques compared to traditional Human-Computer Interaction. Computer-generated Virtual Environments viewed with stereoscopic depth perception and explored using 3D spatial interaction can represent more accurately how humans naturally interact in the real world. Data analysis is a promising area of application for such technologies, holding potential to promote intuitive interaction, user engagement, collaboration, and data curiosity, as well as to foster appropriate contextual visualization. Even when techniques such as Machine Learning and Data Mining assist with the analysis of data, human interpretation, contextualization, and meaning making are still needed. The design of immersive data visualization and interaction is challenging due to the complexity of the involved technologies and human factors, which calls for an interdisciplinary research effort.

The focus of this thesis is to investigate means of exploration, interaction, and collaboration using Virtual Reality and 3D gestural input in immersive environments within the context of spatio-temporal data analysis. Based on existing literature as well as following an applied and interdisciplinary research approach, a design space for this type of *Immersive Analytics* is defined. The emphasis on spatio-temporal data is relevant across various real-world contexts and scenarios, such as sociolinguistics and climate analysis, given that data collected nowadays commonly feature descriptors of where and when they were captured. An immersive data analysis system has been implemented and evaluated across three virtual environment iterations. Two core themes from a user-centered perspective are interaction and collaboration. The design of useful and engaging 3D gestural interaction techniques support the conduction of typical analytical tasks that aid the data exploration and thus the discovery of insights. Furthermore, data analysis is seldom a solitary activity, but can be conducted in collaboration with multiple analysts, who combine their knowledge to interpret and discuss the discoveries. For this purpose, the concept of *Hybrid Asymmetric Collaboration* is defined, aiming to facilitate an envisioned broader analytical workflow that assumes a mixture of immersive and non-immersive interfaces (hybrid) as well as distinct user roles (asymmetric). To bridge data analysis across heterogeneous interface types, the design of visual information cues is investigated to support foundational aspects of collaboration, such as awareness, common ground, reference, and deixis.

The conducted research has been empirically evaluated using a combination of standardized and custom methods in a total of six main studies. The outcomes of these studies allow for reflections and the proposal of design guidelines for collaborative data interaction in immersive spaces.

Keywords: 3D gestural input, 3D radar charts, 3D user interfaces, empirical evaluation, head-mounted display, hybrid asymmetric collaboration, immersive analytics, spatio-temporal data interaction, user interface design, virtual reality

Svensk sammanfattning

De senaste åren har så kallade immersiva skärm- och interaktionsteknologier utvecklas i snabb takt. Sådana teknologier erbjuder mer avancerade tekniska lösningar jämfört med mer traditionell människa-datorinteraktion. Immersiva inslag som datorgenererade virtuella miljöer med stereoskopisk djupuppfattning och 3D-interaktion kan på ett mer exakt och naturligt sätt representera de interaktioner vi utför i den verkliga världen.

Dataanalys är ett lovande användningsområde för immersiva teknologier, med potential att underlätta intuitiv interaktion, skapa och uppmuntra engagemang, samarbete och nyfikenhet för data, samt främja relevant kontextuell visualisering. Även när teknologier som maskininlärning och datautvinning används är det fortfarande nödvändigt att komplettera resultatet med en överordnad mänsklig nivå av tolkning, kontextualisering och meningsskapande för att göra analyserna kompletta och användbara. Designen av immersiv visualisering och interaktion är en krävande utmaning på grund av de teknologiska och mänskliga aspekternas komplexitet och fordrar därför en tvärvetenskaplig forskningsmetod.

Avhandlingens syfte är att utforska metoder som kan stödja undersökning, interaktion och samarbete med hjälp av *Virtual Reality* och 3D-handinteraktioner i immersiva miljöer inom ramen för analys av datamängder som innehåller rumsliga och tidsmässiga data. Denna typ av *Immersive Analytics* baseras på teoretiska utgångspunkter från tidigare forskning, samt en tillämpad, tvärvetenskaplig forskningsmetod. Fokuset på datamängder som innehåller rumsliga och tidsmässiga data är relevant för flera olika ämnesområden och sammanhang, som till exempel sociolingvistik och klimatanalys, eftersom data som samlas in i nu för tiden ofta innehåller beskrivningar av var och när de mättes eller observerades. Ett immersivt datanalys-system har för detta arbete implementerats och utvärderats i tre iterationer. Två grundläggande områden för ett användarcentrerat perspektiv är interaktion och samarbete. Designen av användbara och engagerande interaktionstekniker via 3D-handinteraktioner stödjer utförandet av typiska analytiska uppgifter och underlättar undersökningen av data och därmed upptäckten av nya insikter. Dataanalys är sällan en aktivitet som utförs ensam, utan sker snarare i ett samarbete där flera analytiker med sin kombinerade kunskap tolkar och diskuterar upptäckter tillsammans. För detta ändamål har begreppet *Hybrid Asymmetric Collaboration* definierats, vilket syftar till att beskriva ett betydligt bredare analytiskt arbetsflöde som förutsätter och omfattar en blandning av nya immersiva och nuvarande icke-immersiva gränssnitt (*hybrid*) såväl som distinkta användarroller (*asymmetric*). För att kombinera och integrera dataanalys mellan heterogena typer av användargränssnitt utforskas designen av visuella informationssignaler med syfte att stödja grundläggande samarbetsaspekter såsom medvetenhet, gemensam grund, referens och kontext.

Den genomförda forskningen har utvärderats empiriskt med en kombination av standardiserade och anpassade metoder i totalt sex större studier. Resultatet av dessa studier möjliggör reflektioner kring samt förslag på olika designriktlinjer för interaktion och samarbete i immersiva miljöer.

Nyckelord: 3D-handinteraktion, 3D-radardiagram, 3D-användargränssnitt, design av användargränssnitt, empirisk utvärdering, huvudburen display, hybrid-asymmetrisk samarbete, immersiv analys, virtuell verklighet

Acknowledgments

Setting out to obtain a doctoral degree is a major endeavor that takes many years and efforts. Naturally, as with everything in life, it is a time of constant progression and exciting changes that encourage growth while maintaining a healthy portion of youthfulness. At least that is how these past years appear to me in the grand scheme of things.

This thesis would not have taken on the shape it has without the support – direct and indirect – of some incredible people that I am thankful for. And now it is time to acknowledge them.

The without a doubt most vital role throughout the doctoral journey played my supervisor Aris Alissandrakis whose support, guidance, and dedication is beyond words. He always made time for long and detailed research conversations, was enthusiastic about all the various subject matters at hand, and created an environment in which one can thrive to the fullest. His supervision and values are defining, and I sincerely hope that I can pass them down one day myself. Working together with Aris was a fantastic time that I will always gladly look back at in the years to come.

I am also very grateful for the support by Andreas Kerren who provided valuable impulses, feedback, and reflections throughout all his years on the supervision team.

Special thanks go to Jukka Tyrkkö who was an essential part of the interdisciplinary research efforts throughout my thesis and with whom it always was a pleasure to collaborate.

I would also like to express my thanks to Kostiantyn Kucher who was not just inspiring in the way he approached his doctoral studies, but also provided substantial feedback on the initial thesis draft.

Over the years I had the pleasure to meet, work with, and study along some great individuals at Linnæus University, in particular David Johansson, Myrtha Voges, Maximilian Müller, Angelos Chatzimparmpas, Anna Wingkvist, Daniel Ihrmark, Alisa Lincke, Oskar Petterson, Björn Zimmer, Mattias Davidsson, Koraljka Golub, Mikko Laitinen, Magnus Levin, and Jenny Ström Herold. I am so glad to know you all.

I want to universally thank all the participants who partook in the empirical work throughout the years and interacted with the various developed Virtual Environment systems. They did not just take time off for their participation, but also contributed with constructive feedback and insights without which this thesis would not have been possible.

Furthermore, I would like to give a shout-out to the colleagues of the Center for Ubiquitous Computing at the University of Oulu, Finland. Their hosting of the International UBI Summer School, in which I had the pleasure to participate multiple times, provided a friendly and constructive environment to connect and collaborate with other enthusiastic researchers, students, and professionals in a week of full of "*hard work and fun*". These weeks were certainly a highlight of my academic development so far, learning about various topics, obtaining valuable impressions, and making great memories.

Finally, I have the pleasure to acknowledge those I hold closest and dearest, without whom life would be inconceivable, and I am so glad that they followed along on this journey.

A tremendous amount of love and appreciation goes to Susanna, Ellinor, Per-Olof, and Rasmus who not just partook and encouraged this journey, but who have become an integral part of my life in Sweden that I would not like to miss for anything in the world.

My thanks go to Onkel Mario and Tante Sylvia along with Melina, Sophie, and Janice with whom one can always kid around, laugh, and have a good time.

I have so much appreciation for Tante Christa and Onkel Ado who were supportive, interested, and caring for as long as I can remember – not just during the doctoral studies but throughout life in general.

My love and gratitude for Oma Erika, Opa Jürgen, and "Mamma" Britta is beyond words. For the past 34 years they have created an environment in which one can but thrive – all day every day. Through their love and support they made it possible to set out into the world, inspired, open-minded, full of curiosity, eager to learn, and with sincere values. For all that, one could not possibly feel more loved and appreciated, which is why it is impossible to put my love and appreciation for them into words!

Nico Reski
Växjö, Sweden · 2022

Table of Contents

List of Abbreviations

CIA	Collaborative Immersive Analytics
CSCW	Computer-Supported Cooperative Work
CSV	comma-separated values
CVE	Collaborative Virtual Environment
FKS	Flow Short Scale (*German original*: Flow-Kurzskala)
HCI	Human-Computer Interaction
HMD	head-mounted display
IA	Immersive Analytics
InfoVis	Information Visualization
JSON	JavaScript Object Notation
NTS	Nordic Tweet Stream (*dataset*)
PWt	Plant-Weather timelines (*dataset*)
SSQ	Simulator Sickness Questionnaire
STCQ	Spatio-Temporal Collaboration Questionnaire
SUS	System Usability Scale
TLX	Task Load Index (*sometimes also as* NASA TLX)
UES-SF	User Engagement Scale - Short Form
UI	user interface
VA	Visual Analytics
VE	Virtual Environment
VR	Virtual Reality
VWG	Virtual World Generator
2D	two-dimensional, two dimensions
3D	three-dimensional, three dimensions
3D UI	user interface with support for 3D interaction

List of Figures

List of Tables

List of Publications

This dissertation is based on the following refereed publications in chronological order (I have contributed to all stages of work as the lead author):

1. Nico Reski and Aris Alissandrakis. Open data exploration in virtual reality: a comparative study of input technology. *Virtual Reality*, 24:1–22, March 2020. `doi:10.1007/s10055-019-00378-w`. Materials appear in Chapters 4 and 5.

2. Nico Reski, Aris Alissandrakis, Jukka Tyrkkö, and Andreas Kerren. "Oh, that's where you are!" – Towards a Hybrid Asymmetric Collaborative Immersive Analytics System. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society* (NordiCHI 2020), pages 5:1–12, Tallinn, Estonia, 25–29 October 2020. Association for Computing Machinery (ACM). `doi:10.1145/3419249.3420102`. Materials appear in Chapters 3, 4, 5, and 6.

3. Nico Reski, Aris Alissandrakis, and Andreas Kerren. Exploration of Time-Oriented Data in Immersive Virtual Reality Using a 3D Radar Chart Approach. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society* (NordiCHI 2020), pages 33:1–11, Tallinn, Estonia, 25–29 October 2020. Association for Computing Machinery (ACM). `doi:10.1145/3419249.3420171`. Materials appear in Chapters 3, 4 and 5.

4. Nico Reski, Aris Alissandrakis, and Andreas Kerren. An Empirical Evaluation of Asymmetric Synchronous Collaboration Combining Immersive and Non-Immersive Interfaces Within the Context of Immersive Analytics. *Frontiers in Virtual Reality*, 2:743445:1–29, 17 January 2022. `doi:10.3389/frvir.2021.743445`. Materials appear in Chapters 3, 4, 5, and 6.

5. Nico Reski, Aris Alissandrakis, and Andreas Kerren. Designing a 3D Gestural Interface to Support User Interaction with Time-Oriented Data as Immersive 3D Radar Charts. *In preparation*. Materials appear in Chapters 3, 4 and 5.

6. Nico Reski, Aris Alissandrakis, and Andreas Kerren. User Preferences of Spatio-Temporal Referencing Approaches For Immersive 3D Radar Charts. *In preparation*. Materials appear in Chapters 3, 5 and 6.

This dissertation is also based on the following software presentations with published abstract (I have contributed to all stages of work, including hands-on demonstration for on-site event attendees):

1. Aris Alissandrakis, Nico Reski, Mikko Laitinen, Jukka Tyrkkö, Magnus Levin, Jonas Lundberg. Visualizing dynamic text corpora using Virtual Reality. In *The 39th Annual Conference of the International Computer Archive for Modern and Medieval English: Corpus Linguistics and Changing Society* (ICAME 39), Book of Abstracts, page 205, Tampere, Finland, 30 May – 3 June 2018. International Computer Archive of Modern and Medieval English (ICAME). `urn:nbn:se:lnu:diva-75064`. Materials appear in Chapter 5.

2. Nico Reski, Aris Alissandrakis, and Jukka Tyrkkö. Collaborative exploration of rich corpus data using immersive virtual reality and non-immersive technologies. In *The 2nd International Conference: Approaches to Digital Discourse Analysis* (ADDA 2), Book of Abstracts, pages 7–9, Turku, Finland, 23–25 May 2019. University of Turku. `urn:nbn:se:lnu:diva-83858`. Materials appear in Chapters 6.

This dissertation partially uses the materials of the following refereed publication (I have contributed to some stages of work in conceptualization, implementation, and writing):

- Aris Alissandrakis, Nico Reski, Mikko Laitinen, Jukka Tyrkkö, Jonas Lundberg, and Magnus Levin. Visualizing rich corpus data using virtual reality. *Studies in Variation, Contacts and Change in English: Corpus Approaches into World Englishes and Language Contrasts*, 20:online, December 2019. `urn:nbn:se:lnu:diva-90516`. Materials appear in Chapter 5.

Further publication not related to this dissertation (I have contributed to all stages of work as the lead author):

- Nico Reski and Aris Alissandrakis. Using an Augmented Reality Cube-Like Interface and 3D Gesture-Based Interaction to Navigate and Manipulate Data. In *Proceedings of the 11th International Symposium on Visual Information Communication and Interaction* (VINCI '18), pages 92–96, Växjö, Sweden, 13–15 August 2018. Association for Computing Machinery (ACM). `doi:10.1145/3231622.3231625`.

Chapter 1

Introduction

Contents

While immersive display and interaction technologies have fascinated researchers for many years, there has been a renaissance of interest in recent times (Lanman et al., 2014; Sutherland, 1968). In comparison to arguably more traditional non-immersive interfaces, such as a two-dimensional (2D) monitor as well as keyboard and mouse, immersive technologies allow for a closer coupling between human user and computer system by utilizing a higher level of sensory fidelity (Bowman and McMahan, 2007). Large wall-mounted displays and projection technologies can be used to create physical spaces that allow one or even multiple users to surround – or "immerse" – themselves with visual contents for various purposes. Virtual Reality (VR) interfaces, such as stereoscopic head-mounted displays (HMDs), enable a user to visually immerse themself in a three-dimensional (3D) interactive Virtual Environment (VE), perceiving the computer-generated contents intuitively as one would in the real world by commonly applying skills learned throughout life. Modern sensor technologies allow for device and body tracking in 3D, in turn enabling spatial interaction in the virtual space, for instance through the direct manipulation of 3D objects. Portable handheld devices can be used to display visual contents in-situ, even by juxtaposing, aligning, and overlaying virtual with real-world objects in 3D through the application of Mixed Reality approaches. Audio processing and rendering have capabilities to create 3D spatial audio feedback, typically in conjunction with respective head tracking sensors, providing either standalone or multimodal experiences, for instance in synergy with respective visual interfaces. These are just a few promising examples outlining the current state of mainstream immersive technologies.

While immersive display and interaction technologies increase not just in availability and affordability but also become maintenance-friendlier, more and more researchers and practitioners gain access to them for their subsequent utilization. Naturally, it is up to the designer to lay out and create immersive experiences that build on such technologies, across different contexts and scenarios, and thus possibly taking on manifold variants. For instance, immersive VR experiences for games and entertainment-related purposes are among the driving forces for the success and establishment of these technologies in the consumer market (Wohlgenannt et al., 2020). At the same time, VR has been utilized for a variety of exciting non-entertainment purposes. Technology-mediated learning and education scenarios appear to be of particular interest for the utilization of VR, especially with respect to health and training tasks (Oyelere et al., 2020). So-called exergames encourage bodily engagement through interactive immersive experiences, facilitating mobility and the use of different physical components that in turn can foster aspects of health and rehabilitation (Costa et al., 2019; Kivelä et al., 2019). Simulated training environments, such as the one for lunar exploration as presented by Olbrich et al. (2018), provide safe and risk-free alternatives for the practice of specific tasks as more realistic experiences compared to non-immersive instruction methods. With 3D models already existing and being used in various contexts and industry branches, such as architecture, immersive VR approaches not only enable a more representative first-person viewing but even interactive modification in real-time to prototype design in-situ directly in the VE, either alone or collaboratively (Sugiura et al., 2018; Wolf et al., 2017). There is also potential to apply VR technologies in non-commercial and public settings, such as libraries, to provide engaging multisensory experiences that build upon and complement the respective real-world counterpart in meaningful ways, while educating the public audience about these types of interfaces at the same time (Holappa et al., 2018; Pouke et al., 2018). Immersive VR experiences have also been used to promote peer-to-peer learning by offering viable digital alternatives to in-person meetups, for instance within the context of urban agriculture, allowing not just verbal experience report and information exchange but also interactive demonstrations of relevant mechanisms (Parikh et al., 2022).

Another promising area of application for immersive display and interaction technologies is data analysis, which has the potential to be enhanced through intuitive, fluid, and integrated sensory interaction (Roberts et al., 2014). The emerging field of Immersive Analytics (IA) has been of increased interest to data visualization and interaction researchers in recent years (Fonnet and Prié, 2021). IA aims to support data understanding, analytical reasoning, and collaborative decision making through the utilization of immersive human-computer interfaces as engaging embodied tools in virtual spaces (Skarbez et al., 2019; Dwyer et al., 2018; Hackathorn and Margolis, 2016; Chandler et al., 2015). Skarbez et al. (2019) as well as Hackathorn and Margolis (2016) highlight that while techniques such as machine learning and data mining can be useful for the processing of large multivariate datasets in order to identify patterns, discover potential points of interest, and extract first insights, a need for human interpretation, contextualization, and overall meaning making remains, not least because of human semantics capabilities. As such, immersive interfaces have the potential to synergize, complement, and enhance data analysis conducted by human users in various ways (Dwyer et al., 2018). Immersion can facilitate user engagement, allowing to actively explore and reflect on data (Büschel et al., 2018; Millais et al., 2018). 3D visual displays allow for an intuitive spatial understanding through the perception of depth cues (Bowman and McMahan, 2007), which in turn can be used to facilitate the visualization of data with spatial embeddings (Marriott et al., 2018). The creation of large virtual interaction spaces has also the potential to decrease information clutter, reducing the amount of overlapping graphical artifacts that often require active rearrangement through the user (Bowman and McMahan, 2007). Immersive data analysis environments can also be facilitated through the support for collaboration, enabling multiple users to explore, analyze, and interpret data together in order to establish joint conclusions (Ens et al., 2019; Billinghurst et al., 2018). And naturally, new types of technologies also invite researchers and practitioners to explore novel and intuitive applications for data visualization and interaction in general (Kraus et al., 2022; Sommer et al., 2017; Febretti et al., 2013; Reda et al., 2013). Following these examples, the vast variety of potential approaches for the design of immersive data analysis tools becomes apparent, demanding for an interdisciplinary expertise, among others from research areas such as Information Visualization (InfoVis), Visual Analytics (VA), Computer-Supported Cooperative Work (CSCW), Collaborative Virtual Environments (CVEs), VR, 3D User Interfaces (3D UIs), and Human-Computer Interaction (HCI) in general. Consequently, there are many exciting research challenges to further investigate and empirically evaluate the use of immersive display and interaction technologies for data analysis purposes (Ens et al., 2021), including an overall critical reexamination of utilizing the interactive 3D space for data analysis, which traditionally has been less commonly applied outside of Scientific Visualization (Marriott et al., 2018).

1.1 Motivation for Immersive Data Interaction and Collaboration

Two reoccurring themes in the IA research agenda are concerned with interaction and collaboration (Ens et al., 2021; Skarbez et al., 2019; Dwyer et al., 2018; Marai et al., 2016). Besides being able to perceive and infer visual structures, Streeb et al. (2021) emphasize on the importance of interaction, providing means to access and explore data as well as to learn how to interpret visual patterns. Research areas such as InfoVis and VA have comparatively established theories and practices for data interaction, commonly in 2D, displayed through a normal monitor, and operated with keyboard and pointer input (Fikkert et al., 2007). Within the context of IA however, investigations are needed towards the support of embodied data exploration and spatial immersion in 3D (Dwyer et al., 2018), either applied and adapted from existing interaction techniques or built from the ground up. The design of user-friendly 3D UIs is generally regarded as challenging in itself, requiring a multitude of considerations, from technology properties to interaction technique characteristics to human factors and ergonomics (LaViola, Jr. et al., 2017, Chapter 10.1). The interaction with IA systems is also inherently challenging with respect to providing support for a variety of typical data analysis tasks (Yi et al., 2007; Shneiderman, 1996), thus requiring rich sets of features that in turn increase overall system complexity (Ens et al., 2021; Büschel et al., 2018). The need for more applied research to provide best practices and guidelines for the interaction with IA systems is also highlighted by Fonnet and Prié (2021), encouraging the incorporation of foundational knowledge from research communities such as VR and 3D UIs into IA application scenarios.

Furthermore, data analysis is nowadays seldom an activity that is conducted in isolation, but instead in collaboration with multiple analysts. In fact, enabling multiple users to collaboratively explore and interpret data is often desired. For instance, the analysis of large datasets commonly requires a broad expertise that is unlikely to be covered by just a single analyst (Zimmer and Kerren, 2017; Isenberg et al., 2011). Collaboration has been shown to be more effective compared to working alone (Billinghurst et al., 2018), arguably also because it is inherently anchored within the human nature (Neumayr et al., 2017). Visual analysis and meaning making involve next to perceptual and cognitive also social processes, such as the communication with other analysts to discuss the interpretation of the data, each providing their individual and contextual knowledge (Billinghurst et al., 2018; Heer and Agrawala, 2008). However, providing collaboration support within the context of IA, or more specifically Collaborative Immersive Analytics (CIA), can pose several design and interaction challenges, particularly as immersive display and interaction technologies are often by default rather tailored to be experienced by a single user (Skarbez et al., 2019; Cordeil et al., 2017b; Hackathorn and Margolis, 2016). For instance, following a VR approach and wearing a HMD, the user is visually isolated from the physical real-world surroundings and presented with computer-generated graphical contents. This introduces rather remote-like characteristics in regard to potential collaboration, even in co-located scenarios, as important visual information cues, such as facial expressions, body language, or spatial references, are no longer easily accessible right out of the gate. Instead, collaborative features and information cues need to be appropriately integrated through the CIA system designer, elevating the importance for nonverbal communication features (Cruz et al., 2015). While research areas such as CSCW and CVEs are comparatively well studied over many years and provide important foundational concepts and approaches, Fonnet and Prié (2021) as well as Billinghurst et al. (2018) highlight the lack of applied CIA research today.

In general, there is one aspect concerning IA that cannot be stressed enough: IA research *is not intended to substitute* existing 2D and non-immersive data analysis practices, but instead to *add* immersive 3D approaches that *synergize and complement* the overall workflow. Particularly with respect to the analysis of large multivariate datasets, there is no single tool to satisfy all of a user's needs, but rather multiple ones are desired, each for their own purpose, using different visualization and interaction techniques, to support diverse problem-solving strategies, and as such composing an overall greater analysis workflow (Cavallo et al., 2019; Wang et al., 2019; Isenberg, 2014). Consequently, there is huge potential for the integration of InfoVis, VA, and IA tools, not just with respect to extended workflows but also active cross-platform collaboration, enabling multiple users to assume different roles, perform different tasks, and have different perspectives on the same data (Fröhler et al., 2022; Ens et al., 2021; Fonnet and Prié, 2021; Billinghurst et al., 2018).

1.2 Research Problem, Scope, Goal, and Objectives

Following the first pages of this thesis, it becomes apparent that the main interest of the presented research is concerned with the application of immersive display and interaction technologies within the overall context of data analysis. The **research problem** is that there is an overall lack of empirical research in the area of IA, arguably also due of its highly interdisciplinary characteristics, investigating the transfer and application of existing foundational knowledge and principles into new data analysis tools and approaches. In turn, the insights gained from the design, development, and evaluation of these new tools may be utilized to generate best practices, guidelines, and general strategies for the implementation of future data analysis systems and solutions.

Naturally, IA as well as CIA can concern a vast variety of different subjects, among others depending on (1) the technological approach, both conceptually as well as practically based on the applied output and input devices, (2) the data type, context, scenario, and use case, (3) the purpose of the data analysis tool and its supported tasks, (4) the role and knowledge of its user, and (5) the setup and purpose of collaboration. All these aspects may imply their own individual requirements, hence demanding careful considerations and a clear depiction of the **research scope**. With that in mind, this thesis is concerned with the application of a VR approach, implemented through the utilization of HMD devices for visual immersion as well as 3D gestural input, commonly referred to as hand or mid-air interaction, for 3D spatial interaction in the VE.

Furthermore, this thesis is focused on the analysis of spatio-temporal data as a specific type of multivariate data. Spatio-temporal data is a common type of data that is highly relevant for the measurement and observation of various real-world phenomena, typically featuring descriptors in regard to where (spatial) and when (temporal) a measurement or observation was made. The presented research adopts an overall user-centered perspective insofar as to focus on the design, development, and evaluation of immersive data analysis tools that support intuitive, usable, and engaging *interaction* and *collaboration*. While overall visual perception related matters are considered naturally along the way as part of human information processing in general, for instance influencing aspects of a user's ability to interpret the visualized data in the VE and as such to solve tasks, the presented research is not focused on the deeper investigation and evaluation of fundamental human visual perception matters. Independent of whether the user is a novice or an expert in regard to a particular data context, they should be able to utilize the developed tools without the need for extensive training to learn the interface. Furthermore, rather than assuming that multiple users utilize the same type of interface, the presented research focuses on the investigation of cross-platform collaboration, specifically combining immersive and non-immersive interfaces for collaborative data analysis.

Following the Goal/Question/Metric paradigm as described by Basili et al. (1994), and aligned with the overall described research problem and scope, the **research goal** of this thesis is defined as follows:

Purpose	Design, develop, and evaluate VR-based data analysis tools that utilize HMD devices and 3D gestural input
Issue	to provide empirical insights and to derive design guidelines
Object	for the immersive interaction and collaboration around spatio-temporal data
Viewpoint	from a user-centered perspective, following an applied and inter-disciplinary research approach.

In order to practically approach and assist with the investigation of the defined research goal, three **research objectives** are defined.

Research Objective 1

Design and implementation of a system that allows for multivariate data analysis using immersive display and interaction technologies.

The first research objective is concerned with general aspects of developing an immersive data analysis system. This is achieved through careful contemplation and identification of various requirements that are relevant within the presented context, for instance as functional, non-functional, and user experience requirements with respect to data processing, visualization, interaction, and collaboration. These in turn aim to facilitate the conceptual as well as technological design of a general system architecture, providing important building blocks that serve as an overall foundation to aid the practical implementation of immersive data analysis tools.

> **Research Objective 2**
>
> Investigation of 3D UI design approaches to support immersive interaction with spatio-temporal data.

The second research objective is focused on the investigation of general 3D UI design aspects that are relevant for the interaction with spatio-temporal data in an immersive VE. Three primary aspects within this context are concerned with (1) the data entity visualization design, i.e., the visual representation of individual data items in the virtual space, (2) the VE composition, i.e., the overall structure and arrangement of all virtual artifacts in the immersive data analysis environment, and (3) the interaction design, i.e., the support for the conduction of typical analysis tasks under consideration of general requirements inherent from the applied display and interaction technologies. Aligned with these aspects, respective immersive data analysis tools can be designed, developed, and subsequently evaluated, for instance within the scope of user interaction studies. The utilization of standardized as well as custom methods allows for the collection of empirical data, which can be used to provide insights and reflections for the design of future similar tools.

> **Research Objective 3**
>
> Extension of the immersive data analysis system to support collaboration using heterogeneous interfaces and user roles.

Finally, the third research objective is concerned with the combination of immersive and non-immersive interfaces to enable multiple users to collaboratively analyze data, each using individual tools for dedicated purposes, to provide their own views on different aspects of the same data. To bridge the gap between IA and non-immersive InfoVis/VA tools that allow for synchronous collaboration, the respective tools need to be extended through the integration of features that support the users with their collaboration. Naturally, the immersive data analysis tools developed as part of investigating the second research objective serve as a fundamental building block for the exploration of this third and final research objective. The collaborative features design should aim to align with and support relevant foundational concepts, such as awareness, common ground, reference, and deixis. Furthermore, non-immersive data analysis tools are required for the empirical evaluation within the presented context, either based on existing ones that are extended through collaborative features, or developed from scratch as representative prototypes that utilize commonly applied visualization and interaction techniques. With all system components at hand, i.e., IA tool, InfoVis/VA tool, and collaborative features, their application can be evaluated by pairs of users in respective user interaction studies, allowing for the collection of empirical data and the subsequent deduction of insights.

1.3 Thesis Outline

Under consideration of the overall motivation as presented in Section 1.1, the presented thesis aims to address the in Section 1.2 described research problem, scope, goal, and objectives, as illustrated in Figure 1.1.

Chapter 2 establishes important foundational concepts that are relevant within the scope of the interdisciplinary work presented in this thesis. The general approach of VR as well as important definitions and key concepts are introduced, and relevant considerations with respect to human factors and ergonomics inherent from the practical application of VR approaches are described. An overview about 3D UIs is provided, focused on visual displays and 3D spatial input devices, including detailed descriptions in regard to HMDs and 3D gestural input as primary means of output and input hardware utilized within the scope of this thesis. Relevant concepts and terminology with respect to CVEs, and CSCW as closely related part thereof, are described. Furthermore, an overview about the emerging fields of IA and CIA is given, including relevant definitions, research opportunities, and core ideas in general. The chapter also reflects on the empirical nature inherent from the evaluation of immersive technologies, and provides a comprehensive summary about the various HCI evaluation methods that have been applied within the scope of this thesis. Finally, the chapter concludes with the construction and illustration of the thesis design space.

Chapter 3 describes and discusses relevant related work within the context of this thesis, focusing on four general subjects. First, works in regard to immersive data visualizations are explored that are built around a VR approach and under utilization of HMD devices. Second, an overview is provided about past research that is concerned with the investigation of applying 3D gestural input for interaction with data in immersive spaces. Third, relevant existing work is described that concerns hybrid collaboration experiences, i.e., cross-platform collaboration where at least one interface applied a VR approach. Fourth and finally, research in regard to the design of collaborative information cues in VEs is presented, providing inspiration and reflections for the subsequent design of similar information cues.

Chapter 4 is concerned with the presentation of the conceptual and technological system architecture that serves as a foundation for the implementation of

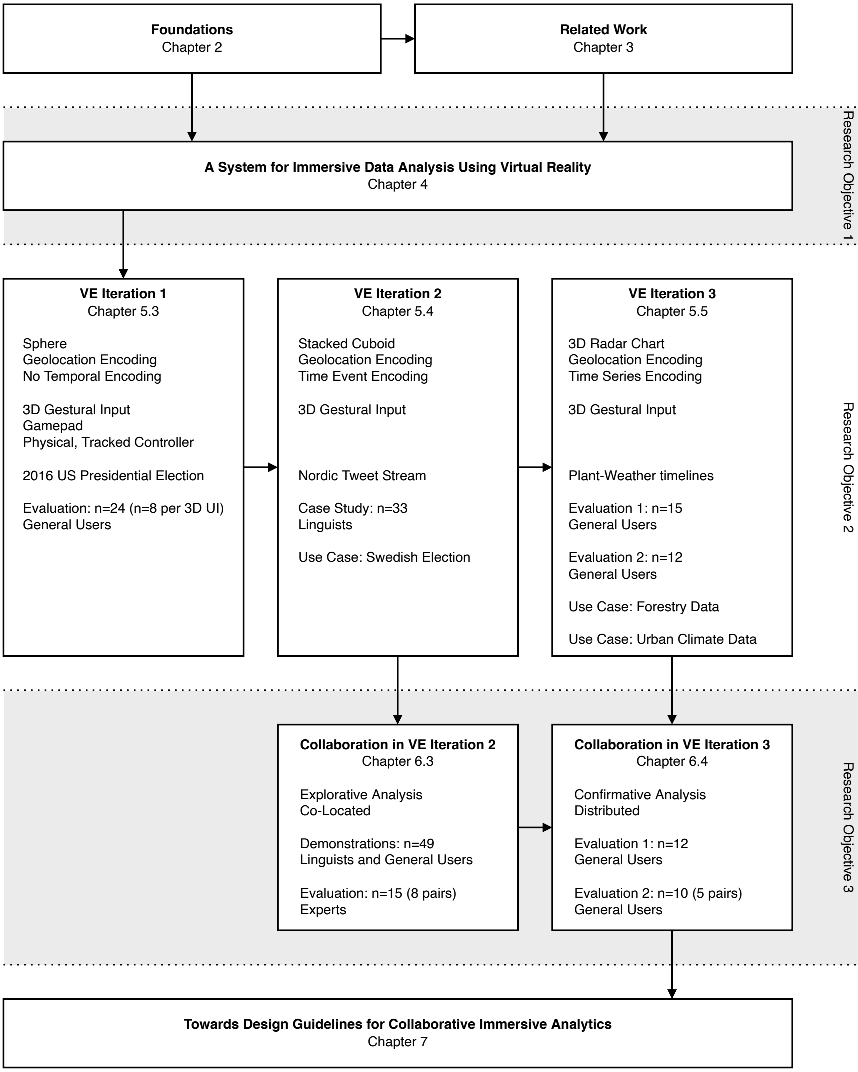


Figure 1.1: Illustration of the thesis structure, presenting the various chapters and their alignment to the three research objectives (see Section 1.2), outlining the logical progression of the presented research, as well as providing an overview of key aspects that characterize the developed and empirically evaluated immersive data analysis systems as VE iterations.

the developed immersive data analysis system presented within the scope of this thesis. It is dedicated to the first research objective, providing an overview of the defined system requirements and detailed descriptions about the four major building blocks that compose the system, namely (1) the *Data Structure Reference Model*, (2) the *Immersive Virtual Environment*, (3) the *User Session Data Transfer*, and (4) the *Collaboration Infrastructure*. The chapter concludes by providing details about the practical implementation of the developed system components.

Chapter 5 is dedicated to the second research objective and begins by reviewing key characteristics about spatio-temporal data and data analysis tasks within the context of interactive visualizations. The chapter presents the design, development, and evaluation of three major VE iterations, used to gradually explore immersive interaction with spatio-temporal data using a VR approach. A key difference between these iterations is how individual data items are visually represented in the VE. The first iteration utilizes *Spheres*, the second *Stacked Cuboids*, and the third a custom developed *3D Radar Chart* approach. The interaction design in each iteration is described, aimed to support the immersed user with the conduction of typical data analysis tasks. Empirical evaluations, case studies, and use cases across the three VE iterations are presented and discussed.

Chapter 6 is concerned with the investigation of the third research objective. It begins with a critical examination of collaboration and cross-platform aspects within the presented context, and leads to the definition of *Hybrid Asymmetric Collaboration*. The introduced concept aims to clearly distinguish between the utilization of immersive 3D and non-immersive 2D display and interaction technologies (hybrid) as well as assuming distinct user roles (asymmetric) in the collaborative data analysis activity. To aid with the empirical evaluation of collaborative activities within the presented context, the development of the *Spatio-Temporal Collaboration Questionnaire* (STCQ) is presented. The remainder of the chapter is dedicated to the description of two collaborative setups, closely aligned and integrated with the second and third VE iterations, describing the developed non-immersive interface prototypes as well as the integrated collaborative features design that enable the users to make spatial and temporal references. Subsequent empirical evaluations are presented and discussed accordingly.

Chapter 7 can be regarded as a synthesis of the previous chapters. It takes the obtained insights from the investigation of the three research objectives, and addresses the overall research goal through the proposal of ten CIA design guidelines, aiming to provide helpful directives for the design and development of future immersive data analysis systems.

Finally, Chapter 8 summarizes the empirical work and the research findings presented within this thesis, and provides an outlook as well as some directions for future work.

1.4 Ethical Considerations

The conducted research is not expected to have any direct ethical implications. However, as the conduction of the various empirical evaluations in the format of user interaction studies within the scope of the presented thesis involve human participants (see Sections 5.3.3, 5.4.3, 5.5.3, 5.5.5, 6.3.3, 6.4.4, and 6.4.6), it is important to address ethical considerations in general. The participants, all of them adults at the age of 18 and above, were involved in the evaluation of the developed immersive and non-immersive interfaces as users, and as such as test subjects. The personal data collected from the participants concerned only their identity and information relevant to assessing their ability to complete the tasks in a user interaction study, for instance their previous experience in a subject area or with various technologies. No information were collected on ethnicity, religion, sexual orientation, or other personal and private topics. The participants as well as the data collected from them were anonymized throughout this thesis, as well as in any presentations and publications as a result of the conducted research.

As with any data visualization, particularly within the context of InfoVis, VA, and IA, care was taken that lengths, areas, and volumes, as well as their proportions, are appropriately visualized in the respective interfaces, enabling the user to interpret the data without distortion or ambiguity.

In terms of the conducted user interaction studies, the participants wore a HMD device, described in detail in Section 2.2.2. In short, a HMD is a device that is worn on the head, similar to glasses or ski goggles, visually isolating the user from the physical real-world surroundings. Instead, the user is presented with computer-generated virtual content, which can be explored by moving the head and thus looking around. Within the scope of this thesis, two common consumer VR HMDs were utilized, i.e., the Oculus Rift CV1 and the HTC Vive. Although these technologies and devices are widely available and established on the mainstream market, it is possible that in rare cases the users may encounter what is referred to as VR sickness. VR sickness, discussed in detail as part of Section 2.1.2, can be described as any unintended and uncomfortable side effects that may occur when using a VR system, for instance nausea, dizziness, or fatigue. None of the conducted user interaction studies involved the targeted evocation of such symptoms. Instead, any developed immersive interface presented within the scope of this thesis was developed under consideration of established VR design guidelines and best practices as provided by the respective HMD manufacturers and informed by existing research, aiming to actively provide conditions that prevent such symptoms from occurring in the first place. This entailed considerations in regard to human factors and ergonomics in general, as well as ensuring that the developed artifacts were running stable and according to performance recommendations on its respective computer system.

Furthermore, there are some operational practicalities that are worth addressing inherent from the application of HMDs for VR experiences in general. The HMDs utilized within the scope of this thesis were physically connected via cable to an external computer system for respective data transfer. The cable is not visually represented in the computer-generated VE for the user to see. As the immersed user is physically moving around, the cable can in some cases become a tripping hazard. Within the scope of the empirical works presented in this thesis, all VR sessions were supervised, i.e., a researcher was physically present in proximity to the immersed user wearing the HMD, and if necessary realigning the cable to prevent interference. Another aspect is concerned with a VR system's safe interaction area, i.e., the physical real-world area in which they can move freely without obstacles. Any VR session conducted within the scope of this thesis provided such a prepared area, either directly in the research group lab, as for instance described in Section 5.3.3.1, or at open spaces in public demonstrations. Furthermore, based on the HMD firmware and general integration with the computer system, the calibrated safe interaction area is also displayed in the immersive VE. In particular, a visual bounding box is displayed in-situ when the immersed user is moving close to the calibrated area's boundaries, aiming to prevent them from leaving accordingly.

Generally, a strict protocol was followed when conducting user interaction studies that involve the usage of immersive VR technologies. First, participants were introduced to the overall context and scenario of the study and the tasks that were involved. As part of this introductory phase, the participants provided their informed consent to partake in the study, acknowledging that (1) the participation is voluntary, (2) they may terminate their participation at any point in time without explanation, (3) no sensitive personal information are collected, and (4) any data collection by the researchers is confidential. Furthermore, before conducting the actual task as part of a study, each participant was given the time – a "warm-up" phase – to get familiar with the hardware, i.e., the HMD, and the software, i.e., the developed interface. Realistically, a participant was wearing the HMD for a total of approximately 15 to 40 minutes, depending on the study. Once the actual task in the immersive VE was completed, the participants were asked to complete some questionnaires, and potentially answer some questions as part of an interview, which concluded the user interaction study.

All of the above is (1) common practice within the HCI research community, and (2) compatible with the ethical guidelines of the Swedish Research Council (2017) and the Norwegian National Committee For Research Ethics in Science and Technology (2016).

COVID-19 pandemic The global COVID-19 pandemic (SARS-CoV-2; March 11, 2020 – present) required additional considerations and the implementation of supplementary practical precautions for the conduction of user interaction studies with human participants in the controlled laboratory environment. More specifically, within the scope of the presented thesis, this concerns the studies presented in Sections 5.5.5, 6.4.4, and 6.4.6. These studies were conducted during the time from April to June 2021, for which approval through the respective head of department (Computer Science and Media Technology) at Linnæus University was received. During that period, pandemic related matters were closely observed on a daily basis, following (1) the national safety rules and recommendations according to The Public Health Agency of Sweden (Folkhälsomyndigheten),[1] (2) the regional safety rules and recommendations for Kronobergs län according to Emergency information from Swedish authorities (Krisinformation),[2] and (3) the local safety rules and recommendations according to Linnæus University (Linnéuniversitetet).[3] A study session was only conducted if all involved individuals, i.e., moderator and participant(s), reported themselves as symptom-free. The moderator was wearing a face mask at all times. Face masks and hand disinfection gel was freely and voluntarily available to each participant. Physical distance between moderator and each participant was kept at all times during a study, which required no physical contact at any time. The respective windows in the controlled laboratory environment were open at all times, ensuring appropriate ventilation. Furthermore, a study that involved two participants followed a remote setup, i.e., it was organized in a way that the participants were located in different office rooms, so that at no point in time they were located in the same one, ensuring recommended physical distancing at all times. All involved technical equipment was carefully sanitized between each study task.

[1] The Public Health Agency of Sweden. Covid-19. Retrieved June 1, 2022, from `https://www.folkhalsomyndigheten.se/smittskydd-beredskap/utbrott/aktuella-utbrott/covid-19/`

[2] Emergency information from Swedish authorities. Current rules and recommendations. Retrieved June 1, 2022, from `https://www.krisinformation.se/detta-kan-handa/handelser-och-storningar/20192/myndigheterna-om-det-nya-coronaviruset/coronapandemin-detta-galler-just-nu`

[3] Linnæus University. The Coronavirus and Covid-19: Information to students. Retrieved June 1, 2022, from `https://lnu.se/mot-linneuniversitetet/kontakta-och-besoka/kris-och-sakerhet/coronaviruset/`

Chapter 2

Foundations

Contents

To approach the investigation of the presented research goal and its objectives, as described in Section 1.2, a sound foundational understanding of relevant core themes and subjects is required. This chapter serves as means to obtain such fundamental knowledge by thoroughly examining the existing literature, particularly in regard to Virtual Reality, 3D User Interfaces, Collaborative Virtual Environments, Immersive Analytics, and various aspects that concern the evaluation of immersive technologies. The chapter concludes with the construction of the thesis design space, serving as a conceptual outline for the presented research.

2.1 Virtual Reality

The idea and concept of Virtual Reality (VR) have caused fascination for many years. The notion of diving into a virtual world, generated by computer systems, perceiving its artificial three-dimensional (3D) environment and interacting within it as one would in the real world, has been a subject of countless novels, such as *Neuromancer* (Gibson, 1984), *Snow Crash* (Stephenson, 1992), or the more recent *Ready Player One* (Cline, 2011). Except, it is not just an imagination. Researchers and computer scientists have been exploring possibilities for human users to interface with computer systems on a higher sensory level than just using two-dimensional (2D) display technologies that are operated through keyboard and pointer input for a long time. Widely regarded as one of the first such interfaces is the *head-mounted three dimensional display* prototype presented by Sutherland in 1968. Sutherland's (1968) system allowed a user to visually perceive a simple graphical wireframe object, generated by a computer system, in stereoscopic 3D. Using an apparatus with an optical system worn on the head, nowadays commonly referred to as head-mounted display (HMD), the system would translate the physical movements of its user and adapt the perspective to the graphically rendered wireframe object (Sutherland, 1968). Consequently, this would create an illusion of the virtual wireframe object floating in the real-world space, and allowing the HMD user to look at it from different viewing angles by naturally moving around (Sutherland, 1968). It is noteworthy that Sutherland's (1968) HMD did not strictly implement the concept of VR, as it did not provide a Virtual Environment (VE) in a sense that all visual stimuli perceived by the user were computer-generated. Instead, it implemented the concept of Augmented Reality, by (1) combining real and virtual objects in the real-world environment, (2) registering and aligning real and virtual objects with each other, and (3) running interactively in 3D and in real-time (van Krevelen and Poelman, 2010). Sutherland's (1968) HMD is considered the original augmented reality system (Lanman et al., 2014), as it featured all of its typical system components, i.e., display device, image rendering, head tracking, interaction, and model generation. This general concept should later be advanced to visually render entire VEs using a VR approach (Lanman et al., 2014).

2.1.1 Definitions and Key Concepts

Moving forward in time to today, the year of 2022, many technological advances have led to a variety of human-computer interface types that aim to further explore and implement the concept of VR. The definition of VR by LaValle helps to understand its concept formally on a foundational level:

> "*Inducing targeted behavior in an organism by using artificial sensory stimulation, while the organism has little or no awareness of the interference.*"
> – LaValle (2020, Chapter 1.1)

LaValle (2020, Chapter 1.1) highlights and discusses the four key components of his VR definition, i.e., *target behavior*, *organism*, *artificial sensory stimulation*, and *awareness*. In other words, a human or another type of being (organism) experiences something as specifically designed by a creator (target behavior).

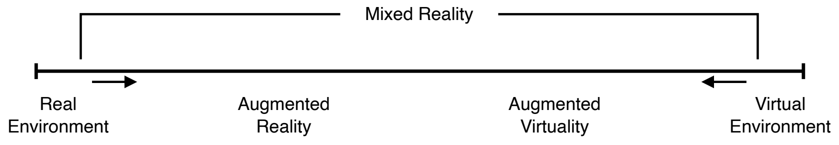


Figure 2.1: Reality-Virtuality Continuum, adapted from Milgram et al. (1995) as well as Milgram and Kishino (1994).

Furthermore, this is achieved by using technology that is able to closely connect to the organism's sensory system and stimulate it in accordance to the intended experience (artificial sensory stimulation). During such a designed experience, the organism has little awareness, if any at all, that the sensory stimulation is artificially engineered, ideally making the experience "feel real". Among others, LaValle's (2020, Chapter 1.1) definition of VR is general insofar as it does not specify, and thus limit, what sensory organ is going to be artificially stimulated. Focusing on human users as the main target group for VR experiences, interfaces exist to stimulate all five human sensory organs (Wallergård et al., 2022, Chapter 3.4.3; Burdea et al., 1996):

1. Visual interfaces to stimulate sight (eye);
2. Auditory interfaces to stimulate hearing (ear);
3. Somatosensory interfaces to stimulate touch (skin);
4. Olfactory interfaces to stimulate smell (nose);
5. Gustatory interfaces to stimulate taste (mouth).

Using such types of interfaces, either standalone by themselves or as a multimodal approach that combines multiple ones, it is possible to blend a user's experience with virtual information to various extents. The Reality-Virtuality Continuum (Milgram et al., 1995; Milgram and Kishino, 1994), illustrated in Figure 2.1, helps to formally understand the differences between various conceptual interface approaches. While one end of the continuum corresponds to a Real Environment, i.e., an environment without any additional computer-generated artifacts, the opposite end corresponds to a VE, i.e., an environment that consists entirely of virtual artifacts (Milgram et al., 1995). In between these two contrasting ends of the continuum, various types of Mixed Reality approaches can be established, essentially blending real and virtual environments to a certain extent (Milgram et al., 1995). Milgram et al. (1995) highlight the concepts of Augmented Reality and Augmented Virtuality as two prominent types of mixed reality. Augmented reality builds conceptually on extending the real environment with virtual information (Milgram et al., 1995; van Krevelen and Poelman, 2010).

Augmented virtuality, which can be regarded as the respective counterpart to augmented reality, aims to extend an environment that is mostly composed of virtual artifacts by integrating objects and information that originate in the real environment (Milgram et al., 1995). Milgram and Kishino (1994) center this taxonomy around three key dimensions:

1. *Extent of World Knowledge*: The extent to which the world is modeled, i.e., unmodeled (real environment) <u>versus</u> partially modeled (mixed reality) <u>versus</u> completely modeled (VE).
2. *Reproduction Fidelity*: The quality of the applied technology to render the designated artifacts, real as well as virtual.
3. *Extent of Presence Metaphor*: The extent to which the user is intended to feel fully integrated in the rendered scene, i.e., the extent of "feeling there".

An analysis of applied technologies towards these key dimensions provides guidance and consequently an important formal classification with respect to mixed reality and VR experiences. Additionally, Skarbez et al. (2021) provide some valuable reflections on Milgram and Kishino's (1994) taxonomy, among others arguing that based on today's technological possibilities and limitations, the original VE end of the continuum could be adapted to represent VEs as External Virtual Environments insofar as that they are computer systems that are still situated in the real world at the end of day.

At this stage, it is also important to clearly distinguish between the two terms VR and VE, as they are often applied interchangeably (LaViola, Jr. et al., 2017, Chapter 1.3). Generally, VR refers to the overall approach of using different types of interfaces to stimulate the human sensory organs in order to allow the user to immerse themself in a VE (LaViola, Jr. et al., 2017, Chapter 1.3). Therefore, a VE refers to the artificially computer-generated environment as experienced through the human user, i.e., from a first-person point of view, typically in 3D, and under real-time control (LaViola, Jr. et al., 2017, Chapter 1.3). Steinicke (2016, Preface) presents three substantial features according to the definition of VR by American computer scientist Frederick Phillips Brooks, Jr., that further assist with the clear distinction between VR and VE. These are arguably of particular relevance within the context of VR approaches that rely on the application of visual interfaces to stimulate sight. According to Brooks, Jr., VR features (Steinicke, 2016, Preface):

1. *Real-time Rendering*: The computer system dynamically updates the visually rendered artifacts in the VE in accordance to the user's respective body and head movements.
2. *Real Space*: The VE provides a genuine 3D space, either concrete or abstract, that is composed of 3D artifacts.
3. *Real Interaction*: The user can interact with the 3D artifacts in the VE through direct manipulation.

Within the context of describing and discussing VR, it is furthermore important to clearly distinguish between two other prominent terms, namely *immersion* and *presence*. Arguably, Slater has investigated and coined the discussion around these terms in the research community like few others (Slater, 2003; Slater and Wilbur, 1997; Slater and Usoh, 1993). Bowman and McMahan summarize the definitions of immersion and presence based on Slater's descriptions as follows:

> *"Immersion refers to the objective level of sensory fidelity a VR system provides. Presence refers to a user's subjective psychological response to a VR system."*
> – Bowman and McMahan (2007), adapted from Slater (2003)

A computer system's level of immersion is directly measurable from an objective point of view (Slater, 2003). That is to say, the more a computer system provides interfaces that artificially stimulate the user's sensory organs in line with their respective real-world modalities, the more *immersive* it can be assessed as (Slater, 2003). Consequently, some computer systems may be more immersive than others. Bowman and McMahan (2007) further highlight that a computer system's level of immersion may be dependent on hardware as well as software aspects. For instance, in the case of visual immersion, aspects such as the user's field of view, display size and resolution, stereoscopy (a display's ability to provide depth cues), realism of lighting, as well as frame rate and refresh rate, may all affect the level of immersion (Bowman and McMahan, 2007).

Presence however refers to the human's reaction to immersion based on the their perceptual and motor system as an individual context-dependent response (Bowman and McMahan, 2007; Slater, 2003). It is inherently subjective, as the same immersive computer system may cause different levels of presence for different users (Slater, 2003). Vice versa, it is also possible that the same level of presence experienced by different users may be caused by computer systems that feature a different level of immersion (Slater, 2003).

To facilitate the understanding of the difference between immersion and presence, Slater (2003) applies a quite helpful analogy, i.e., the description and perception of color. A color can be distinctively described according to its wavelength on the visible spectrum, but may be visually perceived differently by individual humans (Slater, 2003). Following that analogy, the description of color corresponds to immersion, while the perception of color corresponds to presence. Thus, immersion and presence are logically separable, although they are arguably empirically related (Wallergård et al., 2022, Chapter 10.2; Slater, 2003).

2.1.2 Human Factors and Ergonomics

The application of immersive display and interaction technologies for the implementation of VR approaches is aimed to closely interface with the human

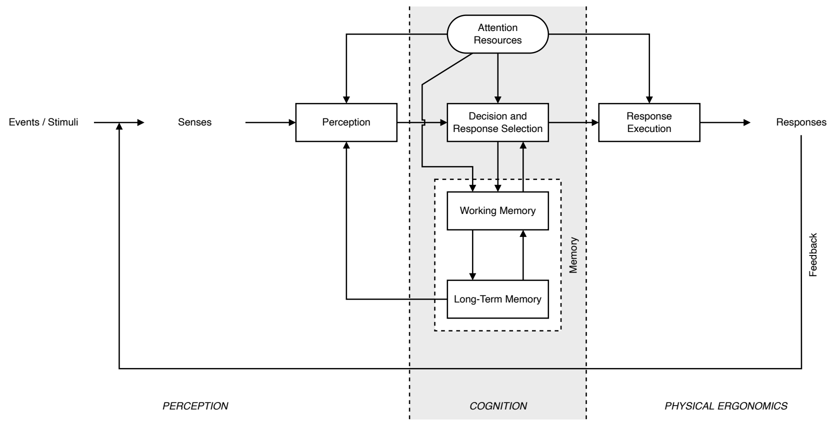


Figure 2.2: Human Information Processing Model, adapted from Wickens and Carswell (2021) and LaViola, Jr. et al. (2017, Chapter 3.2.1).

user to create experiences through artificial sensory stimuli that are under ideal circumstances indistinguishable from real ones. Therefore, a fundamental understanding of *human factors and ergonomics* is required (Rubio-Tamayo et al., 2017). The Human Information Processing Model according to Wickens and Carswell (2021), as illustrated in Figure 2.2, serves as a foundational high level framework to better comprehend this matter. Furthermore, the model can also be mapped to the different stages of human information processing according to psychological and physiological principles, i.e., *perception*, *cognition*, and *physical ergonomics* (LaViola, Jr. et al., 2017, Chapter 3.2.1).

Examining the model, some key processes become apparent (Wickens and Carswell, 2021). *Events or stimuli* are sensed through the human *sensory organs*. *Perception* allows for the meaningful interpretation of such sensed information based on collected *memories* from past experiences. Based on the now interpreted information, the human can make a *decision and select a response* that may be *executed* directly. Alternatively to the direct response execution, the interpreted information may also be stored in the memory, particularly in the *working memory*. The working memory is responsible for the temporary storage of recently interpreted information, has only a limited capacity, and is as such greatly influenced by the human's *attention resources*. In addition to influencing the decision and response selection process, attention resources impact which information are perceived as well as which responses can be performed concurrently. Information stored in the working memory may also be retrieved to further inform the decision and response selection accordingly. The working memory is closely related to another part of the human memory system, i.e., the *long-term memory*. Rather than just storing temporary information, the much higher capacity long-term memory is concerned with the storage of fundamental information about the world, its concepts, facts, and procedures, and thus in turn impacting its perception. Information stored in the long-term memory is retained without the special requirement for attention resources. Furthermore, interpreted information from the working memory may be transferred into long-term memory over time, and be retrieved respectively. Finally, executed responses may produce *feedback* that is sensed in turn, effectively closing the loop of the human information processing model and starting anew.

To summarize, perception is concerned with the understanding of the different information cues as sensed through the human sensory organs, i.e., the various visual, auditory, haptic, olfactory, and gustatory cues. Cognition is concerned with the understanding of how perceived information is processed, interpreted, stored, and recalled in order to initiate a response by deciding and selecting an appropriate action accordingly. Physical ergonomics is concerned with the understanding of human anatomy and physiology, enabling the interaction within a spatial environment, and ideally in a comfortable and effective manner.

With the obtained high level understanding of human information processing and the distinction between perception, cognition, and physical ergonomics, it is now possible to take a closer look at human factors and ergonomics within the specific context of VEs. Stanney et al. state the following question as a guiding principle for the design of VEs under consideration of an anticipated close integration with its human user:

> "*How should VE technology be improved to better meet the user's needs?*"
> – Stanney, Mourant, and Kennedy (1998)

To examine this question, Stanney et al. (1998) categorize human factor research for VEs in terms of *human performance efficiency*, *health and safety*, and *social implications*.

First of all, for a VE to be useful and effective, it has to be usable. In other words, applying designated VR technologies in accordance to the designed experience, the immersed user must be able to perform tasks in the VE. The evaluation of the immersed user's performance in a VE remains a complex endeavor that can be dependent on a multitude of aspects (Stanney et al., 1998). These include, but are not limited to, the user's ability to navigate in the VE (Stanney et al., 1998), the accuracy and speed of various interactions in the VE, for instance travel operations or the selection and manipulation of artifacts (LaViola, Jr. et al., 2017, Chapter 11), the experienced workload (Hart, 2006), or situation awareness (Wickens and Carswell, 2021; Vidulich and Tsang, 2015).

Second, due to the close coupling of user and computer system, human health and safety aspects are of particular importance within the presented context. A common umbrella term for unintended side effects that cause discomfort when using VR technologies is *VR sickness*, originating from *simulator sickness* and historically also referred to as *cybersickness* (Hirzle et al., 2021; LaValle, 2020, Chapter 12.3). Typical symptoms include among others dizziness, nausea, cold sweating, fatigue, or eye strain (LaValle, 2020, Chapter 12.3; Stanney et al., 1998). VR sickness symptoms are usually a result through a mismatch among the senses. A prominent example, not least because it is frequently reported on, is *visually induced apparent motion*, also referred to as *vection*. Vection is the illusion of self-motion due to visual stimuli tricking the brain into believing that one moves without an actual physical movement occurring (LaValle, 2020, Chapters 8.4 and 12.3). Besides sensory mismatches, it is also possible for VR sickness symptoms to occur as a result of flaws in the applied technologies, for instance imprecise sensors and tracking or unexpected happenings in the VE, ultimately causing inconsistencies with expected real-world experiences (LaValle, 2020, Chapter 12.3). It is noteworthy that VR sickness symptoms may endure the actual exposure with the VR system, and continue to be experienced even after having stopped using the system. This phenomenon is commonly referred to as *after effects* (LaValle, 2020, Chapter 12.3; LaViola, Jr. et al., 2017, Chapter 11.3.3; Stanney et al., 1998). Due to the often intricate architecture of VR systems as well as the complex nature of human psychology and physiology, the exact identification of causes for VR sickness symptoms remains challenging (Hirzle et al., 2021; LaValle, 2020, Chapter 12.3; Stanney et al., 1998). The prevention of any type of discomfort, harm, or injury should be the top priority from an ethical standpoint (Swedish Research Council, 2017; Norwegian National Committee For Research Ethics in Science and Technology, 2016).

Finally, while the application of VR approaches provides opportunities, it also poses certain risks in regard to social implications and aspects, demanding careful considerations. For instance, Stanney et al. (1998) highlight potential negative social impacts inherent from the excessive use of VR technologies and their anticipated immersive experiences in regard to escapism, social withdrawal, or change in social behavior. At the same time however, immersive VEs also provide an opportunity to create learning-rich environments that implement a more active "learning by doing" approach (Radianti et al., 2020; Stanney et al., 1998). The integration of VR technologies in real-world learning environments remains intriguing and promising, even if mostly experimental from a practical standpoint at this stage (Radianti et al., 2020). Such environments are comparatively safe and risk-free, allowing for the training of skills as well as the exploration of real-world phenomena that are otherwise not easily accessible to learners (Oyelere et al., 2020; Olbrich et al., 2018; Rubio-Tamayo et al., 2017). Safe and risk-free VEs have also been shown to assist with the treatment of anxiety and phobias (Bowman and McMahan, 2007). The application of VR technologies has also the potential to facilitate social interaction, commonly referred to as *Social VR* (LaValle, 2020, Chapter 10.4). The features of immersive VR technologies allow for more expressive interaction and communication in virtual 3D environments, following both visually abstract and realistic approaches (Wu et al., 2021; Sun et al., 2019; Heidicker et al., 2017). Furthermore, access to the Internet and its high data transfer capabilities allow for remote connection, meet-up, and interaction with other users in shared VEs in real-time, overcoming physical boundaries on a global scale (Sra et al., 2018; Perry, 2016). This is not just interesting from an entertainment perspective (Rubio-Tamayo et al., 2017), but also relevant for remote collaborative work (Snowdon et al., 2001).

2.2 3D User Interfaces

With a general conceptual understanding of VR and VEs, as obtained throughout Section 2.1, it now makes sense to examine various technological aspects, i.e., 3D user interfaces (3D UIs). In particular, the work presented in this thesis is concerned with the implementation of VR through visual interfaces to immerse the user in a computer-generated 3D VE. Rather than just passively looking around, perceiving visual stimuli and events, the user in turn is also expected to respond through active interaction in the VE – in line with the previously described human information processing model (see Section 2.1.2). Consequently, an interface that enables interaction in the 3D virtual space is required. The Human-VE Interaction Loop as presented by Bowman and McMahan (2007), illustrated in Figure 2.3, serves as a conceptual framework to facilitate this matter.

Examining the framework (Bowman and McMahan, 2007), a few key aspects become apparent. A *computer system* is responsible for generating and maintaining the 3D virtual space, i.e., the VE that functions in accordance to the designer's intent. LaValle (2020, Chapter 2.2) highlights that this computational process is commonly referred to as Virtual World Generator (VWG). Typically, the computer system has access to *models and data* that are used within the VE for one purpose or another, such as to populate the virtual space with respective artifacts. The computer system's *rendering software* is tasked with the generation of the VE's visual representation, which is then presented on a dedicated output or *display device*, from where it can be visually perceived by the user. *Input devices*, such as tracking sensors, physical controllers, or microphones, to name just a few, are responsible for detecting and capturing any type of user interaction. The computer system's *software* is responsible for *interpreting* the captured user *input*, and to update the VE, for instance by translating the position of the user in the VE in accordance to their physical real-world movements, or by executing functionalities with respect to the VE's logic in accordance to the user's commands. At that point, the computer system tasks the rendering software to update the visual representation of the VE, and starting the loop again. Figure 2.4 provides an exemplary setup of a VR system, allowing a user to be immersed in a VE through the application of various immersive display and interaction technologies, further illustrating the described human-VE interaction loop.

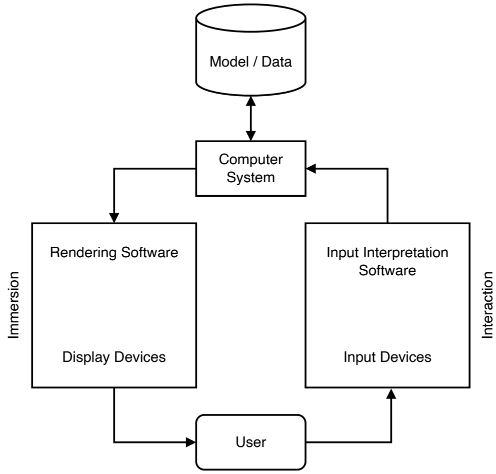


Figure 2.3: Human-Virtual Environment Interaction Loop, adapted from Bowman and McMahan (2007).

To provide further foundational information relevant within the scope of this thesis, the following sections are organized as follows. Section 2.2.1 provides a brief overview of visual displays as output hardware, allowing a user to visually immerse themself in a VE based on a VR approach. Thereafter, Section 2.2.2 describes in more detail the main type of visual display utilized within the presented research, i.e., the HMD device. Section 2.2.3 continues with a brief introduction to 3D spatial input devices, i.e., input hardware that enables a user to interact in 3D in the virtual space. Finally, Section 2.2.4 is concerned with 3D gestural input, as the presented research focuses on interaction through means of hand postures and gestures as main input modality.

2.2.1 Output Hardware: Visual Displays

In order to obtain a better understanding of visual display technologies within the context of 3D UIs and the specific approach of VR, is is important to discern relevant characteristics. LaViola, Jr. et al. (2017, Chapter 5.2.1) provide an overview of such characteristics.

The *field of regard*, measured in degrees, refers to the physical space a visual display surrounds its user with, while the *field of view* specifies the highest possible visual angle a user can see at a given point in time (LaViola, Jr. et al., 2017, Chapter 5.2.1). While the field of regard is a determined characteristic of a given visual display, the user's field of view can vary, for instance dependent on their position and orientation with respect to the display (LaViola, Jr. et al.,

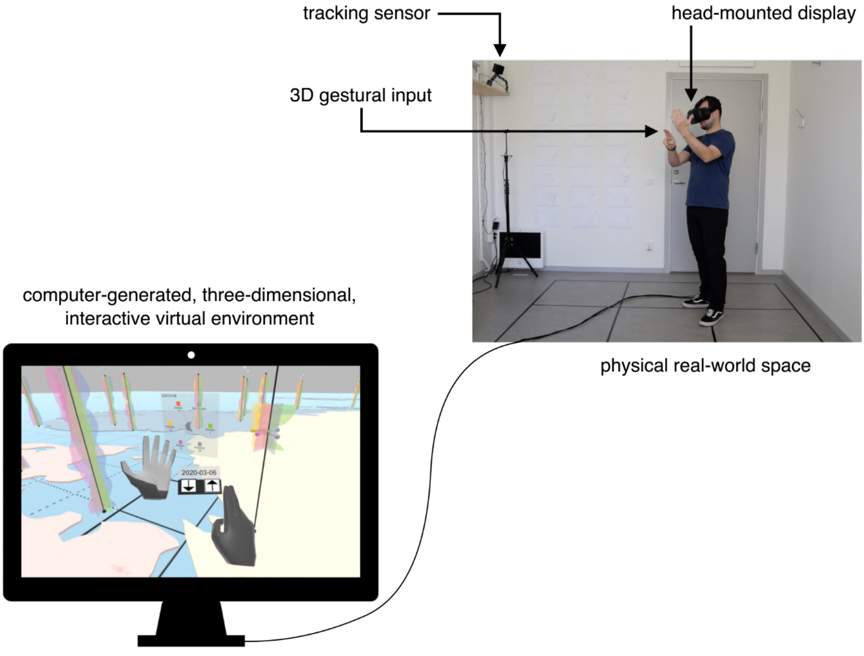


Figure 2.4: An example of a user interfacing with a VR system. A computer system generates a 3D interactive VE that is graphically rendered and displayed on a HMD. The user visually perceives the VE through the HMD by naturally moving around in the physical real-world space. The user's movements are tracked through various sensors, and processed as input by the computer system, tasked with translating these movements and updating the VE accordingly. A tracking sensor attached to the HMD additionally captures the user's hand movements, enabling 3D gestural input and thus allowing for 3D spatial interaction in the VE in accordance to the implemented interactive features.

2017, Chapter 5.2.1). As a reference, the field of view of human vision, i.e., more appropriately referred to as the human visual field, reaches under normal circumstances approximately 200 degrees horizontally and 150 degrees vertically (Steinicke et al., 2011).

The visual display's *spatial resolution* is determined through the consideration of its amount of pixels, commonly measured in pixels per inch, as well as its overall screen size (LaViola, Jr. et al., 2017, Chapter 5.2.1). With respect to the user's visual perception, their distance to the display also affects the perceived resolution, i.e., the closer a user is to the display, the more likely they will be able to recognize individual pixels, and vice versa (LaViola, Jr. et al., 2017, Chapter 5.2.1).

The *screen geometry* of a visual display is also an identifying characteristic (LaViola, Jr. et al., 2017, Chapter 5.2.1). Arguably most common are normal rectangular displays, but there are also L-shaped (basically two rectangular displays in perpendicular arrangement) and curved (circular or hemispherical) displays (LaViola, Jr. et al., 2017, Chapter 5.2.1).

The method of *light transfer*, for instance through a monitor, front or rear projection, or retinal projection through laser light, is another important property within the context of 3D UIs, as it can impact possible interaction modalities on a practical level (LaViola, Jr. et al., 2017, Chapter 5.2.1). For instance, LaViola, Jr. et al. (2017, Chapter 5.2.1) highlight that visual displays based on front projection may be ill-suited for direct 3D interaction as the user's hands are likely to get into the way of the projection, thus casting shadows onto the visual display as well as potentially projecting visual artifacts onto the user's hands.

The speed of a visual display's ability to update its displayed graphical contents is attributed as its *refresh rate* (LaViola, Jr. et al., 2017, Chapter 5.2.1). It is commonly measured in Hertz, describing how often per second a new image can be displayed on a hardware level (LaViola, Jr. et al., 2017, Chapter 5.2.1). A computer system's ability to generate new graphical images on a software level however is described as *frame rate* and measured in frames per second (LaValle, 2020, Chapter 6.2). It is noteworthy that while a computer system could be able to generate high frame rates, a visual display can only display them at its refresh rate limit (LaViola, Jr. et al., 2017, Chapter 5.2.1). Both refresh rate and frame rate are important with respect to the human's ability to visually perceive motion using visual display technologies and computer systems. After all, on a computational level, this is achieved by quickly stringing together still images, which is also referred to as *stroboscopic apparent motion* (LaValle, 2020, Chapter 6.2). Low refresh and frame rates can negatively influence the visual quality of a display, potentially introducing flickering effects or increasing *latency*, i.e., the timed delay it takes for a respective action to be displayed (LaViola, Jr. et al., 2017, Chapter 5.2.1). A particularly important type of latency within the context of immersive display technologies is *motion-to-photon latency*, describing the time it takes for a tracked object and its movements to be updated and appear graphically rendered on the visual display (Stauffert et al., 2020). Low refresh and frame rates as well as high latency can result in poor interaction experiences and even evoke symptoms of VR sickness (Stauffert et al., 2020; LaViola, Jr. et al., 2017, Chapter 5.2.1), as described in Section 2.1.2

Naturally, from the user's perspective, the *ergonomics* of a visual display also need to be considered (LaViola, Jr. et al., 2017, Chapter 5.2.1). This is particularly important in regard to comfort (weight, obstructiveness) in the case that the user themself have to actively wear any kind of "attachment", such as either the display itself or some kind of complementary eyewear (LaViola, Jr. et al., 2017, Chapter 5.2.1).

Finally, visual displays have the potential to facilitate a user's 3D spatial perception and understanding through the provision of various *depth cue effects* (Wallergård et al., 2022, Chapter 5.5; Bowman and McMahan, 2007):

- *Monocular* (static) depth cues can be extracted from images due to geometric distortions as a result of perspective projections (LaValle, 2020, Chapter 6.1.1; LaViola, Jr. et al., 2017, Chapter 3.3.1). Examples of monocular depth cues include occlusion/interposition (an object closer to the viewer occluding parts of another object that is farther away), height in the visual field (objects closer to the horizon appear farther away), perspective cues such as shadows (a distant shadow of an object makes the object appear higher), and atmospheric cues (lower contrast scenery appears to be more distant).
- *Motion Parallax* refers to monocular depth cues as a result of exploiting motion, in particular the phenomenon of visually perceiving closer moving objects as passing through the human visual field more quickly compared to more distant objects (LaValle, 2020, Chapter 6.1.1; LaViola, Jr. et al., 2017, Chapter 3.3.1). Motion parallax can occur as stationary-viewer parallax (the viewer is static, the object is moving), moving-viewer parallax (the viewer is moving, the object is static), or as a combination of both.
- *Stereopsis* refers to the ability to create a single stereoscopic image through the fusion of two slightly different images (binocular disparity) as perceived through two eyes, providing a powerful depth cue for objects in closer proximity (LaViola, Jr. et al., 2017, Chapter 3.3.1). Stereopsis is achieved through focusing both eyes on the same object using *accommodation* and *vergence* (LaValle, 2020, Chapters 4.4 and 5.3; LaViola, Jr. et al., 2017, Chapter 3.3.1). Accommodation describes the process of using eye muscles to physically stretch and relax the eye lens, changing the lens' diopter, and thus allowing to clearly focus on the desired object. Vergence describes the motion of rotating the eyes in order to align their focus with the desired object. Convergence refers to the process of moving the eye pupils closer together in order to focus on an object in closer proximity, while divergence refers to the process of moving them farther apart to focus on an object in the distance.

With the understanding of different characteristics, it is now possible to examine some relevant visual display device types. LaViola, Jr. et al. (2017, Chapter 5.2.1) highlight that basically any type of visual display is able to utilize monocular depth cues, while more specific types of displays, i.e., stereoscopic displays, are required to implement stereopsis by providing dedicated images for both the left and the right eye. For instance, through the utilization of complementary stereo glasses, also referred to as shutter glasses, it is possible to enable stereoscopic viewing even on single-screen displays (LaViola, Jr. et al., 2017, Chapter 5.2.2). This is typically implemented as an active (temporal multiplexing) or a passive (polarization or spectral multiplexing) shutter, a mechanism that essentially filters the displayed content for viewing through the left and right eye respectively (Chen et al., 2012, Chapter 9.1.1). There are even devices that can implement stereoscopic depth cues without the need for complementary eyewear, categorized as autostereoscopic display devices (Chen et al., 2012, Chapter 9.1.1). While potentially providing strong binocular depth cues, one can argue that all these types of visual displays do not provide a high degree of immersion due to the comparatively limited field of regard inherent from their single-screen nature.

Multi-screen displays and surround-screen systems intend to address this shortcoming by expanding the field of regard, aiming to literally "surround" the user with visual display space (LaViola, Jr. et al., 2017, Chapter 5.2.2). These can vary from desktop setups with multiple screens arranged in an angular manner, to installations that entirely encapsulate their user with displays (Sommer et al., 2017; Marai et al., 2016; Febretti et al., 2013). One of the earliest concepts and practical implementations for a surround-screen setup has been the Cave Automatic Virtual Environment, in short CAVE, an audio visual experience presented by Cruz-Neira et al. (1993, 1992). The CAVE utilized five projection-based screens (three walls, one floor, and one ceiling) that physically surrounded the user and generated a visual representation of a VE from the user's point of view to simulate a viewer-centered perspective through respective head tracking (Cruz-Neira et al., 1992). Furthermore, the visual projections were based on stereoscopic computer graphics, enabling the user to wear stereo glasses to experience stereoscopic depth cues (Cruz-Neira et al., 1992). Since it's original implementation, more research around CAVE-type systems has been conducted. For instance, there have been efforts to facilitate integration with modern game engines, enabling more accessible software development for these types of display systems (Lugrin et al., 2012). More advanced installations have also been developed, referred to as CAVE2 systems, among others allowing the display of 2D and stereoscopic 3D graphics as well as enabling multi-user support to provide immersive workspaces for more than one user at a time (Sommer et al., 2017; Febretti et al., 2013).

Compared to these comparatively large and rather stationary visual display setups, there exist also portable display solutions that are directly attached to the user. One such approach is the HMD. As its name suggests, a HMD is directly attached to the user's head in close proximity to their eyes, similar to a pair of goggles (LaViola, Jr. et al., 2017, Chapter 5.2.2; Chen et al., 2012, Chapter 10.4). Section 2.2.2 will provide a comprehensive overview of this type of visual display device. Chen et al. (2012, Chapters 9 and 10) provide an exhaustive overview of immersive display technologies, in particular in regard to 3D displays (including volumetric and holographic displays), mobile displays, microdisplays, projection systems, and headworn displays.

2.2.2 Head-Mounted Display

Within the presented thesis, the HMD device has been used as primary means to allow for visual immersion in a VE through a VR approach. Therefore, this section aims to provide a dedicated overview about this type of technology as well as to elaborate on advantages and potential shortcomings that are relevant to consider when developing VR experiences for HMDs.

As already illustrated in Figure 2.4, a HMD is a device that is "attached" to the user's head in front of their eyes, allowing the display to follow along as the user moves around in the physical real world. It is therefore also referred to as head-worn display or simply headset (LaViola, Jr. et al., 2017, Chapter 5.2.2; Chen et al., 2012, Chapter 10.4).[1] Various types of HMD devices exist, aiming to implement and address different approaches along the reality-virtuality continuum (see Section 2.1.1).

While there are even HMDs based on projection or virtual retina approaches, two comparatively common types of HMDs are *optical see-through* displays and *video see-through* displays (LaViola, Jr. et al., 2017, Chapter 5.2.2). Optical see-through displays typically feature a screen that is based on transparent material, similar to normal glasses or goggles, ultimately introducing a clean layer between the viewer's eyes and the real-world environment (LaViola, Jr. et al., 2017, Chapter 5.2.2). This approach enables the viewer to still visually perceive the real-world space without delay by looking through the transparent layer, while at the same time featuring a screen that can be overlayed with computer-generated graphics (LaViola, Jr. et al., 2017, Chapter 5.2.2). These types of HMDs are particularly interesting for the implementation of augmented reality approaches, allowing for an overlay of the real world with virtual information. Video see-through displays on the other hand utilize one or two small computer screens in front of the viewer's eyes (LaViola, Jr. et al., 2017, Chapter 5.2.2), and are commonly looked at through an additional set of special lenses to better accommodate the comparatively high field of view (LaValle, 2020, Chapter 7.3). To implement the see-through part and to provide stereoscopic depth cues, these devices typically feature a set of two cameras (one dedicated for the left eye, and one for the right) to record and provide real-world imagery (LaViola, Jr. et al., 2017, Chapter 5.2.2). The computer system processes the raw camera feeds, and can generate respective graphics that are displayed as an overlay on screen to the viewer (LaViola, Jr. et al., 2017, Chapter 5.2.2). While optical see-through displays feature no time delay in the view on the real world due to the nature of their approach, they are prone to provide a comparatively limited field of view for the overlay of computer-generated virtual artifacts (LaViola, Jr. et al., 2017, Chapter 5.2.2). Systems based on video see-through display technologies can overcome this limitation by enabling a wider field of view, but instead have to

[1] The term HMD is adopted to refer to this type of visual display device throughout the thesis.

tackle the challenge of providing low latency video throughput that is responsive enough to align the viewer's head movements with the displayed real-world imagery in near real-time in order to avoid sickness symptoms (LaViola, Jr. et al., 2017, Chapter 5.2.2). While video see-through displays are capable of providing augmented reality experiences, they are in fact ideal for the implementation of VR experiences, utilizing their screens to solely display computer-generated graphics to the user, and thus disregarding their camera system during the designed VR experience. Devices for that purpose are also commonly referred to as VR HMDs or simply VR headsets.

VR HMDs generally attempt to implement stereoscopic depth cues by rendering two dedicated images, i.e., one that corresponds to the viewpoint of the user's left eye, and one that corresponds to the user's right eye. With respect to the display, this can be achieved by either using two displays (one for each eye), or by using one larger display unit that simply renders the two images side by side. As briefly described in Section 2.2.1, the perception of stereoscopic depth cues relies on fixating the eyes on an object in the 3D space through accommodation (eye lens adjustment) and vergence (eye rotation). In natural settings, vergence and accommodative stimuli are aligned to fixate on the same object at the same distance (Banks et al., 2013). However, this is not possible with conventional displays that implement stereoscopic viewing (Banks et al., 2013). While eye rotation (vergence) can still be appropriately adjusted to fixate on various objects in the VE to visually perceive depth cues correctly, the focal length (accommodation) remains generally the same due to the static physical distance between the viewer's eyes and the display (Banks et al., 2013). This conflict is known as *vergence-accommodation mismatch* and illustrated in Figure 2.5 (LaValle, 2020, Chapter 5.4; Banks et al., 2013). As the user is likely trying to subconsciously adapt to the mismatching signals, they may experience symptoms of discomfort, fatigue, or eye strain as a result, especially over a longer period of time that goes beyond "just a few minutes"-VR experiences (LaValle, 2020, Chapter 5.4; Banks et al., 2013).

With respect to the human-VE interaction loop as described by Bowman and McMahan (2007), the HMD has been mainly considered as an output device so far, in particular for visual output. However, the HMD also features some important user input, and thus should be more accurately classified as a hybrid device. Considering the overall concept of a HMD, the user's head movements need to be measured in order to appropriately translate their viewpoint in the VE. In practice, this is commonly achieved twofold. First, the HMD device itself features an integrated sensory system that attempts to capture the user's head rotation. This system is referred to as *inertial measurement unit*, and commonly consists of a gyroscope to measure the angular rotation rate and an accelerometer to measure the change in velocity (LaValle, 2020, Chapter 2.1). The inertial measurement unit can also contain a magnetometer to measure the sensor's

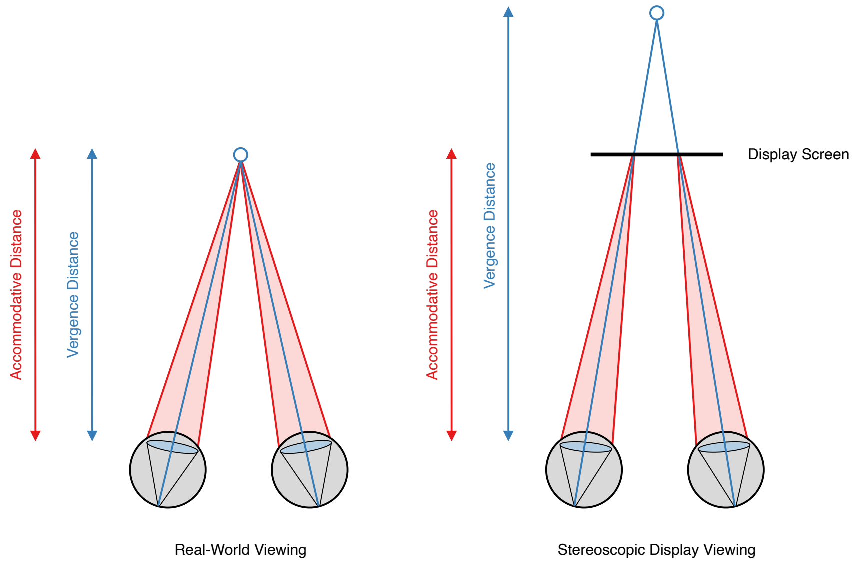


Figure 2.5: Vergence-Accommodation Mismatch, adapted from Banks et al. (2013). **Left**: Normal real-world viewing with no mismatch. **Right**: Stereoscopic display viewing with mismatch.

surrounding local magnetic field (LaValle, 2020, Chapter 2.1). The measurements of both the accelerometer and the magnetometer aim to reduce the potential *drift error* as a result of estimating the overall change in orientation measured through the gyroscope over time, therefore aiming to provide a more precise orientation measurement (LaValle, 2020, Chapter 2.1). In addition to the integrated inertial measurement unit, HMD devices commonly feature some kind of complementary camera system that is responsible for measuring the position and orientation of the HMD with respect to the physical real-world space (LaValle, 2020, Chapter 2.1). Generally, one can discern between two approaches for the implementation of camera-based HMD tracking, i.e., *outside-in* and *inside-out* tracking (Gourlay and Held, 2017). Following an outside-in tracking approach, one or multiple cameras are stationed in a role as sensors *outside* a dedicated tracking area and are looking *in* to determine the position and orientation of the HMD device (Gourlay and Held, 2017). Reverse in concept, following an inside-out tracking approach, the HMD itself features sensors *inside* that are looking *out* in order to determine its position and orientation. Inside-out tracking can be implemented through markers that are placed within the physical real-world environment, or through scanning of the physical real-world space for distinctive features, for instance by using special (infrared) depth cameras (Gourlay and Held, 2017). Noteworthy to mention is also the *lighthouse*-based approach developed by Valve and for instance used within the HTC Vive HMD system (Gourlay and Held, 2017). Their approach can be considered a special type of inside-out tracking that, instead of using markers, utilizes base stations that emit infrared light that in turn is sensed by various infrared sensors attached on the HMD itself (Gourlay and Held, 2017).

A variety of different HMD devices has been released in recent years, from different manufacturers and in different iterations (Kugler, 2021). Within the scope of the research presented in this thesis, two off-the-shelf HMD devices have been utilized, namely the Oculus Rift CV1[2] and the HTC Vive.[3] Figure 2.6 shows a photo of the two HMD devices side by side. Table 2.1 provides an overview of some technical specifications of the Oculus Rift CV1[4] and the HTC Vive,[5] also in regard to important display characteristics.

2.2.3 Input Hardware: 3D Spatial Input Devices

To obtain a better understanding of 3D input devices that enable interaction in VEs, it is helpful to review some typical device characteristics in general, for instance as summarized by LaViola, Jr. et al. (2017, Chapter 6.1.1).

Within the context of using input devices for the interaction in 3D virtual spaces, a particularly descriptive metric is a device's *degrees of freedom*, i.e., a numerical value that represents a device's capability of moving in space (Wallergård et al., 2022, Chapter 8.2; LaViola, Jr. et al., 2017, Chapter 6.1.1; Mackinlay et al., 1990). 3D spatial input devices typically feature three linear degrees of freedom, i.e., the position in space along 3D (x, y, z), as well as three rotary degrees of freedom, i.e., the rotation in space along 3D (yaw, pitch, roll), thus featuring a total of six degrees of freedom (LaValle, 2020, Chapter 3.2; Mackinlay et al., 1990). Generally, an input device's degrees of freedom can be seen as a reference to it's complexity (LaViola, Jr. et al., 2017, Chapter 6.1.1).

The *sensor type* (active or passive) provides another property to describe input devices (LaViola, Jr. et al., 2017, Chapter 6.1.1). Sensors are classified as active if they require some kind of direct manipulation through the user, for instance by pressing a button on a gamepad or moving a device in space (LaViola, Jr. et al., 2017, Chapter 6.1.1). Passive sensors are decoupled from the user and typically placed somewhere in the physical real-world environment, attempting to capture data indirectly, i.e., without the need for the user to actively manipulate the position or orientation of the sensor itself (LaViola, Jr. et al., 2017, Chapter 6.1.1).

---

[2]The first generation of the Oculus Rift HMD released on the consumer market (CV1) has been used, as released in 2016, and developed by Oculus VR (formerly; now Meta Quest).

[3]The first generation of the HTC Vive HMD device has been used, as released in 2016, and developed as a collaboration between HTC Corporation and Valve Corporation (Valve Software).

[4]Based on official manufacturer specifications, and: iFixit. Oculus Rift CV1 Repair. Retrieved June 1, 2022, from `https://www.ifixit.com/Device/Oculus_Rift_CV1`

[5]Based on official manufacturer specifications, and: iFixit. HTC Vive Repair. Retrieved June 1, 2022, from `https://www.ifixit.com/Device/HTC_Vive`

Figure 2.6: Two off-the-shelf HMD devices. **Left**: Oculus Rift CV1. **Right**: HTC Vive. **Note**: Both HMD devices feature an optional attachment in the front for the Leap Motion Controller (described in Section 2.2.4).

Property	**Oculus Rift CV1**	**HTC Vive**
light transfer	2x OLED	2x AMOLED
display resolution	1080x1200 pixels (eye) 2160x1200 pixels (total)	1080x1200 pixels (eye) 2160x1200 pixels (total)
sensors	inertial measurement unit (gyroscope, accelerometer, magnetometer)	inertial measurement unit (gyroscope, accelerometer)
tracking	outside-in: 360° headset tracking via constellation infrared camera	inside-out: 360° headset tracking via lighthouse emitters
field of regard	360°	360°
field of view	>100°	110°
spatial resolution	~456 pixels per inch	~447 pixels per inch
screen geometry	rectangular	rectangular
refresh rate	90 Hertz	90 Hertz
ergonomics	470 grams	555 grams

Table 2.1: Overview of some technical specifications and characteristics of the two HMD devices Oculus Rift CV1 and HTC Vive.

Another descriptive property is the *data frequency* a device utilizes to capture input from the user (LaViola, Jr. et al., 2017, Chapter 6.1.1). Data frequency can be categorized as discrete (a descriptive single value input at a time) or continuous (constant capture of numerical values as input over time) as a result of user interaction (LaViola, Jr. et al., 2017, Chapter 6.1.1). For instance, pressing the physical button on a device corresponds to discrete input, while the on-going tracking of a device in space corresponds to continuous input (LaViola, Jr. et al., 2017, Chapter 6.1.1). Naturally, some input devices may feature both discrete and continuous user input, for instance a physical, tracked controller that is held in the user's hand and that features spatial tracking capabilities as well as physical buttons (LaViola, Jr. et al., 2017, Chapter 6.1.1).

Furthermore, there are arguably some additional properties that are particularly relevant with respect to interactive VR experiences. The *physicality* of an input device may be relevant, i.e., whether or not the user is required to have a physical device in their hands (by holding or wearing it). Additionally, an input device's capabilities for *visual representation* of itself or the respective user input in the VE may also be relevant, for instance affecting the design of appropriate interaction mechanisms or the perceived user experience.

In order to enable spatial interaction in the 3D virtual space, an input device needs to be appropriately tracked, i.e., with the objective to translate its position and orientation from the real-world space into the VE as precisely as possible and ideally in real-time with no perceived latency. Different tracking approaches as user input have already been briefly described in Section 2.2.2 as part of the tracking capabilities of HMD devices. Furthermore, Wallergård et al. (2022, Chapter 9) and LaViola, Jr. et al. (2017, Chapter 6.3.1) provide an exhaustive overview of different tracking approaches, including magnetic, mechanical, acoustic, inertial, optical, radar, bioelectric, and hybrid sensing.

A multitude of different input devices exist in order to capture user input in the 3D space (LaViola, Jr. et al., 2017, Chapters 6.3.2–6.6.2). Under assumption of utilizing hands and arms as main input and thus interaction modality, a general distinction of input devices can be applied as follows:

1. A *physical input device* that the user is *holding* with their hands. Such controllers commonly feature a variety of components, among others buttons, switches, joysticks, touchpads, haptic feedback, and pressure sensitive grips (Kangas et al., 2022; Duane and Þór Jónsson, 2021; Figueiredo et al., 2018). Naturally, any kind of interactive feature provided in the VE needs to be mapped onto such components of the device to enable user interaction.

2. A *physical input device* that the user is *wearing* on their hands or arms, such as a glove or a bracelet. These devices normally aim to provide a more intuitive

Figure 2.7: Various off-the-shelf physical, tracked controllers as 3D spatial input devices. **Left**: PlayStation Move. **Center**: Oculus Touch. **Right**: HTC Vive.

and arguably more natural[6] interaction by using hands as one would do for interaction in the real world. The physical component architecture of these devices commonly attempt to utilize integrated tracking capabilities as well as additional sensory feedback, such as haptic feedback (Liu et al., 2019; Olbrich et al., 2018).

3. *Vision-based input* through hand posture and gesture tracking that does not require the user to actively hold or wear any additional sensors with their hands or arms (Pavlovic et al., 1997). Therefore, such approaches can be described as contact-free or accessory-free input devices, essentially also aiming to provide intuitive hand interaction mechanisms (Koutsabasis and Vogiatzidakis, 2019; Bachmann et al., 2018).

Figure 2.7 shows a photo of some physical, tracked controllers as 3D spatial input devices that can be utilized for interaction in VEs. Within the scope of this thesis, the utilization of 3D gestural input, i.e., hand interaction, has been of particular interest for the interaction in the 3D virtual space. As such, Section 2.2.4 will provide a comprehensive overview of 3D gestural input.

[6] Norman (2010) provides a critical view on the terminology and understanding of *natural user interfaces* that is still relevant today, arguing that they are often not natural but useful.

2.2.4 3D Gestural Input

The research presented in this thesis is centered around the utilization of 3D gestural input as primary modality to allow user interaction in a VE through a VR approach. Thus, this section aims to provide a dedicated overview about this type of spatial user input, aligned with foundational interaction techniques and relevant considerations for the design of 3D UIs that utilize 3D gestural input.

As briefly stated in Section 2.2.3, interfaces that enable 3D gestural input aim to sense and track a user's hand movements. This is most commonly implemented through either physical controllers a user is wearing (Liu et al., 2019) or vision-based approaches without physical attachments to a user's hands (Pavlovic et al., 1997). The overall concept of 3D gestural input, independent of its technological implementation, is also referred to as *mid-air interaction* (Koutsabasis and Vogiatzidakis, 2019). Mid-air interaction has been of interest to Human-Computer Interaction (HCI) researchers for many years, as simply using ones hands to interact in a 3D VE has arguably something appealing, tending to evoke a certain feeling of naturalness (Bolt, 1980).

Human hands are comparatively complex instruments, as illustrated in Figure 2.8, and under normal circumstances utilized in a variety of contexts and for different purposes. It is therefore useful to obtain an understanding of foundational concepts relevant within HCI. Gestural commands are broadly classified as *postures* and *gestures* (LaViola, Jr. et al., 2017, Chapter 9.7). Postures refer to the hand being in a specific (static) configuration, while gestures refer to (dynamic) hand movements, possibly while being in a certain posture (LaViola, Jr. et al., 2017, Chapter 9.7). Modern tracking technologies are often capable of detecting both hands at the same time, allowing for subsequent interaction with either one hand (*unimanual*) or two hands hands (*bimanual*). Hence, bimanual gestural commands can further be classified with respect to hand *symmetry*, i.e., symmetric or asymmetric gestural commands, and hand *synchronicity*, i.e., synchronous or asynchronous gestural commands (Ulinski et al., 2009).

Pavlovic et al. (1997) propose a taxonomy of hand gestures within the context of HCI, illustrated in Figure 2.9, aiming to classify hand and arm movements with respect to their purpose. On a high level, hand and arm movements may be *unintentional*, i.e., without a purpose and intent, or *gestures*, i.e., deliberately performed towards a desired intent. Intended gestures can serve a *manipulative* or *communicative* purpose. Manipulative gestures refer to the direct manipulation of artifacts, such as moving and rotating virtual objects, interacting with menus and widgets, or the like. Inherent communicational purposes are mediated through communicative gestures, typically as *acts*, i.e., in direct relation to a specific movement, or as *symbols*, i.e., as accompanying gestures to an verbal expression.

Similar to the gesture taxonomy proposed by Pavlovic et al. (1997), Nehaniv et al. (2005) present a classification for gestures with respect to a user's intent, specifically within the context of interacting with another entity. Although

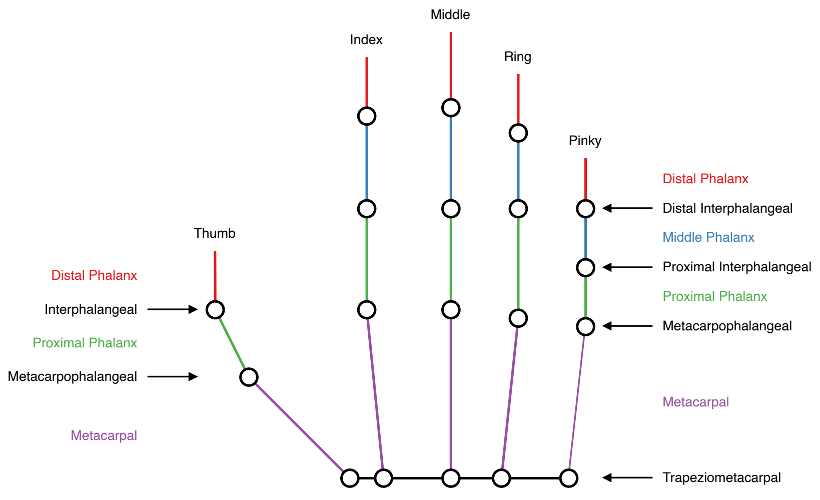


Figure 2.8: Human Hand Model (right hand), adapted from Bachmann et al. (2018) as well as Kuch and Huang (1995). Based on the anatomy of a human hand, it features internally 23 degrees of freedom. Additionally, the base of the palm features three degrees of freedom that allow for the overall orientation of the human hand in the 3D space. Consequently, the human hand model features a combined total of 26 degrees of freedom.

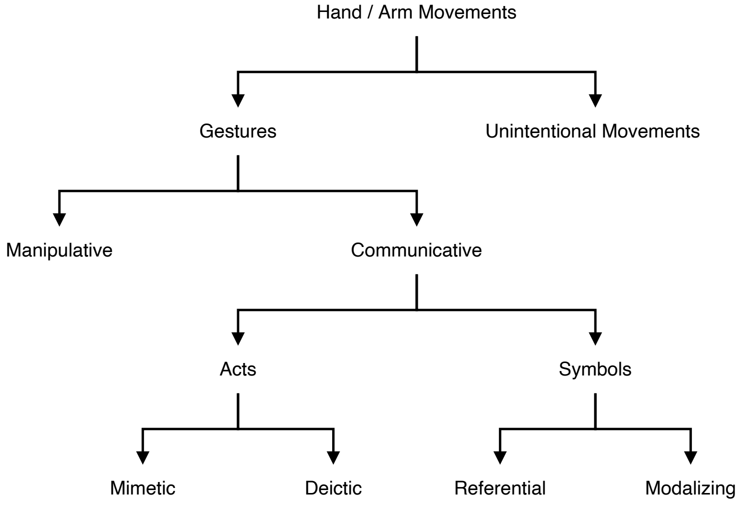


Figure 2.9: HCI Hand Gesture Taxonomy, adapted from Pavlovic et al. (1997).

originating from the context of Human-Robot Interaction (Nehaniv et al., 2005), their classification can be relevant for the interaction in collaborative settings, such as later described in Section 2.3. The classification is centered around the composition of five gestural classes, i.e., (1) *irrelevant and manipulative gestures*, (2) gestures as a *side effect of expressive behavior*, (3) *symbolic gestures*, (4) *interactional gestures*, and (5) *referential and pointing gestures* (Nehaniv et al., 2005). Within the context of interacting with other users in a VE, this classification can provide guiding principles for the design and analysis of multi-user interactions that rely on 3D gestural input. Furthermore, a computer system's ability to infer user intent is also important with respect to the various in-situ contexts a user may find themself in, aiming to implement more robust interaction mechanisms.

With respect to the actual interaction design, LaViola, Jr. et al. (2017, Chapters 7–9) provide an extensive overview about general 3D interaction techniques, describing approaches and metaphors for *selection*, *manipulation*, and *travel* operations as well as *system control* techniques. Within the context of 3D gestural interaction, *grasping* metaphors (LaViola, Jr. et al., 2017, Chapter 7.4) are particularly relevant, allowing the user to simply grab, move, and release artifacts in the VE, as one would in the real world, to enable subsequent selection and manipulation operations. The provision of *gestural commands* (LaViola, Jr. et al., 2017, Chapter 9.7) is likewise relevant, aiming to intuitively map static hand postures and dynamic hand gestures onto desired system control functionalities in the VE. The utilization of graphical menus in the 3D VE is also a common system control technique, enabling user interaction through the selection of menu items that are associated with respective features (LaViola, Jr. et al., 2017, Chapter 9.5; Dachselt and Hübner, 2007). Dachselt and Hübner (2007) provide an in-depth look at existing work, subsequently proposing a detailed taxonomy for the classification of 3D graphical menus that is centered around seven key dimensions, namely (1) *intention of use*, (2) *appearance and structure*, (3) *placement*, (4) *invocation and availability*, (5) *interaction and input/output setting*, (6) *usability*, and (7) *combinability*. Furthermore, one can differentiate between *direct* and *indirect* interactions (LaViola, Jr. et al., 2017, Chapter 7.7). Direct interactions are those that allow the immediate manipulation of an artifact itself, while indirect interactions commonly involve some kind of proxy or middle-layer for the manipulation of an artifact. Interactions with a representative copy that is linked to the original artifact, or 3D widgets that allow artifact manipulations through additionally visible control handles, are typical examples for such indirect interaction techniques (LaViola, Jr. et al., 2017, Chapter 7.7). While direct interaction techniques tend to be perceived as somewhat more natural than indirect ones, as they resemble similar interactions in the real world, indirect interaction techniques can be useful with respect to their intended purpose when designed appropriately (Norman, 2010). Finally, there exist also *multimodal* approaches that rely on the utilization of more than one input modality (LaViola, Jr. et al., 2017, Chapter 9.9), for instance 3D gestural input in combination with voice input (Bolt, 1980). While the implementation of such multimodal approaches can facilitate effective interactions in a VE, it's design should also be carefully considered due to the sometimes increased cognitive load that is required from the user for its operation (LaViola, Jr. et al., 2017, Chapter 9.9). Finally, it is noteworthy that all these techniques are relevant to 3D gestural input and interaction, but not exclusive to this modality. For instance, a mechanism to grab, move, and release an artifact in the VE may also be mapped onto other 3D spatial input devices that feature similar capabilities compared to 3D gestural input, for instance a physical, tracked controller.

With respect to the interface design that utilizes 3D gestural input for spatial interaction in the virtual 3D space, there are several matters worth considering. While one can argue that interaction with visible artifacts in a VE can be intuitively discovered, by allowing the user to simply touch, grab, move, and release them, hand postures and gestures that are invisibly linked to specific functionalities require introduction and possibly training (Delamare et al., 2016). Consequently, appropriate user feedback is helpful to support the user in their understanding of the current interaction state (Delamare et al., 2016). Of particular interest to HCI researchers is also the investigation of gesture appropriateness, the determination of whether the interaction feels intuitive and natural, user capability to memorize and recall implemented postures and gestures, and aspects of comfort (Koutsabasis and Vogiatzidakis, 2019). For instance, requiring the user to frequently apply mid-air interactions in ergonomically rather uncomfortable hand configurations can quickly evoke symptoms of physical fatigue from the extended arms – a phenomenon that is also commonly referred to as "gorilla arms" (LaValle, 2020, Chapter 10.3). Considerations for the design of comfortable hand postures for the utilization in HCI contexts have been reported by Rempel et al. (2014). Their insights originate from an analysis of hand postures and gestures used by sign language interpreters, and are aimed to prevent physical fatigue symptoms in hands and arms (Rempel et al., 2014). The authors identified various comfortable and uncomfortable hand posture configurations that are comparatively common. They recommend the use of comfortable postures for more frequent tasks, while infrequent tasks may also be performed through slightly less comfortable ones (Rempel et al., 2014). Furthermore, Rempel et al. (2014) also highlight the need to investigate aspects of hand posture and gesture learnability more specifically, as this is a topic that is comparatively underexplored. Their argument is in line with the review findings reported by Koutsabasis and Vogiatzidakis (2019), who highlight this matter as well.

Besides the previously mentioned approach of the user wearing an input device to allow for 3D gestural input (Liu et al., 2019; Olbrich et al., 2018), a comparatively popular interface is the Leap Motion Controller that implements a vision-based input approach (Koutsabasis and Vogiatzidakis, 2019; Bachmann

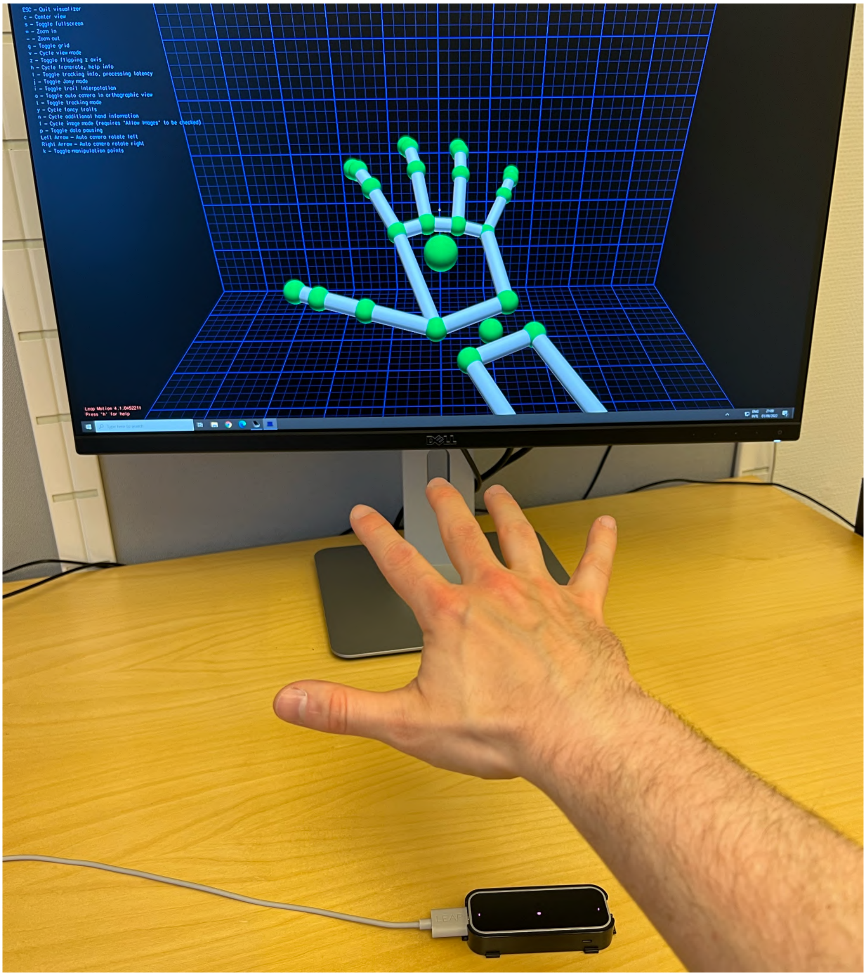


Figure 2.10: An example demonstrating the hand detection using the Leap Motion Controller. **Bottom**: The Leap Motion Controller set up on the desk, connected to a computer system. **Middle**: The user's hand is held mid-air above the Leap Motion Controller in it's interaction zone. **Top**: The Leap Motion Diagnostic Visualizer application is running on the computer system, displaying the detected hand in real-time on the monitor.

Property	Leap Motion Controller
tracking	vision-based (optical): three infrared LED emitters and two near-infrared cameras
interaction zone	~10–80 cm, extended by 120x150° field of view
refresh rate	120 Hertz
degrees of freedom	26 (see Figure 2.8)
sensor type	passive
data frequency	continuous
physicality	contact-free, accessory-free
visual representation	yes (within a VE)
ergonomics	32 grams, tabletop and HMD-attached setups available

Table 2.2: Overview of some technical specifications and characteristics of the Leap Motion Controller.

et al., 2018). The Leap Motion Controller[7] has been used as the primary device to enable 3D gestural input within the scope of the research presented in this thesis. According to Bachmann et al. (2018), interactive systems that utilize 3D gestural input via Leap Motion Controller feature three key processes, namely (1) *data acquisition*, i.e., transfer of the detected hand coordinates into the application's overall 3D space, (2) *feature extraction*, i.e., determine distinct features and attributes of the detected hand, as well as (3) *gesture definition and recognition*, i.e., construct specific hand postures and gestures that the application is tasked to detect at runtime, allowing for execution of the mapped functionalities accordingly. Figure 2.10 presents an example demonstrating the hand detection using the Leap Motion Controller, while Table 2.2 provides an overview of some technical specifications,[8] also in regard to important input device characteristics.

2.3 Collaborative Virtual Environments

As part of examining key concepts related to VR in Section 2.1.1, a VE was defined as an artificially generated environment that is experienced by the user, commonly from a first-person perspective, in 3D, and under real-time control. Immersive display and interaction technologies tend to be by default rather single user-centric in nature (Skarbez et al., 2019; Cordeil et al., 2017b; Hackathorn and Margolis, 2016). For instance, wearing a HMD, the user is visually isolated from their real-world surroundings. At the same time however, enabled by modern communication technologies and the Internet as a networking infrastructure, it is possible to create VEs that can be inhabited by more than one user. Allowing

[7]The first generation of the Leap Motion Controller has been used, as released in 2013, and developed by Leap Motion (formerly; now Ultraleap).

[8]Based on official manufacturer specifications, and the review presented by Bachmann et al. (2018).

users to immerse themselves in such shared virtual spaces holds great potential to remove spatial boundaries, and thus bringing users closer together (LaValle, 2020, Chapter 10.4). Naturally, such multi-user environments can be utilized for different purposes, from games and entertainment-related ones to those that are aimed to provide shared virtual workspaces to enable collaboration (de Belen et al., 2019; Perry, 2016; Churchill and Snowdon, 1998). Those latter ones are also commonly referred to as Collaborative Virtual Environments (CVEs), and particularly relevant within the scope of this thesis as data exploration and analysis are seldom solitary activities but collaborative ones that build on multiple users sharing their knowledge and expertise (Billinghurst et al., 2018; Isenberg et al., 2011; Heer and Agrawala, 2008).

Research that is concerned with CVEs is inherently rooted in Computer-Supported Cooperative Work (CSCW), utilizing computer technologies to allow multiple users in a shared workspace to interact with each other to conduct collaborative tasks through means of communication, cooperation, and coordination (Andriessen, 2003; Churchill and Snowdon, 1998). The definition of a CVE as proposed by Snowdon et al. provides a general perspective without limitations in regard to applied technologies or visual representations that is still relevant today:

> "*A CVE is a computer-based, distributed, virtual space or set of places. In such places, people can meet and interact with others, with agents or with virtual objects. CVEs might vary in their representational richness from 3D graphical spaces, 2.5D and 2D environments, to text-based environments. Access to CVEs is by no means limited to desktop devices, but might well include mobile or wearable devices, public kiosks, etc.*"
>
> – Snowdon, Churchill, and Munro (2001)

In order to allow for appropriate collaborative work through interaction and information sharing, Churchill and Snowdon (1998) highlight some key characteristics that CVEs should strive to provide. For instance, CVEs should allow their users to *transition* seamlessly *between their individual efforts and shared activities* that involve one, or potentially multiple, other users (Churchill and Snowdon, 1998). As such, in order to successfully collaborate, means of *communication and negotiation* are required to allow the users to discuss their findings during phases of shared effort and activity (Churchill and Snowdon, 1998). Furthermore, to enable users to effectively transition between the various different contexts during their stay in a CVE, it is essential to provide means that allow them to be *aware of each other*, and thus understanding the current state of their collaborators (Churchill and Snowdon, 1998). Therefore, a CVE should also provide features that facilitate and support both focused and unfocused collaborative work, allowing for the subsequent establishment of *shared contexts* and understandings (Churchill and Snowdon, 1998). Work on a specific subject commonly also involves taking on *different viewpoints*, allowing the collaborators to inspect a matter from different perspectives to obtain insights – another feature that CVEs should strive to provide according to Churchill and Snowdon (1998).

The description and subsequent evaluation of collaboration and its many involved processes, facets, and dimensions, is a very complex endeavor in CSCW in general, and thus also within the context of CVEs (Snowdon et al., 2001). The distinct dissection of collaborative work is particularly challenging as its many states and actions are inherently interconnected and dependent on each other (Rupprecht et al., 2017; Neale et al., 2004). Consequently, it is foundational to have a sound understanding of relevant terminology in order to implement and investigate collaboration related matters, not least because some terms tend to be ambiguously applied throughout the literature (Schmidt, 2002).

Schmidt (2002) describes *awareness* as a user's ability to align and integrate their own actions with those of other collaborators in the shared workspace, ideally in a rather seamless and organic manner without interruptions and major efforts. Schmidt (2002) also highlights the many different aspects and use cases in which awareness as terminology has been applied over the years, calling for careful consideration and clear definition when applied in real-world collaborative scenarios. For instance, Heer and Agrawala (2008) expand on Schmidt's (2002) description, stating that awareness is also concerned with the user's ability to assess work and task completion, enabling them to decide where to allocate their next efforts. A closely related concept to awareness is *common ground*, referring to the collaborators' shared understanding of their states in the environment as well as the state of their work, allowing for subsequent communication and negotiation (Clark and Brennan, 1991). The process of achieving a state of common ground is typically referred to as *grounding* (Heer and Agrawala, 2008). Andriessen (2003) examined group processes within collaborative scenarios with respect to the various potential interactions between the collaborators, providing a heuristic differentiation between *communication*, *co-operation*, and *co-ordination*. As such, communication can be seen as an interpersonal exchange process between the collaborators, allowing them to utilize respective tools for the exchange of verbal (such as talk) and nonverbal (such as referential pointing; see also Figure 2.9, describing hand gesture classifications in general) signals (Andriessen, 2003). As part of such interpersonal exchange processes, *referencing* is arguably a comparatively frequent communicative action, typically applied in order to indicate a specific object, area, person, or time, through a mixture of verbal and nonverbal signals (Heer and Agrawala, 2008). Heer and Agrawala (2008) highlight and elaborate on the different approaches one can follow when making a references in space. In particular, references may be (1) *general*, e.g., "*six o-clock*", "*southeast by east*", (2) *definite*, e.g., using the specific name of the entity referred to, (3) *detailed*, e.g., describing the attributes of the entity referred to, or (4) *deictic*, e.g., nonverbal pointing to the referred entity accompanied by a complementary verbal

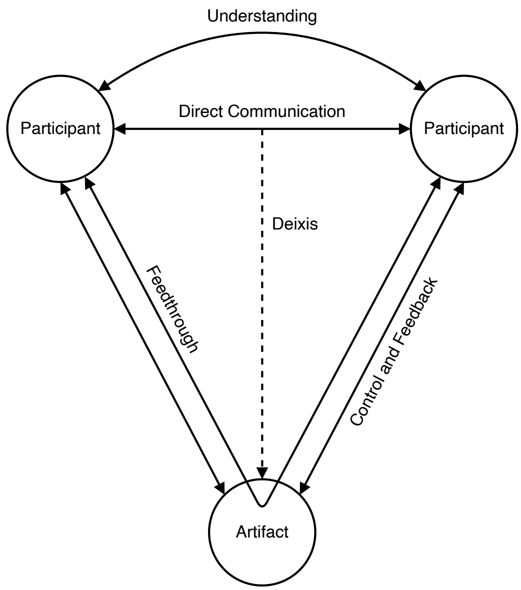


Figure 2.11: CSCW Framework, adapted from Dix (1994).

expression such as "*that one*", "*over there*", and so forth (Heer and Agrawala, 2008). Following, co-operation refers to the task-oriented process of the collaborators actually working together in a joint group effort, making decisions, and potentially co-manipulating data and objects in the shared environment (Andriessen, 2003). The task-oriented process that enables collaborators to adjust their individual and group efforts in order to solve a given task, thus adjusting the overall work of the involved collaborators, can be described as co-ordination (Andriessen, 2003). Within collaborative scenarios, the concepts of *attention* (Shneiderman et al., 2017; Kristoffersen and Ljungberg, 1999) and *focus* (Schmidt, 2002; Snowdon et al., 2001) are commonly rather seen ambiguously, typically referring to a user's cognitive alignment to a specific area or point of interest in the shared workspace.[9]

Several frameworks exist that aim to facilitate the understanding of collaborative systems, allowing for more formal descriptions and classifications. For instance, Dix (1994) introduces a general CSCW framework that focuses on aspects such as *cooperative work*, *communication*, and the *artifacts of work*. As illustrated in Figure 2.11, the framework depicts two *participants* in collaboration, i.e., a joint group effort, as well as the involved work *artifact*. Both participants are able to *directly communicate* with each other, potentially applying *deictic* terminology to

[9]In regard to attention resources, see also the general description of human information processing as illustrated in Figure 2.2 in Section 2.1.2.

refer to artifacts in the shared environment. Typically, both participants are able to *control* the work artifact, for instance by directly interacting with it. Naturally, such interactions generate *feedback* accordingly. Furthermore, while a participant controls an artifact and receives feedback, the other participant is able to passively observe their partner's interactions and thus to receive similar feedback, referred to as *feedthrough*. Finally, all the interactions between the participants themselves and with the artifacts of their work aim to establish a mutual *understanding* between the participants, enabling them to successfully collaborate.

Thematically aligned with Dix's (1994) framework, collaboration coupling styles, i.e., modes describing how participants interact with each other as well as with work artifacts, have been empirically investigated by, among others, Isenberg et al. (2012) and Tang et al. (2006). Although originating from multi-user collaboration around tabletop displays, their descriptions and classifications of the various collaboration states, that the participants may find themselves in during their individual and group work, provide relevant reflections and important considerations for the design of collaborative systems in general. For instance, Isenberg et al. (2012) provide detailed descriptions for various collaboration styles, classified as *close* and *loose collaboration* depending on the participants' activity. Gutwin and Greenberg (2002) examine the complex subject matter of *workspace awareness* more closely, providing a framework to describe a participant's understanding of their collaborator's interactions in the shared workspace. Their framework aims to facilitate the design of CSCW systems by establishing and providing mechanisms for the support of workspace awareness through the analysis of relevant components, such as *environment*, *knowledge*, *exploration*, and *action*, as well as the *interplay* between these (Gutwin and Greenberg, 2002).

In practice, collaboration using computer systems may occur in various different contexts and scenarios. A popular approach to determine these more descriptively is the classification of the collaboration technologies according to time and space. Such a Time/Space Matrix of CSCW systems has been pioneered by the works of Johansen et al. (1988), Baecker et al. (1995), and Dix et al. (2004, Chapter 19.2). As illustrated in Figure 2.12, the matrix generally differentiates between *spatial* (co-located or remote) and *temporal* (synchronous or asynchronous) dimensions in which collaboration can take place. However, as information and communication technologies advance and provide new mechanisms, modalities, and contexts, in which computer interfaces are available, a simple classification with respect to time and space becomes increasingly insufficient in order to capture the variety of potential scenarios and use cases (Ouverson et al., 2021; Neumayr et al., 2018). Lee and Paine (2015) discuss this matter and propose a Model of Coordinated Action that is built around seven dimensions (*synchronicity*, *physical distribution*, *scale*, *number of communities of practice*, *nascence*, *planned permanence*, and *turnover*) that each utilize a continuum rather than binary categorization. Through this less rigid classification of collaborative scenarios and technologies,

	One Meeting Site [Same Space]	**Multiple Meeting Sites** [Different Spaces]
Synchronous Communication [Same Time]	Co-Located / Face to Face Interactions	Remote / Distributed Interactions
Asynchronous Communication [Different Times]	Continuous / Ongoing Tasks	Communication and Coordination

Figure 2.12: Time/Space Matrix of CSCW systems, adapted from Dix et al. (2004, Chapter 19.2), Baecker et al. (1995), and Johansen et al. (1988).

their framework provides practitioners with a wider vocabulary to describe their CSCW systems accordingly (Lee and Paine, 2015). Neumayr et al. (2018) expand on the original concept of the Time/Space Matrix through the definition of Hybrid Collaboration, stating that in modern collaboration setups (1) the temporal and spatial collaboration dimensions are not exclusive to each other, but rather constant transitions between the individual quadrants occur, (2) more than two collaborators exist, including the dynamic allocation into subgroups, thus featuring a variety of different collaboration coupling styles at any given time, and (3) it is likely that more than just one tool or application is used, but multiple different ones, each serving their own purpose.

Ens et al. (2019) reviewed the application of mixed reality technologies for collaborative purposes throughout the past decades (1995 to 2018), and found that the amount of publications related to CSCW and mixed reality have increased significantly since 2012. Arguably, as immersive technologies become more accessible and generally easier to maintain, researchers and practitioners can focus on the application of their collaborative system in real-world contexts and its subsequent empirical evaluation (Ens et al., 2019). Based on the examined 110 publications, Ens et al. (2019) created a set of dimensions to aid with the categorization of these publications. Despite facilitating the identification of popular themes in the research community, potential gaps, and underrepresented topics that require further investigation, these dimensions can also assist practitioners to appropriately position and describe their collaborative systems. In particular, Ens et al. (2019) defined six dimensions to describe collaborative systems that utilize immersive technologies:

- *Time*: When is the collaboration happening, i.e., is it synchronous, asynchronous, or both?
- *Space*: Where is the collaboration happening, i.e., is it co-located, remote, or both?
- *Symmetry*: Do the collaborators have the same roles (symmetric), or has each one a specific role that differs from the other (asymmetric)?
- *Artificiality*: In alignment with the reality-virtuality continuum (see Figure 2.1 in Section 2.1.1), to what extent is the environment synthetic, i.e., is it mostly physical, mostly digital, or hybrid?
- *Focus*: Who is the primary target of the collaborative activity, i.e., is it an environment (surroundings of the collaborator), a workspace (region of interest during collaboration), a person (representation of other collaborators), or an object (real-world object or virtual replica in the VE)?
- *Scenario*: What is the overall concept and use case of the collaborative system, i.e., is it a remote expert, shared workspace, shared experience, telepresence, or co-annotation use case?

Following a similar approach of reviewing collaborative mixed reality research presented between 2013 and 2018, de Belen et al. (2019) additionally report on the *area of application* for collaborative systems. The results of their review indicate the many possible contexts in which collaboration through immersive technologies has been investigated, including prominent ones, such as entertainment, gaming, education, and training (de Belen et al., 2019). Besides these frequently investigated areas of application, there is also research conducted within the context of industrial applications, architecture and construction, medicine, as well as tourism and heritage (de Belen et al., 2019). It becomes apparent that the design, development, and evaluation of CVEs that utilize immersive technologies entails a multitude of considerations, and is thus an inherently complex subject matter due to the many different components involved – both conceptual as well as technological. Naturally, CVEs should aim to support the various foundational CSCW concepts as presented throughout this section in order to facilitate successful collaboration between the participating users (Snowdon et al., 2001). However, the often rather single-user centric nature of the involved display and interaction technologies (Skarbez et al., 2019; Cordeil et al., 2017b; Hackathorn and Margolis, 2016) can strongly contrast with valuable collaborative aspects, especially within the context of VR. For instance, important visual communication cues from a collaborator in a the real-world environment, such as facial expressions, body language, gestures, or spatial references, are no longer conventionally available.

Instead these need to be specifically designed and implemented in order to be supported in CVEs. Such nonverbal communication features are particularly important and relevant instruments within CVEs, but their appropriate design is by far not a trivial task, and highly dependent on the characteristics and capabilities of the chosen display and interaction technologies as well as the context, design, and composition of the VE itself (Cruz et al., 2015; Nguyen and Duval, 2014). One can further argue that due to the single-user centric characteristics of immersive technologies, i.e., each user viewing the VE through their own display and interacting in it with their own dedicated tools, CVEs are often rather remote (distributed) in nature, even in co-located spaces – especially when adapting a VR approach as described in Section 2.1. Essentially, there is not just a challenge with respect to system design and implementation, but also evaluation, which is particularly demanding in the case of systems that feature remote collaboration through complex synchronous and asynchronous interactions (Neale et al., 2004).

2.4 Immersive Analytics

Utilizing immersive display and interaction technologies for the visualization, exploration, and analysis of data in VEs has fascinated researchers at least since the 1990's (Fonnet and Prié, 2021; Dwyer et al., 2018), and is thus not a particularly new research direction. However, comparatively recent technological advances have elevated immersive interfaces into the mainstream, making them generally more affordable, widely accessible, and easier to maintain – both with respect to hardware and software technologies. New research interest has sparked for the exploration of approaches to analyze and interact with data in immersive spaces, utilizing the different characteristics and benefits immersive technologies can hold, such as described throughout Sections 2.1 and 2.2. Researchers have come together under the umbrella term of Immersive Analytics (IA) to establish this research domain more formally. Several definitions of IA have been recently proposed to outline the research domain's general purpose and approach (Skarbez et al., 2019; Dwyer et al., 2018; Hackathorn and Margolis, 2016; Chandler et al., 2015). While all of these are certainly applicable, the definition of IA by Skarbez et al. is adopted within the scope of this thesis:

> "*Immersive Analytics is the science of analytical reasoning facilitated by immersive human-computer interfaces. By analytical reasoning, we specifically refer to computer-aided analytical reasoning as a partner with the human; that is, a process of foraging and sensemaking where part or all of the foraging and / or sensemaking processes are performed in cooperation with a computer. By immersive human-computer interfaces, we specifically mean those interfaces which enable a user to interact with a system using additional or more-immersive displays and user interface techniques.*"
>
> – Skarbez, Polys, Oggle, North, and Bowman (2019)

Among others, and besides being coincidentally the most recently published one of the referred IA definitions, Skarbez et al. (2019) provide a more comprehensive definition compared to the compact ones by Dwyer et al. (2018) and Hackathorn and Margolis (2016), and at the same time appears less rigid compared to the earlier definition by Chandler et al. (2015). It furthermore becomes quickly apparent that IA is a highly interdisciplinary research field that integrates knowledge and methods from various disciplines and research communities, such as Information Visualization (InfoVis), Visual Analytics (VA), HCI, CSCW, augmented reality, mixed reality, VR, and 3D UIs, to name just a few. Particularly through the 3D characteristics of VEs as well as the availability of powerful depth cues in modern visual displays that can assist the user with their 3D spatial perception, as described in Section 2.2.1, it is time to reassess the value of data visualizations and interaction in 3D, as these have been commonly disregarded outside Scientific Visualization (Marriott et al., 2018). Naturally, this may concern utilizing the third dimension for additional visual mapping and encoding of data compared to 2D approaches (Marriott et al., 2018). Another possibility could be utilizing the immersive 3D space for the placement of 2D visualizations and other 2D artifacts, such as text and media, in contextually relevant locations, creating information-rich VEs (Skarbez et al., 2019). However, it is important to highlight that IA aims to provide novel, intuitive, engaging, and purposeful 3D data analysis tools that complement, synergize, and potentially even closely integrate with InfoVis and VA workflows rather than replacing them (Cavallo et al., 2019; Wang et al., 2019; Isenberg, 2014).

Dwyer et al. (2018) describe various opportunities for the design of appropriate IA applications, both from a general HCI as well as a data visualization perspective. For instance, it is possible to integrate immersive technologies closely with objects anywhere and at any time in the physical real-world environment, linking physical and virtual worlds, and allowing for the subsequent display of information and the interaction therewith within in-situ contexts (Dwyer et al., 2018). Due to the 3D spatial characteristics of many immersive interfaces, the analyst is no longer required to, for instance, sit in front of a computer terminal "outside", but instead surround themself with information and taking on instead a more "inside" perspective. Consequently, this allows for a more embodied and user-centered approach, potentially utilizing on a wide variety of different input technologies and interaction techniques, enabling the analyst to explore data in a more engaged manner by actively moving around (Dwyer et al., 2018; Büschel et al., 2018). Naturally, there is also the potential of utilizing the benefits of various interface types for the stimulation of the different human sensory organs, as described in Section 2.1.1. Besides the arguably most commonly applied visual interfaces, data can be mapped and displayed to alternative interfaces that enable multisensory experiences (Dwyer et al., 2018), including spatial 3D audio (Marai et al., 2016). Furthermore, closely aligned with the subject matter of CVEs, as described in Section 2.3, IA has the opportunity to facilitate collaboration between multiple analysts, independent of space and time constraints, potentially supporting a wide variety of collaboration modalities (Dwyer et al., 2018). Finally, there is also the chance to benefit from intuitive and engaging interaction and viewing approaches enabled through the utilization of immersive technologies, allowing non-expert and novice users to engage with IA experiences, facilitating their understanding about presented phenomena in a more narrative-driven approach, and thus informing their subsequent decision-making process (Dwyer et al., 2018; Isenberg et al., 2018).

The high level of interdisciplinarity as well as the various opportunities across many different applied scenarios and use cases pose a multitude of research challenges that require further investigation, as briefly described in Section 1.1. Just recently, in 2021, a collective of 24 researchers from different disciplines and backgrounds came together to formalize in a joint effort a set of 17 challenges in IA, providing an overall summary and agenda for important research directions in the near future (Ens et al., 2021). In particular, Ens et al. (2021) classify these challenges across four major topics as follows:

- *Spatially Situated Data Visualization*: This topic is concerned with the investigation of applying immersive technologies in situated contexts in real-world spaces, i.e., in-situ, in order to link and display relevant information interactively in an appropriate manner. As such, it also concerns the investigation of particular human factors issues as well as the ethical approaches for data visualization in such in-situ contexts.
- *Interacting with Immersive Analytics Systems*: This topic focuses on the examination of interaction modalities and techniques for the purpose of actively operating and engaging with IA systems, and coping with their overall complexity (Büschel et al., 2018). The utilization of novel immersive output and input devices, as described throughout Section 2.2, poses various challenges with respect to human factors and ergonomics in order to design and develop appropriate and effective interaction with data in (partially) virtual 3D spaces that incorporate multisensory user feedback. A lack of guidelines and best practices for the interaction in IA environments is also emphasized by Fonnet and Prié (2021).
- *Collaborative Analytics*: As data analysis and subsequent meaning and decision making is seldom a process that is done in isolation but in collaboration between multiple analysts, the challenges in this topic are concerned with the utilization of immersive technologies to enable collaborative experiences in various ways. As described throughout Section 2.3, there are many facets that have to be considered when designing collaborative systems that also need to be supported within IA scenarios and use cases. Of particular interest in this topic are aspects with regard to the support of human

behavior in collaborative environments, the support of collaboration across technologically different platforms, the integration into existing analysis workflows, as well as the evaluation of collaboration in immersive data analysis environments.

- *User Scenarios and Evaluation*: This topic is generally concerned with the identification of guidelines and recommendations as to which scenarios and use cases are suitable and applicable for IA, and which ones are maybe rather implemented using non-immersive technologies. There are many different aspects to consider for the deduction of such recommendations, for instance with respect to technology, the user, and application contexts. Furthermore, there is a also a need for general strategies and approaches that guide and facilitate the evaluation of any IA system, allowing the research community to draw conclusions with respect to user experience and performance based on empirical evidence.

In addition to the research topics presented by Ens et al. (2021), Skarbez et al. (2019) provide some impulses in regard to an overall IA research agenda. Their areas for IA research overlap for the most part with the ones discussed by Ens et al. (2021). However, two are in particular noteworthy to highlight. First, in line with the overall concept of allowing user interaction with computer systems in a closer and more integrated manner compared to traditional approaches (see Section 2.2), Skarbez et al. (2019) propose the investigation of approaches to combine human and machine intelligence to solve analytical tasks in IA environments. Parts of this investigation are therefore concerned with topics such as Big Data, Machine Learning, and Artificial Intelligence, and their subsequent integration in IA systems. And second, with respect to human behavior in immersive environments, Skarbez et al. (2019) issue the investigation of how the application and exposure to IA systems, in particular in regard to a provided level of immersion, may affect and improve analytical procedures and presentations.

2.4.1 Collaborative Immersive Analytics

As previously highlighted as one of the main challenges in the IA research agenda (see Section 2.4), the application of immersive display and interaction technologies to allow analysts to work together in collaborative scenarios, for joint data exploration and interpretation to collectively extract insights, is highly anticipated. To examine this subject, it is helpful to find a suitable definition as means for further guidance. Based on the general IA definition proposed by Dwyer et al. (2018), and incorporating the definition of Collaborative Visualization proposed by Isenberg et al. (2011), Billinghurst et al. derive and suggest the following definition for the term of Collaborative Immersive Analytics (CIA):

> *"The shared use of immersive interaction and display technologies by more than one person for supporting collaborative analytical reasoning and decision making."*
>
> – Billinghurst, Cordeil, Bezerianos, and Margolis (2018)

While one can argue that this definition of CIA is straightforward and to the point, it is at the same time arguably also rather generic, implicitly highlighting the vast amount of opportunities as well as the potential complexity of CIA. After all, the design of collaborative systems that utilize the benefits of immersive technologies for analytical purposes require many different considerations that have to be carefully attended. Besides the actual data context and scenario, these include also the choice and availability of technology, the knowledge and subsequent roles of the involved analysts, and general aspects of the collaborators as a group, such as their locations, common workplace, and so on. As designed IA systems can potentially utilize any kind of immersive technology and approach, possibly situating themselves anywhere on the reality-virtuality continuum (see Figure 2.1 in Section 2.1.1), so can in turn CIA systems. The before mentioned vast amount of potential configurations, scenarios, and use cases, becomes apparent.

Cernea (2015, Chapter 1) discusses the concept of User-Centered Collaborative Visualization, incorporating aspects of user-centered design in the proposed definition of collaborative visualization by Isenberg et al. (2011), similar to the definition approach by Billinghurst et al. (2018). Aligned with the strong focus on allowing users to closely immersive themselves with a computer system through the use of respective display and interaction technologies, as described throughout Section 2.2, it is worth having a closer look at Cernea's (2015, Chapter 1) concept. As such, the definition of user-centered collaborative visualization has been proposed as follows:

> *"User-Centered Collaborative Visualization is the shared use of computer-supported, interactive, visual representations of data that considers knowledge about the abilities and needs of both the involved users and the group as such, their task(s), and the environment(s) within which they work, in order to capture vital contextual information and support the process of completing the user group's common goal of contribution to joint information processing activities."*
>
> – Cernea (2015, Chapter 1)

Besides highlighting aspects of social processes and the interpersonal interaction as a consequence thereof during the collaborative analysis activity, in line with the discussions of general design considerations for collaborative systems by Heer and Agrawala (2008), Cernea's (2015, Chapter 1) definition hints at two other aspects that are arguably of particular importance within the context of CIA, namely the occurrence of potentially multiple different tasks as well as multiple different environments. Ens et al. (2019) highlighted symmetry as a defining characteristic for collaboration using immersive technologies, i.e., users having the same roles (symmetric) or different ones (asymmetric). Symmetry aspects arguably address and affect implicitly the task distribution between the collaborators, as different user roles are likely to infer different responsibilities and tasks during the joint data analysis. As a logical follow-up question, one may ask *what* tools, applications, or systems are the collaborators using, and *where* with respect to each other. For instance, are they using the same tools in the same environment, different tools in different environments, or any permutation thereof. Naturally, collaboration may occur across different times and spaces, as presented together with the many different aspects of collaboration within the context of CVEs as well as CSCW in general throughout Section 2.3. Examining the literature in regard to CIA, it becomes apparent that the term *environment* requires some further reflections. In particular, it subjectively appears that the terms environment and *space* are often applied interchangeably when discussing collaboration contexts, commonly referring to the physical real-world environment or space a collaborator finds themself in to partake and conduct their work. Within the context of immersive technologies however, common terms are VE, CVE, or sometimes simply *immersive environment* (see Sections 2.1.1 and 2.3). As such, within the context of CIA, one could argue that a different *environment* actually refers to a different type of immersive application or system, independent of the collaborators' physical real-world location. In other words, collaborators could be either co-located or remote, and find themselves in the same or different VEs, using the same or different display and interaction technologies, depending on their tasks and role distribution (Rogers et al., 2021; Lee et al., 2020; Grandi et al., 2019; Clergeaud et al., 2017). Thus, while certainly valid within the context of Cernea's (2015, Chapter 1) definition of user-centered collaborative visualization, the term environment should be carefully used within the context of describing CIA concepts and systems, potentially under consideration of additional clarifications.

The definition of hybrid collaboration by Neumayr et al. (2018), as described in Section 2.3, incorporates next to time and physical space dimensions also aspects of cross-device and cross-application collaboration. The description and specification of the involved device types that compose a CIA system seem therefore useful and worth considering. Fröhler et al. (2022) investigated the use of heterogeneous interface types for data analysis purposes as Cross-Virtuality Analytics, aiming to provide tools that support collaboration through the seamless integration in the data analysis activity independent of their platform. Hybrid Virtual Environments (Wang et al., 2019) and Collaborative Hybrid Analytics (Cavallo et al., 2019) are similar concepts that incorporate both 2D and 3D visualization environments. Particularly within the "bigger picture" of any analytical workflow, and also in regard to complex datasets, individual analysis tools that utilize different display and interaction technologies are nowadays rarely used in isolation, but instead as composition of multiple different ones, each serving their own purpose (Wang et al., 2019). Such a multimodal composition of tools, each providing their own dedicated advantages and perspectives, has the potential to greatly facilitate data exploration and analysis (Isenberg, 2014). Neumayr et al. (2018) already highlighted some considerations for hybrid collaboration across different devices in general, such as with respect to *working styles*, *user territoriality*, and *awareness*. Hybrid collaboration environments that utilize immersive and non-immersive technologies require however some further ones, especially in regard to *complementarity*, *transition*, *interaction*, and *collaboration*, when using different data analysis modalities (Cavallo et al., 2019; Wang et al., 2019; Isenberg, 2014). Considering such anticipated workflows that incorporate an interplay between tools that are built around heterogeneous display and interaction technologies, both immersive and non-immersive, the specification of device and application type within the scope of an own dedicated framework dimension, to facilitate the classification and description of CIA systems, seems logical henceforth. Among others due to its highly interdisciplinary nature as well as the complexity of CIA systems in general, Fröhler et al. (2022), Fonnet and Prié (2021), and Billinghurst et al. (2018) all emphasize that more research and empirical evaluations are needed in this domain.

2.5 Evaluation of Immersive Technologies

Examining the contents throughout all the sections in this chapter so far, the close coupling between human user and technology becomes apparent. Hence, HCI research is inherently empirical by default, aiming to systematically investigate phenomena and matters of interest as a result of using technological artifacts, and with a focus on the human user (Boyd and Bogen, 2021; Hansson, 2021), i.e., human factors and ergonomics as described in Section 2.1.2. Evaluation is a fundamental part of the iterative design process for technology (Franssen et al., 2018), and enables the assessment of a developed technological artifact with respect to the initially determined goals and requirements. Evaluation purposes may include (1) the identification of design problems with respect to the developed user interface (UI) or user experience in general, (2) the deduction of design guidelines by obtaining a deeper understanding of the technological artifact's application in practice, or even (3) the development of performance models with the aim to predict anticipated user performance when using the technological artifact (LaViola, Jr. et al., 2017, Chapter 11.1.1). HCI research has a long history of user evaluation (Barkhuus and Rode, 2007), and as such a rich corpus of diverse evaluation methods and metrics, both standardized as well as custom-made.

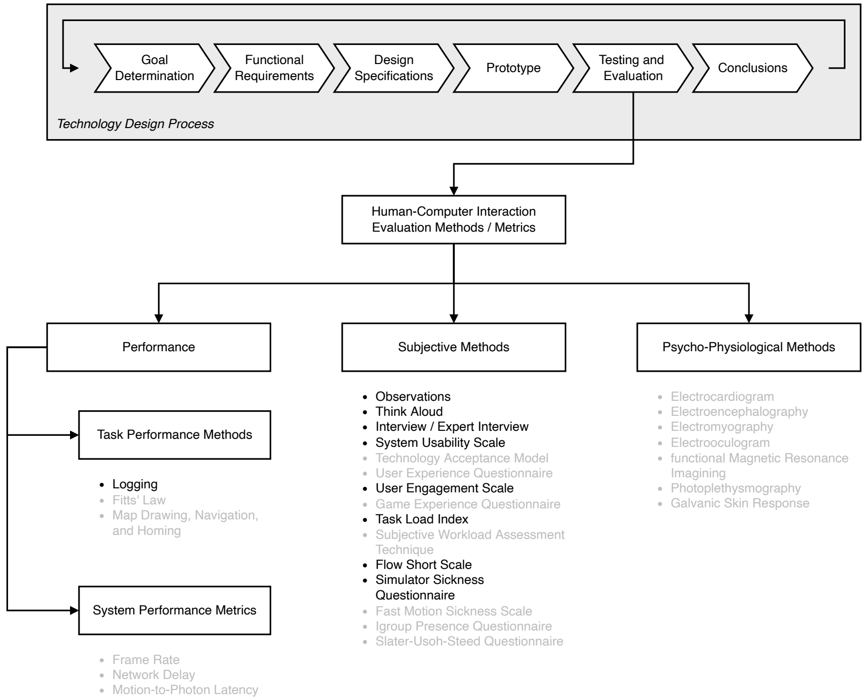


Figure 2.13: Technology Design Process, adapted from Franssen et al. (2018), and overview of HCI evaluation methods and metrics. The selection of evaluation methods and metrics is not meant to be exhaustive, but rather provide a general overview of widely recognized options that are deemed relevant within the context of this thesis. **Note**: Highlighted methods have been used for the empirical evaluation of developed artifacts within the scope of this thesis.

Figure 2.13 illustrates the evaluation stage as part of the Technology Design Process (Franssen et al., 2018), and provides an overview of evaluation methods and metrics that can be relevant for the research presented in this thesis, i.e., within the context of evaluating developed artifacts that utilize immersive technologies. Generally, evaluation methods and metrics can be divided according to *performance* related methods and metrics, in particular *task* and *system performance*, as well as *subjective methods* and *psycho-physiological methods*.

Assessments in regard to performance are generally *objective*, i.e., they are independent of the user's personal feelings and opinions. Aligned with the H and the C in HCI, performance can be assessed both with respect to the human user (task performance) and the computer system (system performance). Task performance related measurements are commonly presented as metrics such as speed and accuracy. LaViola, Jr. et al. (2017, Chapter 11.3.2) emphasize on the implicit relationship between these two, i.e., the faster one performs an action, the less accurate that actions tends to be, and vice versa. Other common task performance metrics include the amount of errors a user made in order to successfully complete specific tasks, or the amount of specific interactions with a UI. Although the human user is commonly in the focus of any HCI evaluation, it is sometimes important to measure performance metrics of the tested computer system. A developed artifact that performs poorly may impact how the user experiences the interaction with it, and subsequently influence the results of the human-centered metrics. For instance, human perception of motion is closely coupled to a visual display's frame rate, i.e., its ability to update the displayed graphical contents (see Section 2.2.1). Based on the intended use, certain frame rate thresholds should be met in order to ensure smooth visual perception of the displayed media, for instance, 90 frames per second for modern VR HMDs (Wallergård et al., 2022, Chapter 5.6.7; LaValle, 2020, Chapter 6.2). If a computer system is not powerful enough to run the developed artifact, thus not meeting such recommended thresholds, the result may be a less smooth and "jittery" visual perception through the user that potentially negatively impacts their ability to perform tasks as well as their subjective assessment of artifact. As such, system performance metrics should be generally regarded as complementary to the data collected from human users within the context of HCI evaluations, serving as a validation to confirm the appropriate functioning of a technological artifact.

Data collected through *subjective* methods are influenced by the personal feelings and opinions of the human that reports the data. That can be a user that self-reports their experience with a prototype on the one hand, or on the other a researcher themself through taking notes and making observations. In practice, data collected through subjective methods can be both *quantitative* and *qualitative*. Quantitative data can be described as numerical values, often within the context of a continuum or otherwise interpretable scale, and are as such commonly *close-ended*. One of the arguably most famous scales has been introduced by Likert (1932), namely the Likert scale, a bipolar scale comprised of multiple rating options, typically 5, 7, or 9.[10] Accompanied by a statement, the reporting user can select the rating option on the scale that reflects most their impression, allowing for comparison among multiple different users. Collections of such Likert scale items are often compiled into questionnaires, which can be custom-made by the researcher to specifically investigate those aspects they want to evaluate, or provided by the community in the format of standardized benchmarks for a more generalized investigation of a subject (Sauro and Lewis, 2012). Qualitative data are commonly *open-ended*, such as collected through interviews, observations, or

[10] An odd number of rating options allows for a neural rating through the median item. An even number of rating options may be chosen when a decisive tendency towards one of the two endpoints is desired.

note taking, relying on descriptive expressions of the reporting user or researcher. As such, qualitative data collection allows for a rather free-form capture of experienced or observed phenomena. Based on the evaluation objective, a *mixed methods* approach may be applied, incorporating both quantitative and qualitative subjective data collections methods.

Pyscho-physiological methods are used to investigate the relationship between the human user's cognition and physical ergonomics, i.e., how a user's body physically reacts based on the prior perceived stimuli as processed through cognition (Baig and Kavakli, 2019; Park, 2009). Among others, measurements such as heart rate, body temperature, muscle activity, brain signals, skin conductance, eye movements, and facial expressions, can provide data that allow for an analysis with respect to emotion and affect recognition, or towards a general assessment of a user's cognitive state (Baig and Kavakli, 2019). Within the context of HCI research, pyscho-physiological methods are particularly relevant in regard to the assessment of *stress*, *attention*, and *emotion* (Park, 2009). Schmälzle and Grall (2020) describe the application of pyscho-physiological methods within the context of media theory and research, providing an overview of relevant data collection methods and discussing their validity. Common pyscho-physiological data collection methods include electrocardiogram (ECG), electroencephalography (EEG), electromyography (EMG), electrooculogram (EOG), functional magnetic resonance imagining (fMRI), photoplethysmography (PPG), and galvanic skin response (GSR) (Schmälzle and Grall, 2020; Baig and Kavakli, 2019; Park, 2009). Park (2009) highlights the complexity of modern HCI systems, and emphasizes that results from the application of pyscho-physiological methods should always be interpreted specifically within their applied context. Furthermore, Baig and Kavakli (2019) found as a result of their review that psycho-physiological measurements commonly show strong correlations with a user's self-reported data.

2.5.1 Applied Evaluation Methods

Section 2.5 and Figure 2.13 have provided a general purpose overview of evaluation methods and metrics as well as their classifications within the context of HCI. The remainder of this section is dedicated towards the description of those evaluation methods that have been applied to empirically evaluate developed artifacts within the scope of this thesis.

Logging The utilization of a logging system that is directly integrated as part of a developed artifact allows for a comprehensive data collection of the interactions between the human user and the computer system. Rather than observing and manually counting how often a user makes use of a specific feature in a UI, that kind of data can be automatically collected through the computer system itself. Naturally, it is the researcher's responsibility to integrate such mechanisms in a systematic and structured manner, ideally following a predefined protocol, aiming to keep track of all interactions that are relevant for their analysis. Collecting data in a human readable format, such as comma-separated values (CSV) or JavaScript Object Notation (JSON), can facilitate processing and analysis using desired tools of the researcher's choice. Furthermore, each entry collected with a logging system should feature a timestamp as a general reference, allowing for the analysis of user interaction over time. Finally, an integrated logging system as part of a developed artifact will consume computer system resources. Consequently, any logging system should be implemented non-disruptively, i.e., not affecting the performance of the developed artifact in a for the human user noticeable manner.

Observations Conducting observations, and subsequent note taking, is arguably among the most basic, but nevertheless effective, subjective data collection methods (Goodman et al., 2012, Chapter 9). Observations allow researchers to naturally follow a user's interactions with a technological artifact in a way that is not disruptive to their experience. They can be free-form with respect to the researcher's ability to capture observed phenomena as they naturally occur – after all, unexpected things may happen at any point in time. As part of designing an evaluation, it can be beneficial to think critically about certain phenomena that the researcher anticipates to observe, compiling a complementary list that aids their observation and note taking process. This can allow the researcher to more easily detect re-occurring patterns and phenomena across multiple user study sessions, such as usability issues or the use of features for unintended purposes. If the researcher is taking notes of observations in real-time as they occur, it is important to highlight that this process may temporarily distract the researcher. As they focus on their note taking, the next noteworthy event may happen that could thus be missed. The comparatively free-form nature of observations commonly require further preparations for the results presentation and analysis, such as bundling and categorizing observations into themes or topics, enabling an appropriate presentation of re-occurring matters.

Think Aloud The Think Aloud technique encourages the human user to actively comment on their interactions with a developed artifact in-situ by speaking aloud (Fonteyn et al., 1993). This allows the researcher to get a better idea of the thought processes and the intentions of the user during their interactions. The technique is comparatively easy to learn and to apply, but requires some considerations. For instance, some users may be more talkative than others. Some users may also feel uncomfortable expressing negative feedback, potentially omitting such comments, and rather focusing on positive aspects of the evaluated artifact. Furthermore, the technique can be somewhat disruptive to the overall flow of interaction, as users are tasked to verbally express themselves and to describe what they are doing, which can also feel unnatural. As such, applying the think aloud technique as part of an overall evaluation design should be carefully planned beforehand. For instance, combining performance measurements where users are tasked to complete interactions as fast and accurate as possible, in combination with the think aloud technique where they are encouraged to actively comment on what they are doing, is inherently conflicting. However, in other scenarios where task performance measurements are disregarded, the think aloud technique may be very valuable to easily capture user feedback, especially in iterative design evaluations. Similar to the presentation of observation results, user feedback captured using this technique should also be bundled and categorized into re-occurring themes and topics.

Interview/Expert Interview Interviews are a common approach to inquire the feelings and opinions of the human user, typically after they had the opportunity to engage and interact with a developed artifact (Goodman et al., 2012, Chapter 6). Interviews can be conducted in a *structured* manner, where all questions are prepared by the researcher in advance, and all users are asked the same questions according to a prepared protocol. There is also the possibility to conduct informal interviews that are rather *unstructured*, allowing the researcher to improvise and come up with questions in-situ, for instance based on observations made prior to the interview or as a follow-up to previous answers by the interviewee. *Semi-structured* interviews combine the best of structured and unstructured approaches, allowing the researcher to prepare questions every user is asked beforehand, while maintaining the freedom to compose follow-up questions during the interview to inquire more details if necessary. Interviews are commonly conducted with representative users, i.e., those expected to be within a target group of the evaluated artifact. However, another approach to inquire insights is to interview relevant experts in order to build upon their domain knowledge (Bogner et al., 2010). Interviewed experts may also come from different domains. For instance, an interactive climate visualization tool could be evaluated with a climate data domain expert to inquire insights in the usefulness and appropriateness with respect to the overall data context, while a HCI domain expert may provide constructive feedback on the implemented interaction techniques. It can also be beneficial to conduct interviews with multiple experts of the same background at the same time, potentially enabling them to organically discuss among themselves as they comment on each other's thoughts.

System Usability Scale The System Usability Scale (SUS) presented by Brooke (1996) is a standardized questionnaire aiming to evaluate the usability of a developed artifact, i.e., with respect to how easy and how pleasant the interaction with the artifact is.[11] It consists of ten 5-point Likert scale items that are generalized, i.e., they are formulated to allow application to any kind of evaluated artifact without the need for additional changes. The answers can be easily calculated into an interpretable usability score between 0 (bad) and 100 (good).

[11] Jakob Nielsen. Usability 101: Introduction to Usability. Retrieved June 1, 2022, from `https://www.nngroup.com/articles/usability-101-introduction-to-usability/`

The SUS has been widely used and is well established within the HCI research community (Brooke, 2013). Tasking users to complete the SUS after they interacted with the developed artifact is comparatively straightforward. Its ten items are easy to understand, making the questionnaire not too lengthy nor too demanding. In addition to the original numerical score, Bangor et al. (2009) present adjective ratings to further facilitate the SUS score interpretation. In particular, Bangor et al. (2009) propose the following ratings and their corresponding SUS score thresholds: worst imaginable (25), poor (39), ok (52), good (73), excellent (85), best imaginable (100). Another complementary score interpretation approach is presented by Sauro and Lewis (2012), proposing the transfer of the numerical score into traditional school grades (A+ to F).

User Engagement Scale O'Brien et al. (2018) describe user engagement as part of an overall user experience, stating that it represents the depth of a user's investment when interacting with a developed artifact. Thus, O'Brien et al. (2018) present the User Engagement Scale (UES) as a standardized tool to measure engagement for the purpose of utilizing the results within an artifact design and evaluation process. The 2018-revision of the UES evaluates user engagement across four dimensions, i.e., *focused attention*, *perceived usability*, *aesthetic appeal*, and *reward*. There are an extensive Long Form (LF) as well as a more condensed Short Form (SF) of the UES questionnaire available (O'Brien et al., 2018). The UES-SF features twelve 5-point Likert scale items, while the UES-LF consists of a total of 30 items. The answers can be analyzed on a scale from 1 (bad) to 5 (good) for each of the four dimensions, and as a combined *overall user engagement* score.

Task Load Index Hart and Staveland (1988) present a method to allow for a self-reported estimation of workload when operating an interactive system, namely the Task Load Index (TLX).[12] The TLX is a two-step approach that first tasks users to weigh and then rate six different workload related factors, i.e., *mental demand*, *physical demand*, *temporal demand*, *effort*, *frustration*, and the user's *own performance*. A calculation of a final score (weighted rating) is possible, representing the user's perceived *workload* on a scale from 0 (extremely low workload) to 100 (extremely high workload). Hart and Staveland (1988) argue that the overall concept of workload is rather versatile, likely resulting in multiple users having a different understanding and interpretation of what workload means to them. The *weighing process*, a comparison of 15 possible factor pairs, is intended to address this, essentially assigning weights to the individual workload contributing factors according to a user's understanding. The subsequent *rating process* involves a simple rating of all the six factors on a scale from 0 to 100, typically in steps of five. Based on the determined weights and ratings, scores for the individual factors (*adjusted ratings*) as well as a total workload score (*weighted*

[12]Hart and Staveland (1988) developed the TLX as part of their work at the National Aeronautics and Space Administration (NASA). Therefore, the TLX is alternatively referred to as NASA TLX.

rating) can be calculated accordingly. Applying the TLX in practice post-study is comparatively convenient, although users should be briefed with respect to the six different factors. The TLX has been widely used and thus become an established method to measure workload (Hart, 2006). Alternatively to the original two-step approach, the TLX is sometimes applied without the weighing process as Raw Task Load Index (RTLX), inquiring only user ratings for the six different factors to streamline the data collection process (Hart, 2006).

Flow Short Scale Based on the overall flow theory as discussed by Csikszentmihalyi and Csikszentmihalyi (1992) as well as Rheinberg (2010), it is possible to subjectively measure and investigate the perceived state of *interaction flow* a user can experience when interacting with a developed artifact. For this purpose, Rheinberg et al. (2003) developed the Flow Short Scale (FKS; German original: Flow-Kurzskala), adopted and shortened based on the work by Csikszentmihalyi and Csikszentmihalyi (1992). The FKS consists of 13 7-point Likert scale items, ten of which are concerned with the *flow experience* in general and can thus be analyzed with respect to the *smooth automatized process* and the *ability to absorb*. The last three of these 13 items make inquires regarding potential *concern* a user may experience. Furthermore, three 9-point Likert scale items may be included at the end of the FKS that are concerned with some additional self-assessments of the user in reference to their interaction with a developed artifact or activity.

Simulator Sickness Questionnaire The Simulator Sickness Questionnaire (SSQ) was originally introduced by Kennedy et al. (1993) within the context of aviation psychology with the aim to investigate symptoms a user may experience due to the exposure to a simulator-like environment. The questionnaire consists of 16 4-point Likert scale items, each inquiring user assessments with respect to different symptoms across three dimensions, namely *nausea*, *oculomotor*, and *disorientation*. The answers can be analyzed with respect to each individual symptom, each dimension, and as a total *simulator sickness* score. The SSQ gained popularity over the years and has been applied across different types of simulator-like experiences, including VR and HMDs (Hirzle et al., 2021; Rebenitsch and Owen, 2016). For the purpose of evaluating VR experiences, Bouchard et al. (2007) propose an alternative score calculation of the original 16 SSQ items across only two dimensions (nausea and oculomotor) as *cybersickness*. LaValle (2020, Chapter 12.3) discusses some practical implications of utilizing the SSQ for VR research that are in line with the reflections and suggestions by Bimberg et al. (2020), including the subjective results based on the user's self-reported symptom assessment and its rather disruptive nature, as it should normally be applied multiple times during a user's exposure with a developed artifact. Hirzle et al. (2021) also highlight the need for improvements with respect to facilitating the capture of factors that are more closely related to VR experiences that utilize HMD devices. VR sickness as described by LaValle (2020, Chapter 12.3) has been discussed in Section 2.1.2.

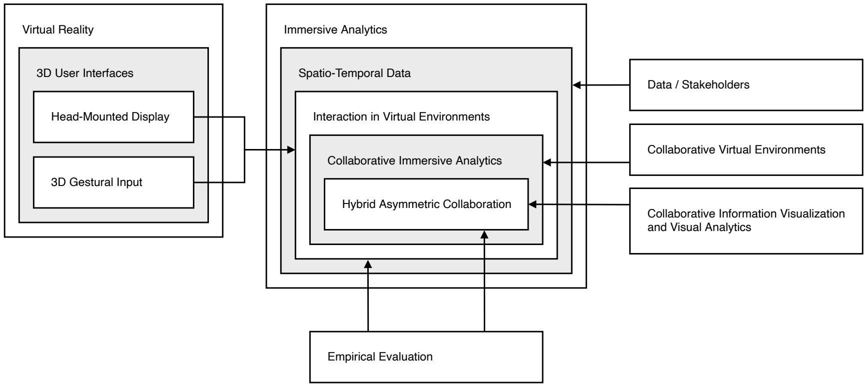


Figure 2.14: The design space of this thesis, aligned with the presented research problem, scope, goal, and objectives (see Section 1.2), and the various foundational subjects presented throughout Chapter 2, aiming to illustrate the interdisciplinarity of the presented research in a modular manner.

2.6 Thesis Design Space

Under consideration of the presented research problem, scope, goal, and objectives, as described in Section 1.2, as well as the obtained foundational understanding of the various research areas and relevant subjects throughout this chapter, the overall design space for this thesis can be proposed. Consequently, Figure 2.14 illustrates the thesis design space, constructed by carefully aligning all relevant topics and their relationships in a modular manner.

IA is positioned at the center of the design space as a generally overarching theme. More specifically, the represented research is concerned with the investigation of spatio-temporal data that is provided through some external data source or stakeholder, depending on the overall context and scenario. Furthermore, the first core theme of this thesis is concerned with the interaction in VEs for data analysis purposes. VEs can be implemented using different conceptual approaches and technological interfaces. The research presented in this thesis focuses on the application of a VR approach that utilizes HMD and 3D gestural input devices as types of 3D UIs. The second core theme of this thesis is concerned with enabling multiple users to collaboratively analyze data in VEs, i.e., CIA. Naturally, this theme is supported through insights from the area of CVEs. A central aspect of the collaborative analysis setup is concerned with the application of immersive and non-immersive display and interaction technologies as well as dedicated roles for each user, aiming to bridge CIA with collaborative InfoVis and VA. As such, the concept of Hybrid Asymmetric Collaboration is introduced and positioned within the context of CIA. Finally, the research presented in this thesis is empirically evaluated, both in regard to the interaction in VEs and hybrid asymmetric collaboration.

The modular representation of the thesis design space is chosen on purpose to reflect on its interdisciplinary nature. It is conceivable that similar other investigations may be conducted in the future. Instead of utilizing a VR approach and the presented 3D UIs, one could investigate an overall similar subject, such as the interaction with spatio-temporal data within the context of IA, with other conceptual approaches and technologies. Consequently, the respective modules on the left side of the illustrated design space could be replaced, for instance through an augmented reality approach and portable tablet devices as 3D UIs, to name just one example.

Chapter 3

Related Work

Contents

After obtaining a foundational understanding of important concepts, terminology, and research areas that are relevant within the scope of this thesis, as described throughout Chapter 2, it is possible to examine related work of other researchers to extract helpful insights and impulses. Researchers have presented interesting work that is informative for the design and development of interactive and collaborative systems, including aspects of immersive data visualization in general, in anticipation of investigating the presented research goal and objectives (see Section 1.2). It is noteworthy that the related work presented throughout this chapter is selected in accordance to their relevance within the defined design space of this thesis (see Section 2.6).

In particular, Sections 3.1 and 3.2 are centered around the exploration of Immersive Analytics (IA) systems that incorporate aspects of Virtual Reality (VR) through head-mounted display (HMD) devices and three-dimensional (3D) gestural input. The insights and impulses obtained from the presented work in these sections are relevant, among others, for the visualization and interaction design as well as aspects of general Virtual Environment (VE) composition across the three major VE iterations described throughout Chapter 5. Sections 3.3 and 3.4 are concerned with the investigation of related work that focuses on collaborative aspects, particularly relevant within the contexts of Collaborative Immersive Analytics (CIA) and Collaborative Virtual Environments (CVEs) that involve at least one user being immersed in a VE through a VR approach. The presented work in these two sections provides guidance for the design and development of collaborative data analysis experiences as described throughout Chapter 6.

3.1 Immersive Data Visualization Using Virtual Reality

Researchers have conducted some interesting work in regard to the immersive visualization of data using VR (see Section 2.1) over the years (Kraus et al., 2022; Fonnet and Prié, 2021). Facilitated through the recent advancements in regard to immersive display and interaction technologies (see Section 2.2), more and more researchers reassess the utilization of 3D virtual spaces as interactive data analysis tools (see Section 2.4). The insights of the existing empirical work in the domain of IA can provide useful considerations and reflections for the design of future immersive data analysis environments.

Among others, Donalek et al. (2014) describe their immersive *iViz* tool that is designed to visualize large multivariate datasets in an immersive 3D environment. Through utilization of a HMD and 3D gestural input, a user can interactively explore the displayed data, following an overall abstract data visualization approach (Donalek et al., 2014). An interesting feature of their presented tool is that it enables the user to dynamically map the various data variables of the dataset to different graphical attributes, such as the position in 3D, colors, shapes, sizes, transparency, or textures (Donalek et al., 2014). This allows them to reuse their tool across various datasets and analysis scenarios. The reported insights from the development of their exploratory proof-of-concept tool (Donalek et al., 2014) can be important for the design and implementation of similar tools, especially with respect to its data-agnostic capabilities and the dynamic visual mapping of data that encourage reusability of the tool.

Similar to the overall abstract data visualization approach as presented by Donalek et al. (2014), Wagner Filho et al. (2018) investigated the utilization of 3D scatterplots within the context of IA. Using their system, the data variables of a dataset can be visually encoded according to position, shape, and color (Wagner Filho et al., 2018). Under utilization of various analytical tasks that required the selection and identification of data entities, the authors conducted an empirical investigation to compare three display and interaction configurations, i.e., (1) 2D visualization using a normal monitor with keyboard and pointer input, (2) 3D visualization using a normal monitor with keyboard and pointer input, and (3) immersive 3D visualization using a HMD and physical, tracked controllers (Wagner Filho et al., 2018). The results of their evaluation indicate, among others, that the immersive display and interaction setup required less effort and navigation to solve the analytical tasks compared to the two non-immersive configurations, while also being subjectively more engaging (Wagner Filho et al., 2018). The positive results of their investigation in favor of the IA tool are promising, indicating an overall usefulness of such immersive data analysis solutions, and as such encouraging further investigations in similar directions.

Nguyen et al. (2019) demonstrate the utilization of star coordinates and star plot techniques within the context of immersive data visualization. Interestingly, rather than positioning the individual data entities in the empty 3D space, their placement is facilitated through the display of respective data variable axes that provide additional information cues to assist with the spatial data interpretation (Nguyen et al., 2019). The authors also provided several features to address analytical tasks, for instance to explore, select, filter, and zoom, as well as some preliminary collaborative features (Nguyen et al., 2019). Overall, the work by Nguyen et al. (2019) provides impulses for the design and development of IA tools that are centered around abstract data visualizations similar to Donalek et al. (2014) and Wagner Filho et al. (2018).

Moran et al. (2015) explored the visualization of a multivariate social network dataset, including geolocation data variables, in a more realistic-looking VE, arguably following a situated visualization design (Bressa et al., 2022; Thomas et al., 2018). In particular, they displayed satellite imagery on the virtual floor and placed 3D representations of the buildings on their university campus accordingly on the top (Moran et al., 2015). The individual data entities, i.e., the visualized data items of the dataset, could then be placed and displayed in-situ in the VE in accordance to their geolocation coordinates (Moran et al., 2015). Various interactive features enabled the immersed user to explore the dataset, for instance by navigating through the 3D space, filtering the displayed data entities, and displaying details-on-demand (Moran et al., 2015). Overall, their presented immersive data analysis environment appears to apply a valuable data entity positioning, directly in-situ and as such juxtaposed to virtual representations of real-world facilities. This arguably allows for a contextually relevant analysis of the dataset in regard to its spatial data variables. In the future, it could be interesting to extend such an approach by incorporating features – both in regard to the visualization and interaction – that better reflect on temporal data variables, which are typically of importance within the context of social network analysis. This could allow the immersed user to not just identify where data items exist, but also in regard to when they were created.

Ivanov et al. (2019) present an immersive tool that enables its user to explore individual data items in a dataset by literally *walking among the data*. Based on dataset and scenario, individual data items are represented as abstract human-like avatars in the VE that differ in appearance according to gender and age (Ivanov et al., 2019). The authors provided several interactive features, for instance to group the displayed data entities dynamically based on the available data variables (Ivanov et al., 2019). To navigate in the immersive 3D space, Ivanov et al. (2019) provide two main mechanisms, i.e., a zoomed out perspective to enable the immersed user to explore the dataset in a more overview-like manner, and a zoomed in one that allows the user to walk among the abstract avatars and to display details-on-demand. Based on the experiences gained from their implemented prototype, the authors are able to highlight some interesting challenges, for instance in regard to the immersed user's orientation, situational awareness, and personal space, as well as with respect to overall visual fidelity and unit appearance (Ivanov et al., 2019). The insights from their work provide design considerations that are particularly relevant within the context of the exploration and analysis of large situated datasets in immersive VEs.

Pirch et al. (2021) present a comprehensive implementation of an IA platform that is centered around the exploration of networks in a 3D VE. Besides the visualization of individual data items as network nodes in the 3D space, the authors also provide a multitude of interactive features, implemented using a mixture of natural 3D interactions as well as adapted 2D graphical menus operated through two physical, tracked controllers (Pirch et al., 2021). According to the authors, they designed the overall user interface (UI) with established 2D and 3D interaction concepts in mind, aiming for the IA tool to be rather self-explanatory to the immersed user (Pirch et al., 2021). In addition to the provided visualization and interaction design considerations, Pirch et al. (2021) describe in detail the various implementation aspects, highlighting the various modules that contribute to the overall platform composition. The modular characteristics of their IA platform implementation arguably facilitated not just the aspects of their practical collaborative work, but also in regard to general feature extension and adaptation. Consequently, their work provides value considerations for the respective implementation of future IA tools.

Furthermore, based on the recent advancements in the IA research community, various frameworks and toolkits have been presented with the overall aim to facilitate the development of immersive data analysis environments. For instance, Cordeil et al. (2019, 2017a) present tools that are based on the Unity cross-platform game engine that can be utilized to, comparatively easily, compose immersive multivariate data visualizations. The presented *ImAxes* toolkit is centered around the transfer of typical axes-based visualizations into the immersive 3D space, such as histograms, scatter plots, scatter plot matrices, linked scatter plots, as well as parallel coordinate plots (Cordeil et al., 2017a). A key feature of the toolkit is centered around the interactivity of the data axes, enabling the immersed user to manipulate, reconfigure, and filter the visualized data directly in the VE (Cordeil et al., 2017a). In comparison, *IATK* focuses on providing accessible configuration options for similar immersive visualizations based on a graphical user interface that requires only minimal programming efforts (Cordeil et al., 2019). Similarly, the *DXR* toolkit, also based on Unity, has been developed specifically keeping developer novices without extensive experiences in 3D graphics, Augmented Reality, and VR development in mind, allowing for approachable authoring and rapid prototyping of IA tools and concepts (Sicat et al., 2019). Rather than attempting to support a variety of basic visualization techniques, the *MIRIA* toolkit presented by Büschel et al. (2021) focuses on the visualization and subsequent analysis of spatio-temporal user interaction data. As such, their Unity-based toolkit enables the display of 3D trajectories, position heatmaps, and scatterplots, for instance representing a user's spatial movements over time, allowing the immersed analyst to follow and reiterate on these trajectories with the goal to obtain an understanding of the user's activity (Büschel et al., 2021). Currently, the authors focus on the specific use case of displaying spatio-temporal user interaction data in-situ using augmented reality (Büschel et al., 2021). Naturally, it would be intriguing to adopt their approach to allow an analyst to similarly follow and reiterate on the interactions of a user in an immersive VE that is based on a VR approach. The *UMI3D* toolkit, presented by Casarin et al. (2018) and also based on Unity, differs from the other presented solutions insofar that it focuses on providing dedicated mechanisms that aim to facilitate the implementation of collaborative features, which are of particular relevance within the context of CIA. Arguably similar in comparison to IATK, Butcher et al. (2019) present *VRIA*, a toolkit for the generation of immersive visualizations that is solely based on web technologies with open standards. Their toolkit supports visualization techniques such as 3D bar charts and multivariate scatterplots, and integrate well with other established web-based Information Visualization (InfoVis) and Visual Analytics (VA) libraries, such as D3.js, potentially facilitating the overall development workflow.

3.2 Immersive Data Interaction Using 3D Gestural Input

Over the years, various researchers have investigated the application of 3D gestural input (see Section 2.2.4) as interaction modality within the context of immersive data analysis environments (see Section 2.4). These provide insights and impulses for design considerations of new IA tools that utilize 3D gestural input for the interaction with data in the virtual 3D space.

For instance, LaViola, Jr. (2000) describes an interface that utilizes a multimodal approach of 3D gestural input and voice commands to interact with a scientific data visualization in stereoscopic 3D. Different analysis tools can be attached to the user's hands and moved around in the VE (LaViola, Jr., 2000). Interestingly, rather than selecting these tools from a graphical menu, they implemented voice commands that allow the user to say aloud the tool they want to interact with, following a *show and ask* metaphor (LaViola, Jr., 2000). They also implemented several hand-based grasping configurations to provide navigation features, i.e., user movement as well as translation, scaling, and rotation of the visualization (LaViola, Jr., 2000). An evaluation indicated that their participants valued the tool's ease-of-use after an initial learning phase (LaViola, Jr., 2000). Their results also indicated that the voice command interface worked well in single-user scenarios, while having detection problems in collaborative ones that featured auditory input from more than one user (LaViola, Jr., 2000). As such, the design of multimodal interaction techniques should be carefully considered, especially in regard to potential collaborative scenarios that involve multiple analysts, as described in Sections 2.3 and 2.4.1, in order to avoid potential conflicts with other aspects of the immersive data analysis activity.

Fittkau et al. (2015) explored gestural command design for the interaction with an immersive data visualization following the *software cities* metaphor, implementing several unimanual and bimanual gestural commands to support translation, rotation, zoom, selection, and reset tasks. The results of their evaluation indicate that the users favored one-handed gestures (translation, rotation, selection) over the two-handed (zoom) one that was performed through a rowing motion (Fittkau et al., 2015). Interestingly, the authors attempted to utilize more elements of embodied interaction for the zooming command, such as rotating the user's torso or walking back and forth in the VE (Fittkau et al., 2015). However, such movements would inherently result in a change of the user's field of view, which was not appreciated during early design iterations (Fittkau et al., 2015). Even though more empirical evaluations with a focus on the use of embodied interaction within the context of IA are required, one should arguably carefully consider the purpose and design of whole body interaction techniques for immersive data analysis and interpretation. Mapping features that modify the immersive visualization itself to input that require the user to change their own field of view in the immersive VE, may arguably prevent the user to make useful observations and data interpretations.

Similarly to the work presented by Fittkau et al. (2015), Streppel et al. (2018) explored 3D interaction techniques within the software cities context as well, comparing 3D gestural input, physical controllers, and virtual controls. Their results indicate similar preferences for direct manipulation through 3D gestural input and physical controllers as opposed to virtual controls that were based on adapted 2D graphical menus (Streppel et al., 2018). Even though the physical controller condition received better usability scores, participants stated that they would rather like to use the 3D gestural input in a real-world scenario, as it was subjectively perceived as more natural and appropriate for interactions in a VE (Streppel et al., 2018). The expressed desire for better 3D gestural input controls is quite interesting, indicating that more work in that direction should be undertaken to further improve usability aspects of 3D gestural input within the context of IA.

A VR system developed by Betella et al. (2014) featured 3D gestural input for manipulation and filter operations within a large network visualization. Their interface utilized a hand-based grasping technique for the overall interaction in the immersive VE (Betella et al., 2014). Interestingly, the authors chose to map distinct features to each of the user's hands following a somewhat asymmetric hand interaction approach, i.e., the user's right hand featured a cursor function to highlight and select nodes in the network, while the left hand was used to operate task parameters, such as filter strength and complexity (Betella et al., 2014). Their asymmetric feature mapping strategy is interesting insofar that the authors differentiate between left and right hand interactions instead of following a symmetric approach where the same features are provided independent of which hand performs the posture or gesture. In turn, this allows the potential reuse of simple and comfortable hand configurations for subsequent interactions in the VE, which could be useful considering how many features are available in the VE and how often they are anticipated to be used. Naturally, when following such an asymmetric feature mapping approach, general accessibility aspects for the user should be taken into account as well, for instance allowing the user to chose which features are mapped to the left hand and which to the right, accommodating to the user's dominant hand.

Osawa et al. (2000) investigated hand-based grasping and gestural command techniques for interaction with an immersive graph visualization. Their system allowed the user to select and manipulate individual nodes of the 3D network (translate, lock position in space, adjust characteristics), to translate the user's position in space (move), and to adjust characteristics of multiple nodes through a "spotlight" approach (Osawa et al., 2000). The latter was operated through pointing one's hand in the general direction of the desired nodes, and creating an arc-like spread through moving index finger and thumb apart, enabling dynamic control of the included network nodes (Osawa et al., 2000). The 3D gestural input was considered intuitive and more appropriate compared to the application of 2D techniques for the interaction with the implemented graph visualization in the 3D space, not least as 3D interaction resembles interaction in the real world (Osawa et al., 2000). Nonetheless, the authors also acknowledged some aspects of frustration during the 3D interaction, particularly with respect to the selection of small artifacts that was sometimes prevented due to the input device's lack of precision (Osawa et al., 2000). 3D spatial input technologies and 3D gestural input thereof (see Sections 2.2.3 and 2.2.4) have evolved significantly since 2000, enabling technically more precise interactions in immersive 3D spaces. Nevertheless, technological aspects, such as the precision of the input technology, should be carefully taken into account for the design of useful interactions within the context of IA. After all, the ability to make targeted selections and manipulations are arguably essential during the immersive data analysis activity.

Huang et al. (2017) reported on the design of a 3D gestural interface for interaction with graph visualizations using VR – in concept quite similar to the work presented by Osawa et al. (2000). Their developed immersive VE provided gestures to move and highlight nodes and edges (unimanual interaction), to rotate and translate the entire graph, and to group nodes (bimanual interaction). An evaluation, comparing the implemented gestures with more traditional pointer input (mouse), revealed positive trends towards the participants' ability to manipulate the 3D graph with the gestures, stating that the interface "*was intuitive, easy to learn, and interesting*" (Huang et al., 2017). While their implemented node/edge movement and graph rotation gestures were appreciated for their learnability, some usability issues were identified for the highlight and group gestures that involved aspects of holding a specific hand posture or performing a gesture very quickly (Huang et al., 2017). The results of their work are another example for the circumstance that the input technology's capabilities and overall hand posture comfort should be considered for the interaction design – particularly when assuming that such developed IA tools aim to be applied more frequently and over longer periods of time beyond "just a few minutes"-experiences.

As part of interacting with an immersive 3D trajectory visualization, Wagner Filho et al. (2020) implemented a mixture of hand-based grasping (scale, translate) and gestural commands (single and double tap via index finger to inspect and select). Interestingly, the authors implemented the 3D gestural input through the utilization of physical, tracked controllers that the user had to hold in their hands, while instead displaying an abstract visual hand representation in the VE (Wagner Filho et al., 2020). The authors evaluated their system in comparison to a desktop one, revealing generally better usability scores for the immersive VE (Wagner Filho et al., 2020). Participants overall agreed that the 3D gestural input enabled them to easily and comfortably manipulate the data, resulting in an engaging and intuitive experience (Wagner Filho et al., 2020). Room for improvement was identified towards the index finger tapping mechanism that required to be comparatively precise (Wagner Filho et al., 2020). The authors' interaction design is interesting insofar that they utilized the same overall technique for multiple contextually different features, for instance reusing hand-based grasping as a unimanual technique for translation as well as a bimanual configuration for temporal and spatial scaling (Wagner Filho et al., 2020). However, based on the participant feedback, aspects of its implementation arguably need to be revisited as the similar operation was sometimes perceived as too constraining (Wagner Filho et al., 2020). The authors' work indicates that the reusability of the same interaction techniques for different features in the immersive data analysis environment should be carefully considered in order to enable the user to perform the desired operations without mistake and constraints. Arguably, the VE's ability to interpret user input and infer their in-situ context and intent, as highlighted by Nehaniv et al. (2005) and discussed in Section 2.2.4, could help overcome such challenges.

Austin et al. (2020) investigated common gestural commands for the interaction with large immersive maps that are placed on a virtual floor. In particular, using a participatory design approach, their study participants were asked to come up with hand gestures for typical operations to manipulate the virtual map, such as pan, rotate, zoom, and marker interaction (Austin et al., 2020). Their results indicate that the participants most commonly proposed unimanual gestures for interactions such as pan as well as creating and selecting markers on the map, while proposing bimanual gestures for rotate and zoom operations (Austin et al., 2020). Austin et al. (2020) reflect on their findings, stating that the identified user preferences for these gestural commands need further investigation in regard to performance related matters, such as efficiency, accuracy, and physical fatigue. Austin et al. (2020) also reflect on potential feasibility concerns, stating that an accurate and reliable implementation based on current 3D gestural tracking sensors might be difficult for some of the proposed gestural commands. Naturally, the anticipated frequency of performed features, i.e., how often a specific functionality is used in the VE, depends on the data scenario and task. Examining the proposed gestural commands in regard to their unimanual and bimanual configuration, one could argue that unimanual gestures were utilized for more frequent tasks, such as moving the user's position on the map as well as interacting with markers, while bimanual gestures were proposed for rotation and zoom – tasks that in comparison to user movement and selection may be performed less frequently. Interestingly, bimanual gestural commands for similar features were also utilized in the IA tools presented by Wagner Filho et al. (2020) and Huang et al. (2017), indicating their overall feasibility and appropriateness under consideration of some practical design and implementation aspects, as previously described.

3.3 Hybrid Collaboration Experiences Using Virtual Reality

Concepts of hybrid collaboration, i.e., the application of heterogeneous display and interaction technologies in collaborative scenarios (Fröhler et al., 2022; Neumayr et al., 2018), has been described as part of Sections 2.3 and 2.4.1. Researchers have conducted some interesting work in the past, exploring such cross-device and cross-platform setups that involve at least one type of immersive VR interface.

Wideström et al. (2000) conducted a study to compare two different settings, i.e., (1) a hybrid VR setup consisting of a CAVE-type (Cruz-Neira et al., 1992) and a desktop-based system, and (2) a co-located real-world setup, in regard to collaboration, leadership, and performance aspects within the context of a two-person puzzle solving task. Their results show that the participants reported their contribution to the task completion more unequally in the hybrid VR setup compared to the real-world one. Furthermore, the participants felt a higher degree of collaboration in the real-world setup due to the lack of face-to-face communication in the hybrid VR one (Wideström et al., 2000). Arguably, the integration of additional information cues to better support mutual awareness, as discussed by Benford et al. (1994), could help to overcome such an experienced lower degree of collaboration in the immersive setup, potentially in turn facilitating a more balanced task contribution independent of a user's interface.

Within the context of collaborative educational settings, Thomsen et al. (2019) propose a taxonomy that is the centered around a hybrid VR setup. The authors follow the general concept of one user being immersed in a VE using a VR approach, while one or multiple others are not, defining a distinctly asymmetric user role relationship between *actor* (immersed) and *assistant* (non-immersed) that their taxonomy is designed around (Thomsen et al., 2019). The taxonomy consists of different components (*asymmetric mechanics*, *hardware components*, *game components*, *collaboration mechanics*) in order to address varying degrees of *collaboration asymmetry* (low, medium, high) between actor and assistant (Thomsen et al., 2019). Under consideration of the educational setting, the taxonomy can provide some valuable parameters for the integration of immersive VR technologies in the classroom, actively involving not just the immersed user wearing the HMD device, but also the other "bystanders". Such a setup could arguably also be relevant within other contexts, such as CIA, allowing a subsequent insights transfer of the proposed taxonomy accordingly.

Peter et al. (2018) propose a set of features for a non-immersed user in a guiding role to support communication with an immersed user, similarly to the actor-assistant relationship described by Thomsen et al. (2019). Based on their system setup, they envision the immersed user to have a low degree of control but a high level of immersion, while it is the other way around for the non-immersed *VR-Guide*, who has a high degree of control but a low level of immersion (Peter et al., 2018). The authors describe the design and implementation of a highlighting feature, comparing different variants, with the aim to guide the VR user's attention to specific reference points in the VE based on the non-immersed user's input (Peter et al., 2018). Within the context of CVEs and CIA, such an actor-assistant relationship, similar to the presentations by Thomsen et al. (2019) and Peter et al. (2018), could be useful in data guidance scenarios, for instance when a non-immersed expert guides an immersed novice through the immersive VE, highlighting noteworthy points of interest, that can than be discussed accordingly.

Welsford-Ackroyd et al. (2020) evaluated their proposed system design that allows a non-immersed user, typically in the role of a spectator outside the VE, to actively collaborate with the HMD user through the additional utilization of a large scale display. Camera control and pointing features were provided to the spectator, the later of which – implemented using a virtual laser pointer technique – clearly facilitated the communication between the two collaborators in a task scenario where the immersed user had to place artifacts at certain locations as indicated by the spectator (Welsford-Ackroyd et al., 2020). Their system shows similarities to the VR-Guide by Peter et al. (2018) insofar that the non-immersed user directs the immersed one to a point of interest through visual highlights in the VE. In both cases (Welsford-Ackroyd et al., 2020; Peter et al., 2018), the

HMD user had arguably little to no awareness of the non-immersed user other than through the directed visual references, while the non-immersed user could somewhat "monitor" their immersed partner through the shared point of view, i.e., the HMD user's point of view was displayed on a normal monitor for the non-immersed user to follow. Arguably, this contributes to a rather unequal interplay between the users right from the start.

However, there are also some interesting examples that aim to leverage on more balanced and equal user contributions in hybrid technology setups. For instance, Sugiura et al. (2018) investigated hybrid collaboration between a HMD user and multiple non-immersed users around an interactive tabletop system within the context of interior design. While the immersed HMD user got to perceive the living space from an in-situ real-world perspective, the tabletop system featured a top-down view that allowed its users to see the position and orientation of the immersed user as well as providing an overview of the living space (Sugiura et al., 2018). The HMD user and the non-immersed tabletop users were provided with features that enabled them to point and refer to artifacts of interest in the virtual living space, that were visually highlighted in the respective collaborator interface (Sugiura et al., 2018). The results of their case study indicated that the provided functionality assisted the verbal communication between the collaborators across the two different device modalities (Sugiura et al., 2018). Their presented setup is interesting insofar that next to the hybrid technology setup each interface user assumed also a role that is distinctly different from the other, providing further impulses for the careful design of purposeful collaborative activities that utilize heterogeneous display and interaction technologies.

Clergeaud et al. (2017) explored possibilities for seamless interaction between a HMD user, who is immersed in the VE, and multiple non-immersed users gathered around a meeting table within the scope of an industrial scenario. Based on insights from industry experts, the authors propose and explore several features with the objective to facilitate aspects of awareness, communication, and interaction with the immersed HMD user (Clergeaud et al., 2017). The presented *Interaction through Windows* and *Navigation through Doors* approaches received positive feedback, essentially creating conceptual portals between the virtual and real-world environments, successfully bridging the collaboration between immersed and non-immersed users (Clergeaud et al., 2017). Their collaborative environment aligns well with the overall concept presented by Sugiura et al. (2018) insofar that both systems assume distinct purposes for the involved interfaces, while enabling cross-platform collaboration in a meaningful way.

Gugenheimer et al. (2017) describe design guidelines for co-located hybrid VR experiences that feature an asymmetric user role configuration based on insights from evaluating their developed *ShareVR* prototype. The prototype allowed different types of interaction between a HMD and a non-HMD user based on a combination of VR and floor projection technologies (Gugenheimer et al., 2017). Among others, the authors emphasize on the importance to leverage on asymmetrical aspects, carefully considering each user's role, in order to design meaningful interactions for collaboration accordingly (Gugenheimer et al., 2017). The results of their study are yet another example that highlight the importance of purposefully designing collaborative experiences that are based on the utilization of heterogeneous display and interaction technologies, keeping anticipated user roles in mind.

Insights and experiences of co-located hybrid interaction between a HMD and a non-HMD user are also presented by Lee et al. (2020). The authors designed an application where each user assumed distinct roles, designed under consideration of their respective level of immersion, i.e., assuming a spatial relevant role for the HMD user and a more temporal one for the non-HMD user (Lee et al., 2020). The presented prototype featured a game-like experience that tasked the two collaborators to actively work together in order to navigate successfully through a maze, and was used to evaluate presence, game experience, and different aspects of the users' roles within the scope of multiple experiments (Lee et al., 2020). The results indicate a perceived higher than usual level of immersion of the non-HMD user due to their more active role and involvement in the overall task setup, as well as similar levels of enjoyment and social interaction among both user roles (Lee et al., 2020). Aligned with insights such as presented by Gugenheimer et al. (2017), Lee et al. (2020) clearly defined and distinguished between the different roles and responsibilities for each interface user, resulting in an enjoyable and balanced experience for both collaborators. Consequently, their presented collaborative design can provide useful impulses and considerations for the design of similar future experiences.

Grandi et al. (2019) compared collaborative co-manipulation of 3D virtual artifacts in three different setups, i.e., based solely on augmented reality interfaces, solely on VR interfaces, and a hybrid approach utilizing both augmented reality and VR interfaces. All interfaces provided manipulation features based on unimanual and bimanual techniques to translate, rotate, and scale a 3D artifact in the virtual space (Grandi et al., 2019). The authors evaluated aspects of performance and collaboration in a task scenario where pairs of users had to utilize their respective interfaces to match a 3D artifact's position, rotation, and scale according to a target artifact (Grandi et al., 2019). Interestingly, the sole VR and the hybrid approaches performed better than the setup that was solely based on augmented reality interfaces (Grandi et al., 2019). The positive results around the conditions that involved VR technologies (Grandi et al., 2019) are arguably attributed to more intuitive manipulation features that allowed for spatial 3D interaction, compared to the handheld augmented reality interface that relied on 3D manipulation through a 2D touchscreen. Nevertheless, independent of the technology setup, all pairs showed similar task participation (Grandi et al., 2019), indicating a rather balanced user involvement and collaboration overall.

3.4 Collaborative Information Cues in Virtual Environments

The importance of providing mechanics that aid users with their communication in CVEs has been highlighted in Section 2.3. The support for appropriate referencing, i.e., indicating a specific artifact through a mixture of verbal and nonverbal signals, is arguably even more important in hybrid technology scenarios where not every collaborator is immersed in the same VE. Researchers have investigated various aspects of collaborative information cues in immersive VEs in the past.

Cordeil et al. (2017b) evaluated the collaborative aspects of two IA systems, CAVE-style and HMD-based, in terms of collaboration strategies, shared focus, completion time, self-perception of collaboration, proportion of oral communication, and balance of physical movements. While the users were able to see each other within the physical boundaries in the co-located CAVE-style setup, the HMD-based system required the authors to implement dedicated collaborative information cues in order to visually represent the respective partner in the VE (Cordeil et al., 2017b). As such, Cordeil et al. (2017b) utilized a real-time networking interface to transfer data about the partner's head position, rotation, field of view, as well as their 3D gestural input for display in a user's VE accordingly, aiming to facilitate a sense of presence. Interestingly, even though the users of the HMD-based collaborative system were not able to monitor each others facial expressions or body language, which are important during their communicative efforts to establish a common ground (see Section 2.3), no major issues regarding this limitation were detected within the scope of a low-level graph visualization task (Cordeil et al., 2017b). Even though more empirical studies are required to investigate this matter more thoroughly, their results indicate that the support for facial expressions and body language are not always necessary to successfully complete collaborative data analysis tasks, assuming there are sufficient other collaborative information cues available in the VE.

Considering a hybrid technology as well as an asymmetric user role setup, the VR-Guide interface presented by Peter et al. (2018), briefly described as part of Section 3.3, provided several features to support the guidance of a HMD user immersed in a VE. In particular, their implemented proof-of-concept prototype included three visual approaches as nonverbal information cues, aiming to catch the immersed user's attention and thus guide them towards an artifact as selected through the non-immersed guide (Peter et al., 2018). First, they implemented an *outline* effect, visually highlighting the border of the selected artifact independent of its occlusion, i.e., the outline is visible even with other objects in the VE between user and targeted artifact, leaving it otherwise hidden from the user's view (Peter et al., 2018). Second, they implemented a realistic looking *light beam* technique, similar to a spotlight, allowing the highlighted artifact to be visually distinguished from others (Peter et al., 2018). And third, a *virtual drone* was implemented using a laser beam to point to the selected artifact (Peter et al., 2018). An interesting aspect of their tool is that is allows the VR-guide to customize these signals, for instance by adjusting the color or thickness of the outline effect, or the size and intensity of the light beam (Peter et al., 2018). Their evaluation revealed trends towards a higher acceptance of the outline technique compared to the light beam one, with participants interestingly expressing desire for a combination of outline and virtual drone techniques (Peter et al., 2018). The insights based on their presented reference approaches and subsequent evaluation can provide meaningful impulses for the design of similar collaborative information cues in other contexts, such as CIA.

Compared to the approaches described by Peter et al. (2018), Sugiura et al. (2018) followed a visually different design, implementing a *virtual hand* in a pointing hand posture that literally points to the selected artifact. Their approach was partially inspired by previous work of Stafford et al. (2006), but adapted to support a setup that involved the non-immersed user operating an interactive tabletop application that allowed for touch interaction to select individual objects from a top-down view that are in turn referred to accordingly in the immersive VE (Sugiura et al., 2018). The results of their preliminary user study suggest overall usability of their prototype, but lack a formal evaluation on the effectiveness of the implemented guiding technique (Sugiura et al., 2018).

Similar to the work by Grandi et al. (2019) as presented in Section 3.3, Pinho et al. (2002) investigated various aspects of collaborative co-manipulation of 3D artifacts in immersive VEs. Their presented framework contains an *awareness generator*, i.e., a dedicated module responsible for handling various collaborative information cues (Pinho et al., 2002). More specifically, the authors differentiate between (1) *user information* that indicate the state of a user in the immersive VE, such as their position and orientation, (2) *interaction information* that represent a user's current interaction mode, and (3) *object state information* that provide an indication of what 3D artifact is currently manipulated and by which user (Pinho et al., 2002). Based on the reported preliminary results of a first pilot study, pairs of participants were able to successfully conduct several tasks that involved the co-manipulation of virtual 3D artifacts (Pinho et al., 2002), arguably supported through the implemented awareness information cues. The logical differentiation of awareness cues into the three presented conceptual categories (Pinho et al., 2002) can provide useful design considerations for the structured implementation of collaborative information cues in similar VEs.

Lacoche et al. (2017) investigated different visual approaches with the objective to raise awareness towards other HMD users in a co-located CVE, aiming to prevent physical collisions. First, an *extended grid* representation visualizes the other user's position through a grid-shaped cylinder, allowing a user to avoid to navigate to the same position, while still being able to "look beyond" to prevent occlusion issues (Lacoche et al., 2017). Second, another user was indicated through means of a *ghost avatar*, displaying some features of that user's position and orientation in the VE, i.e., a semi-transparent model of their HMD and hand controllers (Lacoche et al., 2017). Third, a *safe navigation floor* utilizes a heat map inspired visualization approach to display on the virtual floor where it is safe to go and where not, avoiding collision with the other user as well as movement beyond the physical boundaries of the VR system's calibrated safe interaction area (Lacoche et al., 2017). Under introduction of a fourth approach, titled *separated tracked spaces* that limited each user's available safe interaction area as a typical bounding box grid, an evaluation was conducted to investigate effectiveness and user preference for the different approaches (Lacoche et al., 2017). Their results point towards better performance of the extended grid and ghost avatar compared to the safe navigation floor and separated tracked spaces approaches (Lacoche et al., 2017), providing several impulses for the design of collaborative information cues. In contrast to most of the other approaches presented through this section, the scenario presented by Lacoche et al. (2017) is conceptually different insofar that their presented visual approaches do not aim to actively encourage a user to move towards an artifact in the VE, but rather to avoid it.

Ward et al. (2016) compared two visual information cue designs with the objective to guide a HMD user to a specific artifact in the VE. The first design was based on a visual *arrow* that is centrally positioned in the user's field of view, and dynamically updating itself to indicate the direction to the specified artifact in the VE (Ward et al., 2016). The second design was based on the concept of visual *pursuit motion* cues, i.e., an object begins a smooth motion in the user's field of view that urges the user to pursuit the object accordingly until it's target destination in reached (Ward et al., 2016). The results of their comparative evaluation with three conditions, i.e., arrow, pursuit motion, and no cue, indicate a clear favor for visual arrow cues as the majority of participants was not able to react more quickly or reach targets faster using the pursuit motion cues (Ward et al., 2016). Their presented arrow design (Ward et al., 2016) is also interesting insofar that it is always displayed in the immersed user's field of view, even if the targeted artifact is beyond it, for instance behind the user, potentially enabling quick identification through the user accordingly.

Chen et al. (2018) investigated three different modalities as directional information cues during a multitask scenario where a HMD user is immersed in a VE. More specifically, the authors implemented information cues through means of visual, auditory, and somatosensory (vibrotactile) stimuli, which were evaluated across an easy and a hard task (Chen et al., 2018). Based on task completion time and accuracy measurements, the results of their study indicate a favor towards information cues based on visual and somatosensory stimuli over auditory ones (Chen et al., 2018). The reported results in favor of visual and somatosensory stimuli align well with overall design considerations for CIA experiences that anticipate synchronous collaboration between multiple users, including verbal communication for the joint data interpretation and discussion. In such scenarios, auditory information cues may arguably be in conflict and disrupt the users' verbal communication, for instance as discussed by LaViola, Jr. (2000) within the scope of their developed multimodal 3D UI.

Casarin et al. (2018) described a developed toolkit that allows synchronous interaction in VEs through multiple users. To avoid simultaneous manipulation of the same artifact, the authors implemented an *abstract interaction filter* to visually indicate whether or not an artifact is available for interaction (Casarin et al., 2018). Conceptually, their approach is similar to the object state information as part of the awareness generator module presented by Pinho et al. (2002). Three different states, indicated through distinct color coding, provided visual feedback in real-time, i.e., (1) an artifact is *available* for interaction (original color), (2) *hovered* (green color), (3) or currently being *manipulated* (orange color) (Casarin et al., 2018). Their toolkit was validated based on the results of an evaluation that featured a collaborative authoring task, but further investigations are necessary in order to address individual aspects of their toolkit, such as the design of the collaborative information cues (Casarin et al., 2018). Nevertheless, assuming that the user has learned the meaning of the different color codes and the overall composition of the VE allows for easy identification of the manipulated artifacts among all others in the immersive 3D space, the presented color coding design could arguably be a "quick and dirty" approach to indicate targeted artifacts to a collaborator.

Chapter 4

A System for Immersive Data Analysis Using Virtual Reality

Contents

With an overall foundational understanding and various impulses obtained from the discussed related work (see Chapters 2 and 3), the practical design and development of an immersive data analysis system that is centered around a Virtual Reality (VR) approach can begin. After all, in order to investigate more closely the interaction with spatio-temporal data in immersive Virtual Environments (VEs) as well as overall collaborative aspects, respective technological artifacts need to be implemented. Consequently, this chapter is concerned with the presentation of a conceptual system architecture that aims to serve as a foundation for the implementation of these artifacts.

First, Section 4.1 provides an overview of a total of 18 requirements across all three research objectives (see Section 1.2). These requirements have been defined to guide the overall system, interaction, and collaboration design on the one hand, while on the other facilitating a more formal description of the subsequently developed technological artifacts. Thereafter, Section 4.2 begins by presenting an overview of the proposed system architecture, and continues by describing in detail the four major building blocks that the system is composed of. Finally, following the structure of the presented system architecture, details are provided about the implementation of all developed technological artifacts as described throughout Chapters 5 and 6.

4.1 Requirement Analysis

Under consideration of the thesis goal, objectives, and design space, as described in Sections 1.2 and 2.6, the motivation was to build a data analysis system that utilizes immersive display and interaction technologies. In particular, the vision was to situate the immersed user in a computer-generated three-dimensional (3D) VE, populated with multivariate data that could be interactively analyzed accordingly. As a starting point to facilitate and guide the description, design, and implementation of such an Immersive Analytics (IA) system on a general level, independent of data context and scenario, a set of requirements are defined to address aspects of the three defined research objectives as presented in Section 1.2, i.e., (1) the development of an immersive data analysis system, (2) the design of a 3D user interface (3D UI) to allow interaction with spatio-temporal data, and (3) the support for collaboration using heterogeneous interface types and user roles. For the definition and categorization of these requirements, the structure proposed by Shneiderman et al. (2017, Chapter 4.3.1) was followed, differentiating between *functional*, *non-functional*, and *user experience* requirements. Functional requirements assume a rather user-centered perspective, stating concretely what the user shall be able to do using the developed system by specifically describing aspects of its behavior. Non-functional requirements describe specifications in regard to the system's overall operation without being directly linked to a particular feature. Thus, non-functional requirements commonly describe general system aspects, for instance with respect to hardware and software components as well as performance and reliability, to name just a few. Finally, user experience requirements can be considered a subtype of non-functional requirements, specifically those that are related to user interface (UI) and interaction design matters. The overview of the defined *requirements* (REQs), including their type and which of the three research objectives they address, is presented in Table 4.1.

4.2 System Architecture

Facilitated through the defined requirements that describe key aspects of the immersive data analysis system proposed within the scope of this thesis (see Section 4.1), an overall system architecture was designed to aid its implementation accordingly. This architecture, illustrated in Figure 4.1, provides a conceptual overview of all involved technological components that compose the entire IA system throughout its three major VE iterations, as outlined in Section 1.3. The architecture is composed of four major building blocks and can be summarized as follows.

The first building block is concerned with the *Data Structure Reference Model*. In particular, multivariate data that are to be visualized require certain preprocessing and transformation in order for any visualization tool to handle it appropriately,

No.	Requirement Description	Type	Objective
REQ 1	The system should enable its user to visually immerse themself in the VE using a HMD device, enabling them to look around through corresponding head movements as one naturally does in the real world.	UXR	RO 1
REQ 2	The system should enable its user to interact in the VE through the utilization of 3D gestural input, i.e., using their hands to directly interact with visible artifacts or using specific hand postures and gestures to trigger corresponding features.	UXR	RO 1
REQ 3	The system should be based on hardware and software technologies that are widely accessible on the consumer market.	NFR	RO 1
REQ 4	The system should be designed in a modular manner that allows for easy adaptation and feature extension.	NFR	RO 1
REQ 5	The system should be data-agnostic, i.e., it should be able to display data in accordance to a data structure reference model (data transformation). This enables compatibility across different data contexts as well as multiple data sources.	NFR	RO 1
REQ 6	The system should provide features to enable the reception of data from remote repositories and the transmission of data to remote repositories.	NFR	RO 1
REQ 7	The user should be able to move freely in close proximity to data entities, i.e., those in the VE within the corresponding physical boundaries of the VR system's calibrated safe interaction area.	FR	RO 1
REQ 8	The user should be able to travel to faraway data entities, i.e., those in the VE beyond the physical boundaries of the VR system's calibrated safe interaction area.	FR	RO 1
REQ 9	The user should be able to inspect virtual data entities in more detail, i.e., to display further information and details-on-demand.	FR	RO 1
REQ 10	The user should be able to capture data discoveries during their immersed analysis activity in the VE, and be able to revisit those afterwards outside the VE.	FR	RO 1
REQ 11	The system should provide visual feedback about the user's in-situ spatial and temporal contexts, i.e., enabling the user to identify the *where* (spatial) and the *when* (temporal) in regard to their current data analysis context.	UXR	RO 2
REQ 12	The user should be able to manipulate their in-situ spatial data analysis context, i.e., to travel to individual data entities in the VE.	FR	RO 2
REQ 13	The user should be able to manipulate their in-situ temporal data analysis context, i.e., to manipulate the displayed time event or range in the VE.	FR	RO 2
REQ 14	The user should be able to perform typical analytical tasks in the VE, for instance filter and reconfiguration operations.	FR	RO 2
REQ 15	The system should provide features to enable real-time connectivity to other interfaces, i.e., to allow the reception of input signals as well as the transmission of output signals, in order to enable the implementation of collaborative features.	NFR	RO 3
REQ 16	The system should provide visual feedback of incoming collaborative information cues in each interface, i.e., to enable its user to identify the spatial and temporal data analysis contexts of their respective collaborator.	UXR	RO 3
REQ 17	Each user should be able to share outgoing collaborative information cues using their interface, i.e., to provide their in-situ spatial and temporal data analysis contexts to their respective collaborator.	FR	RO 3
REQ 18	The users should be able to verbally communicate with each other during their joint data analysis activity, for instance to coordinate their efforts as well as to interpret and discuss their data discoveries.	FR	RO 3

Table 4.1: Overview of the defined functional (FR), non-functional (NFR), and user experience (UXR) requirements (REQs) to guide the design and implementation of the proposed IA system, in accordance with the defined research objectives (ROs; see Section 1.2).

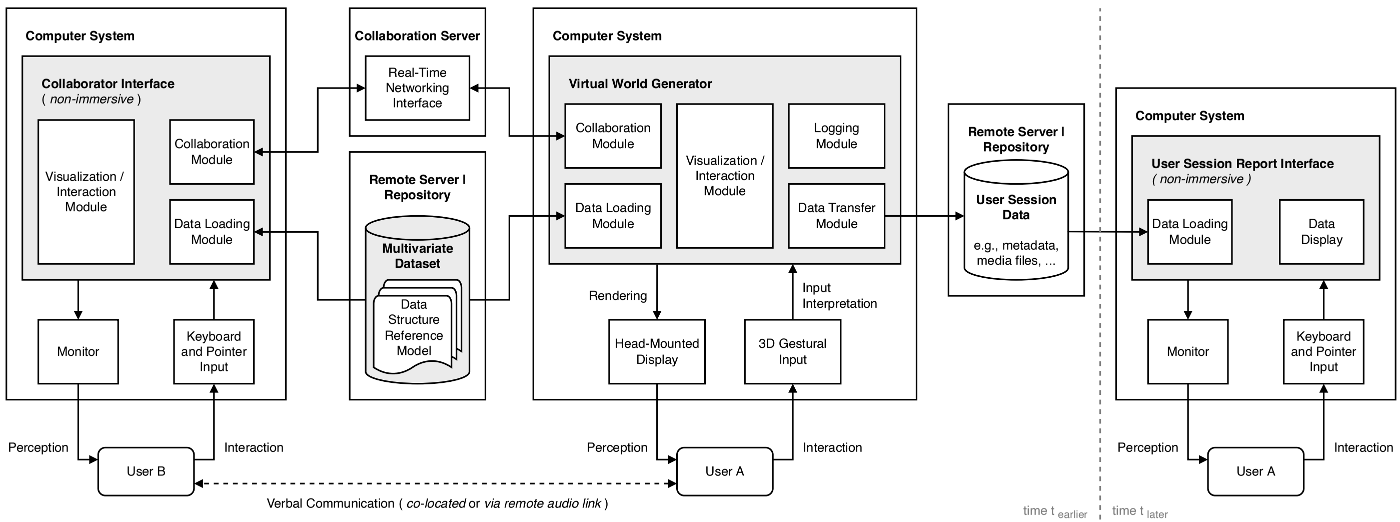


Figure 4.1: System architecture, conceptually illustrating all technological components that are incorporated in the design of the immersive data analysis system as a foundation for the three major VE iterations presented in this thesis, as outlined in Section 1.3.

i.e., to actually generate the corresponding graphical data visualization. As such, a predefined data structure reference model provides an important template for the corresponding mapping of any multivariate data, from various contexts and scenarios as well as potentially multiple sources, preparing the data to be processed and used in the VE or any other type of visualization tool. Within this context, it is useful to keep the actual data separated from the visualization application, i.e., loading the data from an external source, such as a *Remote Server* or *Repository*, rather than having the data directly compiled as part of the application itself.

The second building block corresponds to the *Immersive VE*, dedicated to data exploration and analysis. More specifically, at the architecture's center, a computer system is running the Virtual World Generator (VWG), maintaining the VE that the user is immersed in through the utilization of a head-mounted display (HMD) device and 3D gestural input. Essentially, this VE represents the immersive data analysis interface, and is composed of different modules that handle various key functionalities. While the *Data Loading Module* is essential for the actual import and loading of the data according to the pre-defined data structure reference model, the most important module is the *Visualization/Interaction Module*, responsible for the actual data visualization as well as the provision of the application's interactive features.

The third building block is related to the *Transfer of User Session Data* as a result of the data analysis in the immersive VE. For instance, the user may chose to take notes in order to capture their discoveries for later use outside the VE as input for the next analysis tasks. User session data can be metadata describing the user's data analysis context at a specific point in time in the VE as well as potentially various media files, such as screenshots or voice-over recordings captured directly during the analysis activity. Similar to the loading of data, data that originate from a user's data analysis activity in the VE can be transferred to an external source, such as a remote server or repository. From there, the user session data can be served to other applications, for instance in the format of a *User Session Report Interface* that processes a user's session data to be (re-)viewed in a normal web browser after the immersive activity has finished.

Finally, the fourth building block is concerned with the system's *Collaboration Infrastructure*. In particular, to enable collaboration with a non-immersed user who is using different types of display and interaction technologies for the exploration and analysis of the same data, or aspects thereof, a *Collaboration Server* is needed. That server is responsible for providing a *Real-Time Networking Interface* that allows client applications, such as the immersive VE and a non-immersive desktop interface, to connect with each other. Such a connection is essential for the transmission and reception of signals from a collaborator, allowing for the implementation of collaborative information cues directly in the respective interfaces. The availability of a *Verbal Communication* channel between

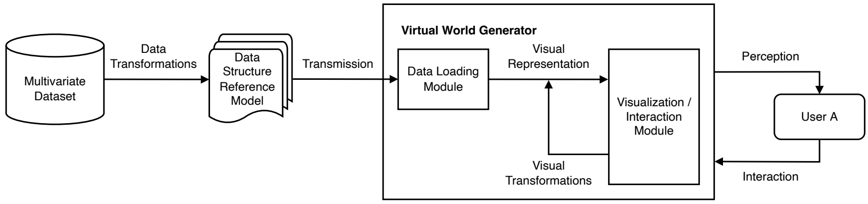


Figure 4.2: The Data Structure Reference Model, integrated in the proposed system architecture, and presented as part of the implemented stages of the visualization process, also referred to as Visualization Pipeline, adapted from Ward et al. (2015, Chapter 4.1).

the collaborators during their synchronous collaboration is anticipated, either due to the collaborators being physically co-located or via a remote audio-link, enabling them to speak to each other.

The modularity of these four building blocks, in regard to their respective modular characteristics internally as well as within the scope of the IA system composition, allows for easy adaptation and feature extension (REQ 4). The remainder of this section provides a more detailed view on the different individual aspects of the designed system architecture, both conceptually as well as practically implemented within the scope of this thesis.

4.2.1 Data Structure Reference Model

A central aspect of any data visualization is concerned with the preparation of the data for the actual visualization through the respective interface (Ward et al., 2015, Chapter 4.1). Naturally, this holds true within the scope of this thesis and the visualization of data in an immersive VE. To support the process of loading data in the VE and consequently creating the visual structures that are representing the data, the concept of a data structure reference model is introduced. It provides a template that enables the VWG to practically import, process, and generate the respective data visualizations in the virtual 3D space. Thus, the data structure reference model also contributes to the VE's data-agnostic capabilities, i.e., as long as (raw) data is transformed in accordance to the template as outlined by the reference model, the VE can support different datasets (REQ 5). This may be useful for the application and reutilization of the immersive data exploration and analysis system across various data contexts and scenarios as well as for supporting data from multiple sources. Figure 4.2 illustrates the overall concept and purpose of the data structure reference model.

Essentially, a predefined data structure reference model needs to capture those aspects of a dataset that are relevant for the anticipated visualization. Within the context of multivariate data in general, and spatio-temporal data in particular, this requires the model to hold potentially various data types (textual, numerical, boolean) in a structured manner in order to accurately represent the original data as well as to allow for convenient computational processing through the visualization application. In practice, various file formats exist that support such structured data representation and convenient processing. Three common file format standards (REQ 3) include JavaScript Object Notation (JSON),[1] Extensible Markup Language (XML),[2] and comma-separated values (CSV).[3] Another aspect and advantage of utilizing an established file format standard such as JSON, XML, or CSV is the wide support by modern software libraries, frameworks, software development kits, and engines, for the serialization and respective deserialization when transferring data from one application to another, even over the Internet using a respective networking protocol (REQ 6).

4.2.2 Immersive Virtual Environment

At the center of the presented system architecture is naturally the immersive VE, representing the computer-generated 3D space that is utilized for data exploration and analysis. Upon closer inspection, the mapping in accordance to the human-VE interaction loop as described by Bowman and McMahan (2007) and presented in Section 2.2 becomes apparent. More specifically, a computer system is running the VWG, i.e., the software responsible for the creation and maintenance of VE's state. Its graphical representation is rendered and displayed on the HMD, worn by the user who visually perceives the VE (REQ 1). User interaction is possible through 3D gestural input (REQ 2). The input signals are interpreted by the VWG and the state of the VE is updated accordingly. Additionally, the position and orientation of the HMD worn by the user are also transferred as input signals to the VWG, allowing for the update of the rendered user perspective accordingly. Furthermore, the visualized data in the VE is based on a multivariate dataset that has been preprocessed in accordance to the respective data structure reference model (see Section 4.2.1). Within the scope of this thesis, its design space, and the particular IA context of the VE, the VWG is designed to be composed of five conceptual and broadly interconnected modules, responsible for the dedicated tasks of data loading, visualization and interaction, logging, data transfer, and collaboration.

The data loading module is responsible for loading the multivariate dataset (REQ 6) in order to create the respective visual representations and thus initiate the overall state of the VE at start-up. After the data in the format as specified

[1]Tim Bray. The JavaScript Object Notation (JSON) Data Interchange Format. Retrieved June 1, 2022, from `https://datatracker.ietf.org/doc/html/rfc7159`

[2]World Wide Web Consortium. XML Core Working Group Public Page. Retrieved June 1, 2022, from `https://www.w3.org/XML/Core/`

[3]Yakov Shafranovich. Common Format and MIME Type for Comma-Separated Values (CSV) Files. Retrieved June 1, 2022, from `https://datatracker.ietf.org/doc/html/rfc4180`

by the data structure reference model is received, the application's internal data structures for the respective handling through the VWG are constructed with the objective to be processed by the visualization/interaction module.

The visualization/interaction module is conceptually the main module of the VWG. It handles the actual composition of the VE, i.e., the visualization of the data in the virtual 3D space as well as the display of additional visual artifacts in the environment and 3D UI, such as the representation of the virtual floor and user feedback. The module utilizes the loaded data to instantiate the individual data entities of the dataset and to populate the VE with them accordingly. The data is visualized in accordance to the defined visual mapping strategy and subsequent implementation, i.e., what visualization technique is applied to visually represent the data and consequently what data attributes affect that visual representation. In addition to the initial data visualization, the module also handles respective state changes, for instance as a result of corresponding user interactions with the visual data representations. As such, visualization and interaction design are conceptually closely coupled. The received and interpreted user input signals are translated as part of this module into the context of the immersive VE, mapping the user interactions to the dedicated functionalities available in the VE. Means of user interaction should be provided with respect to typical analytical tasks, for instance data selection, reconfiguration, or filtering.

The *Data Transfer Module* can been seen as a counterpart to the data loading module. During run time, i.e., when the user is conducting their data analysis in the immersive VE, the data transfer module is responsible for capturing information about aspects that the user desires to take with them from the immersive analysis activity (REQ 6; REQ 10). In other words, its main purpose is to allow the user to retain discoveries about their data exploration in the immersive VE through means of note taking. As such, upon the user's command, the data transfer module is responsible to capture information about the user's state in-situ, for instance by recording contextual information about the currently selected data entity or taking a screenshot of the user's field of view.

Similar with respect to the ability to record in-situ contextual information about the user, the *Logging Module* integrated as part of the VWG takes care of recording important state changing events in the immersive VE. Such events are mainly caused directly as a result of user interaction. However, there may also be system relevant events that are worth considering to capture, such as an incoming signal from a collaborator. Generally, the main purpose of the logging module is twofold. First, it is invisible to the user, collecting logging data without the user's notice in such a way as to neither disrupt nor influence their interactions in the VE. And second, the captured logging data provides an objective measurement for the analysis and evaluation of the user's task performance in the VE (see Section 2.5), potentially even allowing to reconstruct the logged events over time. Within the bigger picture of the VWG, the logging module can be seen rather optional with regard to the actual experience in the immersive VE. However, it is an essential data collection method that is relevant for the empirical evaluation and analysis of user interactions.

Finally, the *Collaboration Module* is responsible for the integration of any collaborative features in the VE. It establishes and maintains a connection to a respective collaboration server that is tasked with the bidirectional transfer of signals to and from the immersive VE, i.e., incoming and outgoing signals (REQ 15). Incoming signals may be received as a result of a collaborator's interaction in their respective interface, sending a signal that requires processing in the immersive VE. For instance, this could be a data entity in the visualization that should be highlighted as a result of the collaborator pointing to it. The same applies vice versa, i.e., the collaboration module is preparing and transmitting outgoing signals as the result of the immersed user's interactions in the VE, allowing respective input in the collaborator's non-immersive interface.

4.2.3 User Session Data Transfer

The main motivation behind the transfer of user session data originates from the desire of taking notes of discoveries during the immersive data analysis activity. For instance, compared to interactive Information Visualization (InfoVis) or Visual Analytics (VA) tools that are based on non-immersive technologies, it is easily conceivable that the analyst is taking notes in order to keep track of data discoveries. Such note taking can take place either directly in the tool itself, via a separate note taking application that is running on the same device and can be easily switched to, or even via pen and paper. However, in a setup that is based on immersive display and interaction technologies, where the user is wearing a HMD and situated in a VE, such note taking mechanisms are no longer easily accessible in a conventional manner. As such, functionalities to take notes need to be integrated directly in the VE in order to allow the analyst to keep track of their discoveries. For this purpose, the system architecture includes a conceptually dedicated procedure to implement this matter (REQ 10).

Any type of note that is taken directly in the immersive VE is transferred via its data transfer module (see Section 4.2.2) to an external server or repository outside the VWG for persistent storage, i.e., to be available even after the immersive data analysis activity has ended. The user session data is stored in a structured and bundled manner as to organize any type of data related to the individually taken notes. For instance, the user may make an interesting observation in the data and choose to take a note about that discovery. As such, a note may include information about the current data exploration context of the user, i.e., where they are located and what data entity is selected, a screenshot of the user's field of view in the VE, as well as even a voice-over recording that allows the user to articulate their current thoughts. Consequently, all these metadata and media files need not only to be transferred to the external storage, but also information about them belonging together.

Finally, with the immersive data analysis activity concluded, the user may choose to revisit and view the notes they have taken. Therefore, a subsequent interface to display and revisit these notes is required. The user session report interface is tasked to provide such functionalities. In the format of an own dedicated interface, it can load the respective user session data from the external repository, and prepare the data for display and potential interaction. For the purpose of the illustration within the scope of the proposed system architecture, the user session report interface is presented utilizing a normal monitor with keyboard and pointer input. However, such an interface could equally well run on a portable tablet device with touch input.

4.2.4 Collaboration Infrastructure

The main purpose of the collaboration infrastructure is to provide the technological foundation and overall workflow to connect the immersive VE with other interactive data analysis tools. Such a connection allows for the creation of collaborative data experiences, essentially providing an interface between different interactive data analysis tools, independent of their display and interaction technologies. Within the presented system architecture, a client-server model is utilized to establish a connection between the immersive VE as one client, and the non-immersive *Collaborator Interface* as another. Both client applications connect to a collaboration server that is responsible to receive (incoming) signals from a client and to transfer (outgoing) signals to the other client. Signals within this context are event messages from a client interface, representing structured information that are forwarded to and processed by the respective other client interface. For instance, the immersed user may select a data entity in the VE, and information about that selection are transferred accordingly to the collaborator's interface that in turn provides visual feedback, and vice versa. This allows for the bidirectional transfer of signals in order to update a client interface in accordance to the state of their respective collaborator (REQ 15).

Compared to the immersive VE, the non-immersive collaborator interface, such as an interactive InfoVis or VA application, has to be composed of at least three conceptual modules. First, it requires its own visualization/interaction module that specifies and implements the visualization and interaction techniques in accordance to the tool's data exploration and analysis purpose. Secondly, it also requires a data loading module that imports the data that are to be visualized in the interface. Finally, aligned with the described functionalities as part of the VWG, a collaboration module is also required as part of the collaborator interface, handling all incoming and outgoing signals accordingly. It is noteworthy that modern InfoVis and VA applications are likely more complex, incorporating additional conceptual modules that assist with the implementation of their purpose. However, the three described modules arguably represent the essentials for the illustration of such applications within the context of the presented system architecture. Furthermore, the non-immersive collaborator interface is presented utilizing a normal monitor as a display device with keyboard and pointer input – an arguably traditional setup. Naturally, the interface may also be based on alternative technologies, such as a tablet device with touch input.

4.2.5 Implementation Overview

The purpose of this section is to provide a general overview about the practical software implementation of the presented system architecture within the scope of this thesis, and as such the various technological artifacts as part of the three major VE iterations (see Section 1.3). All utilized hardware and software technologies are affordable and widely accessible (REQ 3). Additionally, selected parts of the developed system have been published as open source, freely available on the Internet, for other researchers, practitioners, and students to use. Appendix B provides a brief summary of these parts, including references for online access.

Data Structure Reference Model JSON and CSV file format standards have been utilized for the practical implementation of the data structure reference model. In particular, JSON was used for its implementation throughout the first (*Sphere*) and second (*Stacked Cuboid*) VE iterations, providing the required means for a structured data representation as key-value pairs. It is noteworthy that a format such as XML would have been technically suitable for this purpose as well, but JSON was chosen due to convenience with respect to (1) the author's experience with JSON, and (2) the close integration with the software technologies used for the development of the various technological artifacts of the system, such as the immersive VE, the remote server, and the non-immersive collaborator interfaces. For the third VE iteration (*3D Radar Chart*), the CSV file format standard was utilized for a, compared to JSON, more lightweight representation of the time-series data, which was sufficient during that stage. Furthermore, the implementation of the remote server that was tasked with the storage of the multivariate data, the processing of the data into the data structure reference model, and the transmission of the requested data to client applications, was based on Node.js[4] and its provided native features.

Immersive Virtual Environment The Unity[5] cross-platform game engine has been utilized for the software implementation of the VWG throughout all three VE iterations. The source code is written in C#.[6] The utilized Oculus Rift CV1 and HTC Vive HMD devices connected to the VWG have been described in

[4]OpenJS Foundation. Node.js. Retrieved June 1, 2022, from `https://nodejs.org`

[5]Unity Technologies. Unity. Retrieved June 1, 2022, from `https://unity.com`

[6]Microsoft Corporation . C# documentation. Retrieved June 1, 2022, from `https://docs.microsoft.com/en-us/dotnet/csharp/`

Section 2.2.2. The Leap Motion Controller for the utilization of 3D gestural input with the VWG has been described in Section 2.2.4. Unity provides application programming interfaces, either natively or through third party support, for the technological integration with such devices, such as the SteamVR Unity Plugin[7] and the Leap Motion Unity Plugin/Modules.[8] The data loading via respective network communication protocol, such as Hypertext Transfer Protocol Secure,[9] has been implemented using the application programming interface as natively provided by Unity. The logging module also utilizes an application programming interface natively provided by Unity for writing and storage of textual data in a CSV file to the computer system's file storage system.

User Session Data The remote server responsible for the storage of the user session data has been implemented using Node.js and its provided native features to write data to the server's file system. The data transfer module as part of the VWG, tasked with recording user session data and transferring the data to the remote server via respective network communication protocol, such as Hypertext Transfer Protocol Secure, has been implemented using the application programming interface as natively provided by Unity. The user session report interface has been implemented as a web application using HTML5[10] and CSS3[11] as well as the D3.js[12] visualization library.

Collaboration Infrastructure The collaboration infrastructure is based on a client-server model that generally utilizes the WebSocket[13] communication protocol. The collaboration server was implemented using Node.js throughout the second (*Stacked Cuboid*) and third (*3D Radar Chart*) VE iterations. The implementation of the client and server endpoints for the collaboration infrastructure during the second VE iteration was based on Socket.IO,[14] a library that is built on top of the WebSocket standard. The at the time utilized *unity-socket.io* library that allowed for Socket.IO integration with a Unity client application, i.e., the VWG, has unfortunately been deprecated during the development process. Therefore, the developed collaboration artifacts as part of the third VE iteration

[7] Valve Corporation. SteamVR Unity Plugin. Retrieved June 1, 2022, from `https://valvesoftware.github.io/steamvr_unity_plugin/`

[8] Ultraleap. Leap Motion Unity Plugin/Modules. Retrieved June 1, 2022, from `https://github.com/ultraleap/UnityPlugin/releases/`

[9] Eric Rescorla. HTTP Over TLS. Retrieved June 1, 2022, from `https://datatracker.ietf.org/doc/html/rfc2818`

[10] WHATWG (Apple, Google, Mozilla, Microsoft). HTML: Living Standard. Retrieved June 1, 2022, from `https://html.spec.whatwg.org/multipage/`

[11] World Wide Web Consortium. Cascading Style Sheets. Retrieved June 1, 2022, from `https://www.w3.org/Style/CSS/`

[12] Mike Bostock. D3.js - Data-Driven Documents. Retrieved June 1, 2022, from `https://d3js.org`

[13] Alexey Melnikov and Ian Fette. The WebSocket Protocol. Retrieved June 1, 2022, from `https://datatracker.ietf.org/doc/html/rfc6455`

[14] Socket.IO. Socket.IO: Bidirectional and low-latency communication for every platform. Retrieved June 1, 2022, from `https://socket.io`

utilized the native WebSocket Secure API.[15] The Node.js server implementation was facilitated through the use of the ws[16] library. Client applications based on JavaScript integrate with a WebSocket server endpoint natively without the need for additional libraries. The collaboration module as part of the VWG, developed in Unity, utilized the websocket-sharp[17] library in order to connect to the WebSocket server endpoint. The non-immersive collaborator interfaces have been implemented as web applications throughout the second and third VE iterations, running in a normal browser, using a mixture of HTML5, CSS3, the D3.js visualization library, and TopoJSON.[18]

[15] Mozilla Corporation. The WebSocket API (WebSockets). Retrieved June 1, 2022, from `https://developer.mozilla.org/en-US/docs/Web/API/WebSockets_API`

[16] Einar Otto Stangvik. ws: a Node.js WebSocket library. Retrieved June 1, 2022, from `https://github.com/websockets/ws`

[17] Sta. websocket-sharp. Retrieved June 1, 2022, from `https://github.com/sta/websocket-sharp`

[18] Mike Bostock. TopoJSON. Retrieved June 1, 2022, from `https://github.com/topojson/topojson`

Chapter 5

Spatio-Temporal Data Analysis Using Virtual Reality

Contents

The conceptual structure and technological aspects of the architecture for an immersive data analysis system (see Chapter 4) can provide an overall foundation for the development of respective Immersive Analytics (IA) tools, for instance

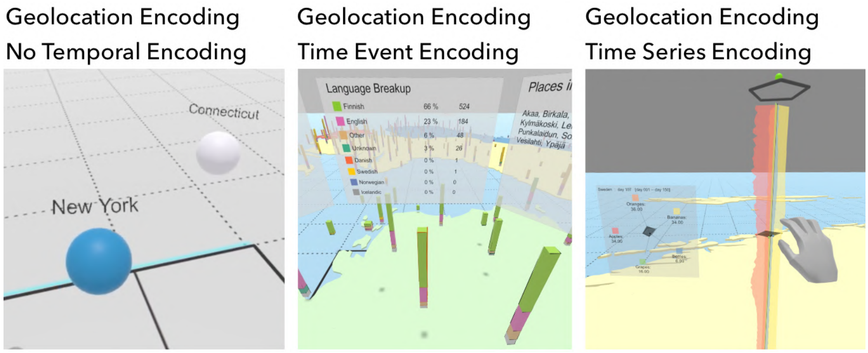


Figure 5.1: Preview of the three VE iterations presented in this chapter, including the applied spatial and temporal visual encoding. **Left**: VE Iteration 1 (*Sphere*). **Center**: VE Iteration 2 (*Stacked Cuboid*). **Right**: VE Iteration 3 (*3D Radar Chart*).

as described throughout this chapter. To investigate immersive interaction with spatio-temporal data based on three-dimensional user interfaces (3D UIs), three major Virtual Environment (VE) iterations have been developed, incrementally building upon each other and increasing in complexity. The first VE iteration visualizes individual data items as *Spheres*, featuring a geolocation encoding, as these spheres are placed in the VE in accordance to the actual geolocation coordinates of the underlying multivariate data. However, these spherical data entity visualizations do not provide any indications of temporal data values. This is instead addressed within the second VE iteration, utilizing a *Stacked Cuboid* design to visually encode the location's temporal context for a single time event at a time. Finally, the third VE iteration expands on that concept, encoding not just a single time event, but an entire time series for each data variable of a location. To achieve this, the immersive *3D Radar Chart* design is introduced. Figure 5.1 provides a preview of these three VE iterations and their data entity visualization designs. Naturally, the works presented throughout Chapters 2 and 3 have informed various design aspects of these VE iterations.

This chapter begins with Section 5.1 by presenting some foundational aspects in regard to multivariate data in general, and spatio-temporal data in particular. Section 5.2 provides an overview of data analysis tasks that are relevant within the context of data interaction. Both sections establish the related terminology that is adopted throughout this thesis.

The first VE iteration is presented in Section 5.3, enabling open data exploration using a Virtual Reality (VR) approach. The design and development allowed, among others, the become practically familiar with the involved technologies and overall concepts. Using the VE, a user could explore data from the 2016 US presidential election. A first set of basic interactive features was implemented, for instance to travel, select, display details-on-demand, and filter. The VE iteration was used to confirm the appropriateness of 3D gestural input as an interaction modality within such contexts, comparing it against gamepad and physical, tracked controllers. The evaluation results informed the choice and focus of using hand interaction throughout the next VE iterations.

The second VE iteration is presented in Section 5.4. Its main context and scenario are centered around the spatio-temporal analysis of language variability on social media from a sociolinguistic perspective. The developed interactive VE was evaluated within the scope of an overall case study with relevant experts and students from the linguistics community. The immersed users were able to extract relevant insights with respect to the spatial and temporal contexts of the data, while at the same time enjoying their engaging experience.

Finally, the third VE iteration is presented in Section 5.5. The introduced *3D Radar Chart* design expanded on the concept of not just encoding a single time event but a time series, facilitating the exploration of the temporal context in the immersive environment. Two empirical evaluations were conducted within the scope of this VE iteration. A first initial prototype with a basic set of interactive features was evaluated to assess and confirm the validity of the overall visualization design. Based on that validation and the received feedback, the VE was extended through additional interactive features, among others, closely aligned with data analysis tasks, interaction techniques, and comfort considerations, and was evaluated once more.

5.1 Spatio-Temporal Data

To obtain a better understanding of spatio-temporal data, it is helpful to first have a high level view on the concept and terminology of *data* in general. Andrienko and Andrienko (2006, Chapter 2) describe data as the result of measurements and observations of phenomena, and consequently *data analysis* as the process of studying these phenomena through the means of analyzing the collected data. Generally, data is described through two fundamental components, i.e., a *referential* component and a *characteristic* component (Andrienko and Andrienko, 2006, Chapter 2.1). The referential component provides the overall context or description of the measurement. The characteristic component on the other hand represents the actual result of the measurement, for instance as numerical, categorical, or ordinal values (Dix, 2020, Chapter 4.1.1). The referential component is also described as *referrer* or *data variable*, while the characteristic component can also be described as *attribute* or *data variable value* (Andrienko and Andrienko, 2006, Chapter 2.1). From a computer science perspective, the analogy to the concept of a *key–value* pair comes to mind to facilitate the understanding of these two basic data components, for instance as a referrer–attribute pair or a data variable–data variable value pair. Consequently, a *dataset* is the collection of at least two, but commonly many more, such data variable–data variable value pairs. Within that context, an individual data variable–data variable value pair in a dataset is also referred to as *data item* or *data entry* (Ward et al., 2015, Chapter 5.1). Furthermore, measured and observed phenomena nowadays are often rather complex, relying on a multitude of data variables that are associated with a data item. A dataset where each data item features two or more data variables is referred to as a *multivariate* dataset, as opposed to a *univariate* dataset where each data item features just one data variable (Ward et al., 2015, Chapter 7.1).

Two common aspects concerned with the measurement and observation of real-world phenomena are *space* and *time*, i.e., where and when a measurement or observation was recorded (Andrienko and Andrienko, 2006, Chapter 2.1). *Spatial* data variables contain geographic information (Andrienko and Andrienko, 2006, Chapter 2.1.2), commonly representing a *location* (point) or *region* (polygon) as latitude and longitude values (Wikle et al., 2019, Chapter 2.1). Alternatively, depending on the overall data context, spatial data variables can also take on the format of lines, trajectories, or objects (Wikle et al., 2019, Chapter 2.1). Aigner et al. (2011, Chapter 4.1.1) provide a comprehensive overview of the characteristics of time that may be used to represent *temporal* data variables. As such, a time domain's *scale* may be *ordinal* (relative order), *discrete* (distinct time unit), or *continuous* (numerical order). A time domain's *scope* is differentiated as *point-based* (a distinct point in time with no duration) and *interval-based* (a range in time with a duration). Furthermore, a time domain's *arrangement* may be *linear* (an ordered model clearly distinguishing between past and future) or *cyclic* (a composition of recurring time elements). The *viewpoint* on a time domain may be *ordered* (time events occur one after the other), *branching* (branching points allow for comparison of alternate scenarios), or based on *multiple* perspectives. A multivariate dataset that consists of data items where each item features data variables with respect to at least one spatial and one temporal context can hence also be specifically referred to as a *spatio-temporal* dataset (Wikle et al., 2019, Chapter 2.1).

The Nordic Tweet Stream (NTS) corpus (Laitinen et al., 2017), in more detail described as part of the data context and scenario in Section 5.4, represents an example for a typical spatio-temporal dataset. The corpus consists of data collected from the social networking platform Twitter. In particular, each data item represents a post on the platform that originated from somewhere within the Nordic region, i.e., Denmark, Finland, Iceland, Norway, or Sweden. As such, each data item in the NTS corpus features a spatial context, i.e., the location from where the post on the platform was published, as well as a temporal context, i.e., the time when the post on the platform was published. Naturally, besides spatial and temporal data variables, each data item contains various other data variables, such as the text of the post, allowing for the analysis of phenomena related to language variability with respect to space and time from a sociolinguistic perspective. Other typical examples for spatio-temporal datasets include weather and climate data (Neset et al., 2019; Lundblad et al., 2010), forestry data (Andrienko and Andrienko, 2006, Chapter 2.3.2), and movement trajectory data (Büschel et al., 2021; Wagner Filho et al., 2020), to name just a few.

Adopted Terminology Examining the terminology described throughout this section, it becomes apparent that some terms are applied interchangeably in the literature. Within the scope of this thesis, the following terminology is adopted for general coherence. A *data item* is a collection of various *data variable–data variable value* pairs, among others featuring data variables that are related to *spatial* and *temporal* contexts. The collection of various such data items represents a *multivariate dataset*, and more specifically a *spatio-temporal dataset*. Additionally, the term *data entity* is adopted to specifically refer to the visual representation of a data item within the context of a visualization environment, such as an immersive VE or a non-immersive desktop terminal.

5.2 Data Analysis Tasks

Naturally, each visualization should be designed to serve a specific purpose and to accommodate the analyst with extracting insights and information by completing desired tasks. Aigner et al. (2011, Chapter 1.1) summarize considerations for the design of information visualizations on a high level with respect to (1) what kind of data are visualized, (2) why are the data visualized, and (3) how are the data going to be visualized. From a user-centered perspective, the specification of the analyst's tasks when interacting with a visualization is particularly interesting, i.e., with respect to why the data are visualized and what purpose does the visualization serve the analyst. Ward et al. (2015, Chapter 1.8) and Aigner et al. (2011, Chapter 1.1) differentiate between three main purposes for the interaction with visualizations:

- *Exploration* or *Explorative Analysis*: The analyst utilizes the visualization and its interactive features to explore an unknown dataset, extract first insights and relevant information with no hypotheses given (undirected search).
- *Confirmation* or *Confirmative Analysis*: The analyst utilizes the visualization and its interactive features to confirm or reject given hypotheses about a dataset (directed search).
- *Presentation of Analysis Results*: The analyst utilizes the visualization and its interactive features to convey and present their findings in the dataset, such as concepts or facts, to an audience.

With respect to the actual design of a visualization's interactive capabilities, Shneiderman's (1996) Visual Information-Seeking Mantra of *overview first, zoom and filter, then details-on-demand* is arguably one of the most famous design guidelines. Based on it, Shneiderman (1996) proposes seven abstract task types that should be supported by the visualization, namely *overview*, *zoom*, *filter*, *details-on-demand*, *relate*, *history*, and *extract*. Another approach by Munzner (2014, Chapters 1–3), Brehmer and Munzner (2013) respectively, describes abstract visualization tasks as a multi-level typology, organizing tasks as to *why* and *how* they are performed as well as *what* a task's input and output parameters are. With respect to why, Munzner (2014, Chapter 3) classifies user actions across four overall groups, i.e., (1) *analyze* (*discover*, *present*, *enjoy)*, (2) *produce* (*annotate*, *record*, *derive)*, (3) *search* (*lookup*, *locate*, *browse*, *explore)*, and (4) *query* (*identify*, *compare*, *summarize*). Depending on scenario and context of the interactive visualization, all these classifications have the potential to be informative for the development, either in isolation or as a mixed and multimodal approach. This allows for guidance and facilitation of the design process towards purposeful interactions with a visualization, and thus with data. Yi et al. (2007) reviewed a multitude of Information Visualization (InfoVis) taxonomies with respect to their described interaction techniques. Based on their analysis of the literature, they synthesized a set formal categories (*select*, *explore*, *reconfigure*, *encode*, *abstract/elaborate*, *filter*, *connect*, *undo/redo*, *change configuration*) to describe a user's intent for the interaction with a visualization in general (Yi et al., 2007). Aigner et al. (2011, Chapter 5.1) further build upon these categories and adapt them to support the more specific context of interacting with *time-oriented* data, i.e., multivariate data where each data item features at least one data variable related to a temporal context. The utilization of such task categories allows to conceptually categorize the interactive features of a developed data analysis tool, similar as presented by Büschel et al. (2018), and thus aiding the tool's description accordingly.

Adopted Terminology Based on the combined work presented by Yi et al. (2007) and Aigner et al. (2011, Chapter 5.1), their data analysis task categories are adopted within the scope of this thesis towards the contexts of IA and the interaction with spatio-temporal data in VEs as follows:

1. *Select – Mark something as interesting*: Select a data entity at a specific spatial location in the VE <u>or</u> modify the displayed temporal context through the selection of a new time event or time range, for instance with the objective to perform various follow-up interactions, such as to display details-on-demand.

2. *Explore – Show me something else*: Look around in the VE with the objective to identify a location/region (spatial) or time event/range (temporal) of interest worth of further inspection <u>or</u> move around in the VE in order to reach data entities, either in close proximity or faraway (outside the physical real-world boundaries of the VR system's calibrated safe interaction area), potentially utilizing virtual travel features.

3. *Reconfigure – Show me a different arrangement*: Perform an interaction that modifies the visual arrangement of the displayed data entities in the VE, for instance with respect to their relative location in the VE or in regard to aspects of their individual visual representation (for instance, sorting the order of the displayed data variables).
4. *Encode – Show me a different representation*: Modify the visualization technique used to represent a data entity in the VE, i.e., mapping a data item's data variables onto a new visual representation and in turn creating a different data entity.
5. *Abstract/Elaborate – Show me more or less detail*: Aligned with Shneiderman's (1996) Visual Information-Seeking Mantra, display details-on-demand (elaborate) to show additional information about a selected data entity, or hide the details (abstract) to enable a more overview-like perspective and interaction mode.
6. *Filter – Show me something conditionally*: Perform an interaction that modifies the visual representation of one or more data entities in the VE to conditionally hide or add information, for instance by deactivating entire data entities or aspects of their individual visual representation (for instance, filtering out undesired displayed data variables).
7. *Connect – Show me related items*: Perform an interaction in the VE that facilitates the inference of relation between and the comparison of data entities, both with respect to spatial and temporal contexts.
8. *Undo/Redo – Let me go to where I have already been*: With respect to the interaction in the VE in general, enable the user to retrace their previous interactions, for instance through undo, redo, history, or reset functionalities.
9. *Change Configuration – Let me adjust the interface*: Perform an interaction that modifies aspects of the user interface on a system level in general or with respect to the particular in-situ interaction mode with one or multiple selected data entities (for instance, temporally accessing and switching between menus and widgets that assist with the interaction in the VE).

5.3 VE Iteration 1: Data Analysis Using Spheres

In order to begin with the empirical investigation in regard to the interaction with data in an immersive VE, and based on the defined requirements and system architecture presented throughout Chapter 4, a first prototype system was developed. This first iteration focuses on the initial design and implementation of an immersive VE that can be utilized for data exploration through VR with a head-mounted display (HMD) device and three-dimensional (3D) gestural input. As such, the system serves as a general proof-of-concept validation with three

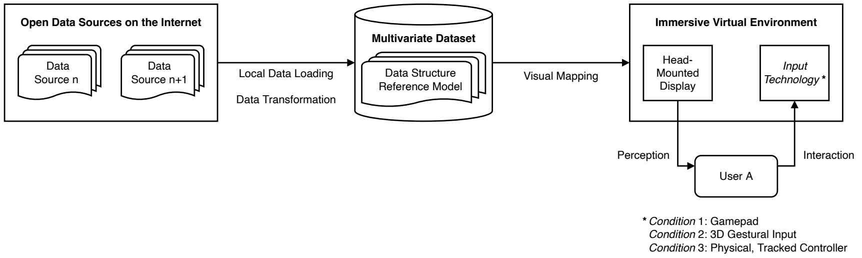


Figure 5.2: Conceptual system overview of the first VE iteration. Detailed descriptions about the various system components are provided in Section 4.2.

objectives, namely (1) validating the overall technological feasibility, (2) validating the users' ability to make use of the immersive VE to explore data and solve related tasks, and (3) investigating the suitability of 3D gestural input and mid-air interaction compared to other input modalities, such as physical controllers. Additionally, the development of this initial immersive data exploration system served as an opportunity to obtain first insights and experiences with respect to its conceptual design and practical implementation. Figure 5.2 provides a conceptual system overview of the developed first VE iteration.

Dataset: 2016 US presidential election The data exploration scenario in this first VE iteration is centered around analyzing the results of the 2016 United States (US) presidential election on a per federal state/district basis. For that purpose, a custom multivariate dataset was created based on relevant data aggregated from various online sources. In particular, (1) the actual voting results were collected from the reporting by The New York Times, (2) overall descriptions and images about the individual states in the US were aggregated from Wikimedia and DBpedia, and (3) Wolfram Alpha was queried in order to extract the geolocation coordinates (latitude and longitude) for each of the state capitals in the US. The dataset served multiple purposes within the scope of this VE iteration. First, it allowed for the conceptual illustration of utilizing the VE as an interface for data from multiple sources. Second, while it is specific enough to create data analysis tasks, it is also generalizable with respect to the overall scenario, i.e., it is conceivable that similar data about other contexts may be aggregated and subsequently explored in the VE. And third, the overall scenario was deemed generally approachable, allowing for inclusive participant recruitment with respect to the empirical evaluation as no specific expertise or domain knowledge is required to explore and make sense of the dataset.

5.3.1 Visualization Design and VE Composition

The overall design of the virtual 3D space and the design of a data entity in the VE, i.e., the visual representation of a multivariate data item, are held intentionally minimalistic in order to direct the user's focus to the data rather than distracting them with unnecessary visual elements in the VE. Similar to immersive data visualization approaches as presented by, among others, Donalek et al. (2014), Wagner Filho et al. (2018), and Pirch et al. (2021), an individual data entity is visualized in the format of a *Sphere* with a complementary textual *name label* anchored directly above it. Figure 5.3 conceptually illustrates the presented data entity visualization design as spheres in the VE. The placement of each data entity in the 3D space of the VE is based on the prior visual mapping process, and as such dependent on the dataset. In case of the presented US election dataset, each data entity is placed with respect to their geolocation coordinates (x- and z-axis) at a fixed height (y-axis) in the VE. Naturally, different placement and arrangement mappings may be applied based on the purpose of the immersive data visualization environment.

To display more information about a data entity (REQ 9 in Table 4.1), the VE features capabilities to toggle, i.e., display and hide, a composite of three *information panels*, similar to a three-monitor desktop setup in the real world. The composition of these panels is conceptually similar to the approach described by Wirth et al. (2018), who utilized various two-dimensional (2D) planes in the *content zone* of their VE to display additional information related to a 3D model. Based on the dataset and thus the available data variables for each data item, each of the panels can be populated with the respective data variable values. For instance, in the case of the presented US election dataset, the three panels are arranged as follows. The *center panel* displays the name of a state as well as a textual description about it. The *right panel* displays various images about the state, while the *left panel* is utilized to display additional information in a listing format, among others the voting results in percent.

Furthermore, the VE is composed as an open space with no visual walls or ceiling. However, rather than letting the user visually "float" in mid-air in the VE, which would arguably feel rather unnatural in comparison to the real world, a virtual floor is placed in the VE. This floor is calibrated and aligned with the position of the real-world floor with respect to the VR system's calibrated safe interaction area. With the intention to facilitate the immersed user's spatial understanding, the virtual floor displays a grid that is composed of both solid and dashed lines. In particular, a three-by-three grid of solid lines represents the user's safe interaction area in the VE, i.e., the area in which they can move freely without obstacles in correspondence to the physical real-world boundaries. Additionally, the virtual floor is extended far beyond the user's physical reach using a grid of dashed lines, providing the impression of a horizon in the VE as important visual orientation cue for the immersed user.

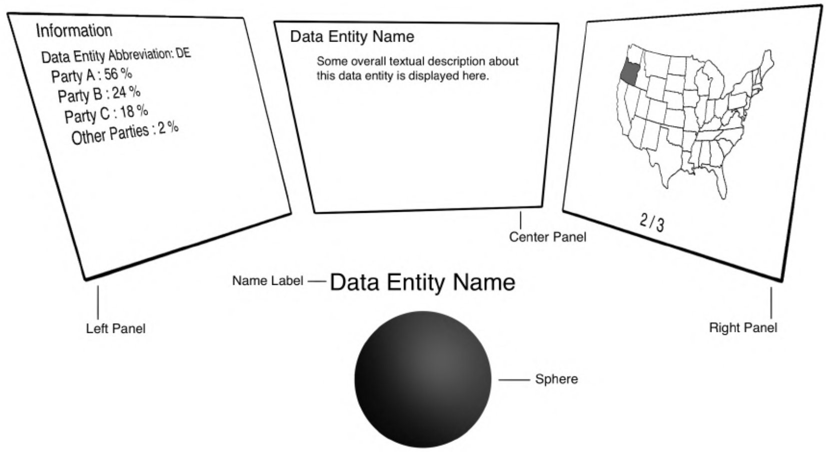


Figure 5.3: *Sphere* visualization design of the first VE iteration.

5.3.2 Interaction Design

In alignment with the data entity visualization design (sphere) and the overall composition of the VE as described in Section 5.3.1, various interactive features are available to the user to support typical data analysis tasks in the VE. Table 5.1 provides an overview of all implemented features, their analysis task classification, as well as how these features are mapped onto the three different input modalities (*Gamepad*, *3D Gestural Input*, and *Physical, Tracked Controller*) with respect to the subsequent empirical evaluation (see Section 5.3.3). Figures 5.4 and 5.5 present some impressions of the interactive features in the VE for the 3D gestural input condition. The remainder of this section provides additional descriptions about the implemented features.

Travel and Selection Rather than allowing the user to arbitrarily transition to any position in the VE, for instance utilizing mechanisms as presented by Pirch et al. (2021) or Streppel et al. (2018), the designed *travel* feature is centered around the approach of *target-based travel* as described, among others, by Lai et al. (2021), Medeiros et al. (2016), and Ragan et al. (2012). The user can indicate a data entity in the VE they would like to travel to, and initiate an automatic transition in linear motion that virtually moves the user to the respective data entity. More specifically, the user's virtual safe interaction area, naturally with the user themself in it, is transitioned in such a way that the targeted data entity is positioned in its center at the end of the virtual transition (REQ 7 and REQ 8 in Table 4.1). This mechanism serves two purposes. First, it ensures that the user can freely move around the targeted data entity in the VE. And second, restricting the user to travel via such target-based travel also explicitly links the user's location to a respective data entity at all times, allowing to easily identify the current in-situ spatial data context (REQ 12 in Table 4.1). Any targeted and visited data entity is also automatically *selected*. The selected data entity, i.e., sphere, features a dark blue color, while all other unselected spheres are displayed in white.

Information Panels Once the user desires to inquire further information about a selected data entity, similar to the details-on-demand concept described by Shneiderman (1996), they can *toggle* the display of a composite of three semitransparent information panels (REQ 9 in Table 4.1). The position and rotation of these panels when toggled for display are derived from the user's head position and rotation in the VE, appearing and aligning conveniently with the user's in-situ field of view. Once toggled and until dismissed, the information panels are statically anchored in the VE, enabling the user to inspect and possibly interact with the panels accordingly. For instance, the user may interact with the right panel in order to *browse* through all associated *information panel images* on a step-by-step basis. Furthermore, while the information panels are displayed, travel and select operations for other data entities are temporarily disabled with the intention to allow the user to remain in their current data analysis context.

Filter In order to provide an element of exploration guidance during the data analysis activity in the VE, a *filter* mechanism is provided that allows for comparison based on the available data variables of the data items in the dataset (REQ 14 in Table 4.1). In particular, the user can compare the currently selected data entity to all others on a *is greater than* and *is less than* basis. The filter results are intentionally limited to indicate only those data entities that are (1) minimal, (2) medial, and (3) maximal different. This concept accommodates to not overwhelm the user with feedback to all, possibly many other, data entities in the VE. The data entities included in a filter result set are highlighted, visually connected to the user's currently selected data entity, and color-coded: green indicating a minimal, yellow a medial, and red a maximal difference. Additionally, travel and selection operations are restricted exclusively to those data entities that are included in the filter result set. In case of the presented US election dataset, such filter operations are available for comparison of the voting results along the Democratic and Republican party as well as the combined other parties.

Bookmark With the intention to enable a user in the VE to keep track of an interesting data entity, while not concerning about potentially forgetting about it or not finding back to it during the data analysis activity, a *bookmark* feature is provided. In particular, the user can mark one data entity at a time as bookmarked. A data entity with the bookmark status appears visually in the format of a large cylinder, conceptually similar to a spotlight, for instance as illustrated by Peter et al. (2018). This visual design should allow the user to easily identify and thus find back to the bookmarked data entity independent of their location in the VE.

Feature	**Gamepad**	**3D Gestural Input**	**Physical, Tracked Controller**	**Analysis Task**
Travel (incl. Select)	Look at a Sphere until it is visually highlighted (color light blue, size scaled up), and press the A button.	Look at a Sphere until it is visually highlighted (color light blue, size scaled up), and point towards it (left/right hand index finger pointing forward).	Use the Pointer Tool to highlight a Sphere (color light blue, size scaled up), and press the Trigger button.	Explore, Select
Select	Utilize the Travel feature (see above).	Utilize the Travel feature (see above).	Touch a Sphere in close proximity (color change to light blue as visual feedback), and press the Trigger button to select the Sphere. For faraway Spheres, utilize the Travel feature (see above).	Select
Toggle Information Panels	With a Sphere selected (color dark blue), press the Y button.	With a Sphere selected (color dark blue), make a “thumbs up” posture (thumb pointing upward, fingers not extended) with the left hand.	With a Sphere selected (color dark blue), press the Grip button.	Abstract/Elaborate
Browse Information Panel Images	While looking at the image in the right panel, press the A button.	Right hand index finger point left or left hand index finger point right (thumb and other fingers not extended).	Touch the image or utilize Pointer Tool.	Reconfigure
Toggle Filter Menu	Press the B button.	Make a “thumbs up” posture (thumb pointing upward, fingers not extended) with the right hand.	Press the Menu button.	Change Configuration
Apply Filter Option	Utilize the D-Pad to navigate to a filter option button in the menu, and press the A button to confirm.	Touch the filter option button.	Touch the filter option button, and press the Trigger button to confirm.	Filter
Set/Remove Bookmark	Look at a Sphere until it is visually highlighted (color light blue, size scaled up), and press the X button.	Look at a Sphere until it is visually highlighted (color light blue, size scaled up), and hold the left or right hand in a “stop”-like posture (hand palm facing outwards, thumb and fingers extended).	Touch a Sphere in close proximity (color change to light blue as visual feedback), and press the Grip button.	Select, Redo
Pointer Tool	n/a	n/a	Press and hold the Trackpad for a vector-based pointing technique (via ray casting), similar to a laser pointer. This is a helper utility, available only for the Physical, Tracked Controller.	n/a

Table 5.1: Overview of the interaction features available in the first VE iteration, including descriptions on how to perform them via the three implemented input modalities (conditions) and their respective data analysis task classifications.

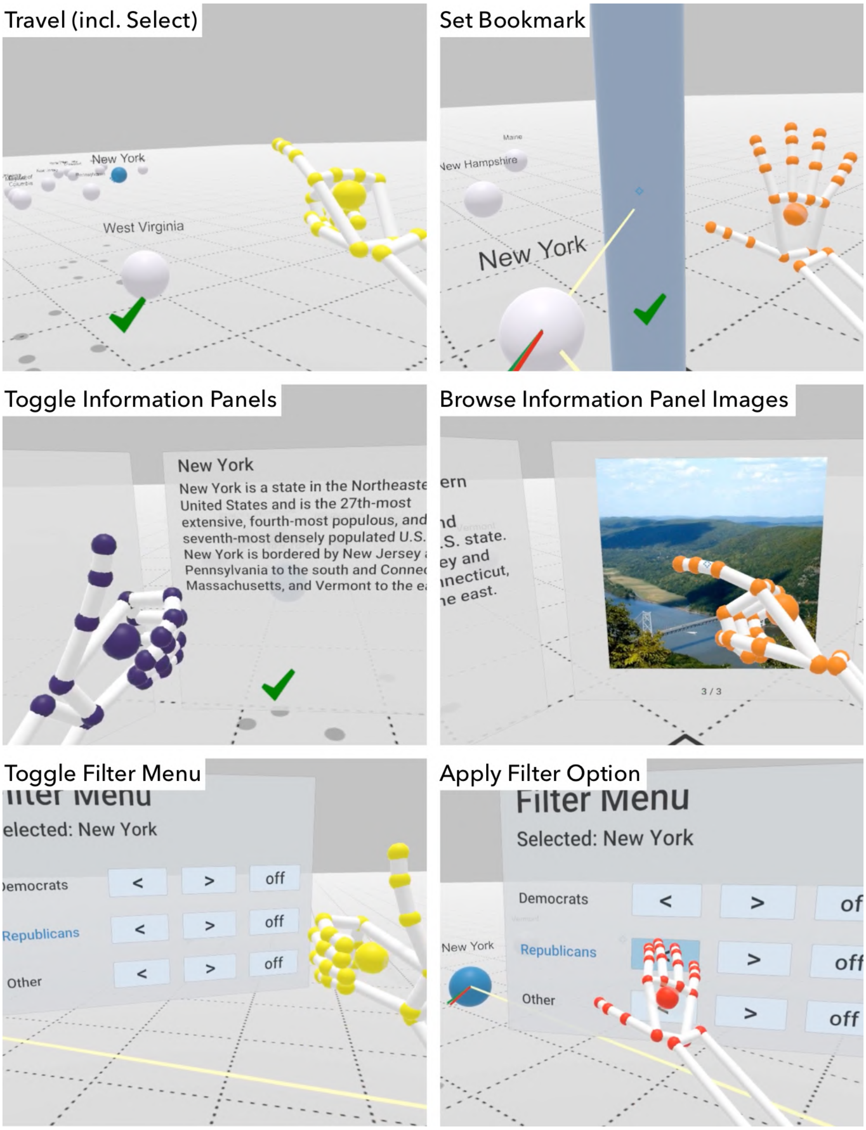


Figure 5.4: Impressions of the various features (see Table 5.1) available in the first VE iteration, from the immersed user's field of view, and utilizing the 3D gestural input condition.

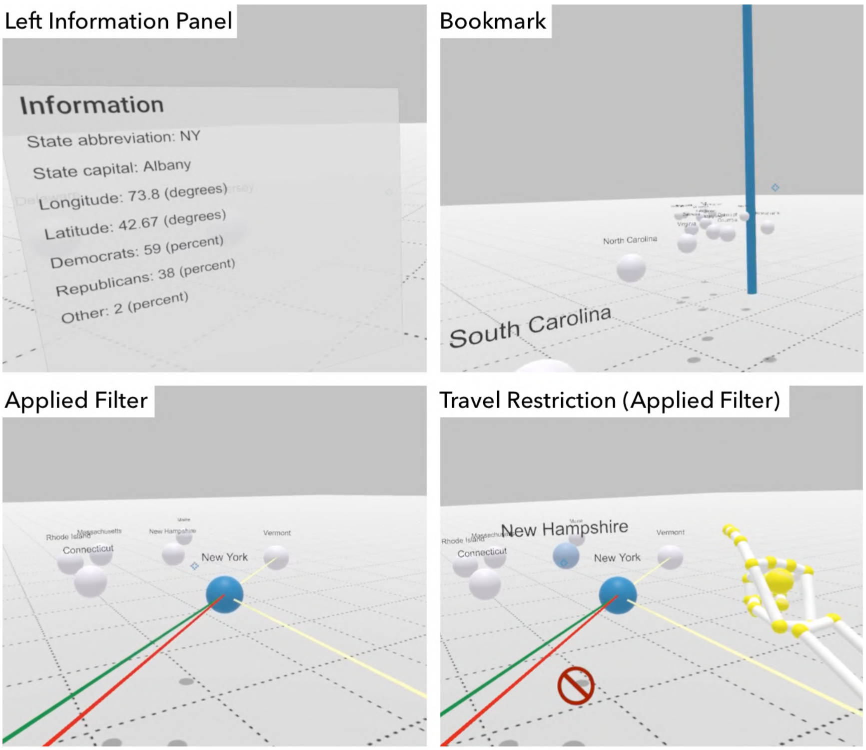


Figure 5.5: Additional impressions of the first VE iteration and the available features (see Table 5.1 and Figure 5.4), from the immersed user's field of view.

5.3.3 Evaluation: Input Technology Comparison

The empirical evaluation of the presented data entity visualization and interaction design of this first VE iteration, in the format of an exploratory user interaction study, was designed with two overall objectives in mind. First, the evaluation should validate the user's ability to explore data and solve related tasks in the VE. And second, the evaluation should investigate the suitability of 3D gestural input in VEs within the scope of the presented IA context, especially compared to other input technologies. To accommodate these objectives and in order to gain first insights and experiences, the evaluation was centered around the comparison of three different input technologies that enable the immersed user to interact in the VE, posing the following research question and hypotheses:

RQ: How do different input technologies (gamepad; 3D gestural input; physical, tracked controller) affect user experience and behavior in VEs using current state-of-the-art VR technologies?

H1: A visual representation tied to the input technology in the VE will have a positive impact on user experience and behavior.

H2: A physical controller tied to the input technology in the VE will have a negative impact on user experience and behavior.

The three chosen input technologies were selected as common and representative modalities for the interaction in VEs that utilize a VR approach. Table 5.2 provides an overview about the key characteristics of these three conditions and how they differ from each other. A detailed overview of how the interactive features in the VE have been mapped to these technologies has been presented in Table 5.1. The study design of the empirical evaluation features three *independent* variables, i.e., gamepad, 3D gestural input, and physical, tracked controller. Measurable differences with respect to experienced workload, perceived flow of interaction, and VR sickness (*dependent* variables) were expected as a result of the participants interacting in the VE using the different input technologies. Within the scope of the presented IA scenario and the developed VE, using these measurements allows for the identification of potential differences in user experience and behavior between these technologies as well as potential advantages and disadvantages of an input technology over another. A *between-group* design was applied, with each participant using one of the three input technologies. In order to have the same number of participants per input technology, the different conditions were cycled with respect to the scheduled study sessions. Individual study sessions were conducted in an one-on-one scenario between one participant and one researcher at a time. The conduction of one session was aimed to a duration of approximately 45 to 60 minutes, whereof the participant would spend approximately 20 to 30 minutes immersed in the VE, wearing the designated HMD. All study sessions were conducted at the VRxAR Labs research group lab at Linnæus University.

Input Technology Characteristic	Gamepad	3D Gestural Input	Physical, Tracked Controller
Visual representation (in VE)	No	Yes	Yes
Physical controller	Yes	No	Yes
Sensor type	Active	Passive	Active and passive
Input device data frequency	Discrete	Continuous	Discrete and continuous

Table 5.2: Overview of the input technology characteristics for the three conditions in the empirical evaluation of the first VE iteration.

5.3.3.1 Physical Study Space

The physical study space that was utilized for the conduction of the empirical evaluation features a designated square two-by-two meter area with a visual three-by-three grid on the floor for the immersed user to move freely without obstacles (see Figure 5.6). It is noteworthy, that the visual three-by-three grid on the floor of the physical study space is aligned with the representative three-by-three grid on the floor in the VE as described in Section 5.3.1. Besides a desk and several chairs provided for the participants, for instance to complete user consent and questionnaires via pen and paper, the physical study space also features a dedicated workstation for the researcher. From their workstation, the researcher can moderate the study, collect data, and monitor the involved hardware and software setup to ensure that all components work as intended. Generally, the researcher remained at their workstation throughout the study session. Aligned with the overall study procedure (see Section 5.3.3.5), the participant was seated twice at their desk, i.e., before and after their interaction in the developed VE.

With respect to the three applied input technology conditions, some additional descriptions are required. First, the gamepad interaction was supported through the utilization of a Xbox One controller[1] and an Oculus Rift CV HMD. Second, the 3D gestural input interaction was supported through a Leap Motion Controller that was attached in front of an Oculus Rift CV HMD. And third, the physical, tracked controller interaction was supported through the utilization of the HTC Vive system that consists of the respective HMD and a physical, tracked controller. Figure 5.7 shows all three input technology devices. Furthermore, the physical, tracked controller condition featured a room-scale VR setup that utilized the entire two-by-two meter physical area, enabling the user to move freely in that area. Even though the users were situated in the designated 2-by-2 meter in the gamepad and 3D gestural input conditions, their setup was generally based on a

[1] iFixit. Xbox One Wireless Controller (Model 1697). Retrieved June 1, 2022, from `https://www.ifixit.com/Device/Xbox_One_Wireless_Controller_1697`

Figure 5.6: A photo of the VRxAR Labs research group lab at Linnæus University. The floor features a three-by-three grid in a physical two-by-two meter area, representing the VR system's calibrated safe interaction area, and thus enabling the immersed user to move freely without obstacles. This physical study space was utilized throughout all empirical evaluations presented within this thesis.

rather stationary VR experience that utilized two external tracking sensors and the Oculus Rift CV HMD.[2]

5.3.3.2 VE Setup

Generally, the VE was setup as described throughout Sections 5.3.1 and 5.3.2. In particular, under utilization of the presented 2016 US presidential election dataset, the VE featured a total of 51 data entities as spheres, representing the 50 federal states as well as the one federal district in the US. All interactive features were available to the immersed user with respect to their assigned input technology for the study session.

5.3.3.3 Task

Utilizing one of the assigned input technologies, each participant was asked to complete a single task, the same for all participants, i.e., to explore the 2016 US presidential election dataset using the provided technologies and identify two data items where both the Democratic and Republican party voting results were close to 50 percent, indicating a tight election race. The data entity representing

[2]A room-scale VR setup that appropriately supports 360° tracking with the Oculus Rift CV HMD requires three external tracking sensors.

Figure 5.7: A photo of the input technology devices for the three conditions in the empirical evaluation of the first VE iteration. **Left**: Xbox One controller (gamepad). **Center**: Leap Motion Controller (3D gestural input). **Right**: HTC Vive Controller (physical, tracked controller).

the state of Alabama was chosen as the starting point in every study session. Each participant was encouraged to utilize all the interactive features at their disposal to freely explore the data entities in the VE with their own strategy and pace. To complete the task, each participant had to name (spoken aloud to the researcher) at one point in time the two data items they assessed fit the task criteria. A total of ten answers were identified that satisfy the task criteria within a reasonable margin (see Table 5.3). This overall task design with no precise answer available was chosen to support the overall inductive and exploratory state of the developed VE at this stage, enabling the participants to interact with the data in the VE in a meaningful way, and allowing for the data collection with respect to the different input technologies.

5.3.3.4 Measures

To obtain a general understanding about the participants, a custom pre-task questionnaire was applied to inquire some demographic information in regard to their self-assessed prior experiences with VR technologies in general, and their self-assessed prior experiences with the assigned input technology. Within the scope of this first VE iteration, the evaluation was centered around the utilization of three subjective methods. First, the Simulator Sickness Questionnaire (SSQ) was utilized as a tool to inquire assessments about the users' perceived VR

Federal State	Democratic %	Republican %	Distance to 50
Florida	48	49	3
Pennsylvania	48	48	4
North Carolina	47	51	4
New Hampshire	48	47	5
Michigan	47	48	5
Wisconsin	47	48	5
Georgia	46	51	5
Arizona	45	50	5
Nevada	48	46	6
Virginia	50	44	6

Table 5.3: Acceptable task answers ranked according to *close match*, i.e., the *combined* distance of each percentage value to 50.

sickness, providing helpful complementary insights regarding the general design of the immersive VE. Second, the Task Load Index (TLX) was applied with the objective to obtain an understanding about the users' cognitive workload during the immersive data analysis activity in the VE. And third, the Flow Short Scale (FKS) was used to measure the users' flow of interaction with the intent to reveal similarities and differences between the three input technologies. Additionally, all user interactions in the VE were logged directly by the system itself, allowing for task performance analysis in general. The researcher also made observations during the users' task solving activity, taking notes accordingly. Finally, an informal interview with each participant was conducted, encouraging them to state additional noteworthy feedback themselves as well as enabling the researcher to pose questions based on the prior made observations. Foundational aspects of these evaluation methods are described in Section 2.5.1.

5.3.3.5 Study Procedure

Each individual study session followed the same procedure of five stages:

1. Introduction (10 min);
2. Warm-up (5 min in the VE);
3. Task (20 min in the VE);
4. Questionnaires (15 min);
5. Interview (5 min).

Although there was no explicit time limitation from the researcher's side, a duration of approximately 45 to 60 minutes per study session was anticipated. During the *introduction*, each participant was welcomed and asked to complete an informed user consent with regard to their participation in the evaluation and the custom pre-task questionnaire. Thereafter, they were introduced to the developed immersive data analysis environment. For this purpose, each participant was shown a short pre-recorded introduction video with regard to their assigned input technology, presenting all the features, and providing all participants the same briefing beforehand.[3] After watching the VE introduction, each participant was given a short *warm-up* time to allow them to familiarize themselves with the composition of the VE and the provided interactive features as well as to become comfortable wearing the HMD and utilizing their respective input technologies. Once the participant felt ready, the researcher initiated the *task* stage as described in Section 5.3.3.3. To prevent a potential insights transfer from warm-up to task stage, different datasets were used, i.e., the task stage featured the setup as described in Section 5.3.3.2, while the warm-up stage featured a dataset with representative dummy data. During the task stage, the researcher made observations and took notes. Once the participant completed their task in the VE, they were asked to complete the provided *questionnaires*, i.e., in order, SSQ, TLX, and FKS. Finally, the informal *interview* was conducted, after which the participant was thanked and sent off.

5.3.4 Results

5.3.4.1 Participants

A total of $n = 24$ participants were recruited to partake in the evaluation to compare the different input technologies within the scope of the first VE iteration that utilized the sphere data entity visualization design. Thus, based on the between-group design, data from $n = 8$ participants per input technology condition were collected. Figure 5.8 shows various participants during their task completion, immersed and interacting with their assigned input modality, in the physical study space. Table 5.4 summarizes their prior experiences per input technology.

5.3.4.2 Task Assessment

Based on the task design, each participant was asked to provide two answers, i.e., two data items in the format of federal states. Table 5.5 presents a detailed summary of the received answers. With 21 participants, the majority provided two answers that were considered acceptable, and as such appropriate. Three participants (one physical, tracked controller, and two gamepad) provided as one of their answers a data item that was not within the anticipated target results.

[3] Three short introduction videos were prepared, one per input technology, of approximately 3 minutes in length each (see Appendix A).

Figure 5.8: Immersed participants during their task completion in the evaluation of the first VE iteration, wearing a HMD and interacting in the VE using one of the three input modalities. **Left**: Gamepad. **Center**: 3D gestural input. **Right**: Physical, tracked controller.

Input Technology	3D UI		VR	
	No EXP	EXP	No EXP	EXP
Gamepad	1	7	0	8
3D Gestural Input	2	6	2	6
Physical, Tracked Controller	4	4	0	8

Table 5.4: Demographic information about the participants in the evaluation of the first VE iteration with respect to their self-assessed prior experiences (EXP), per input technology condition, both in regard to their assigned input technology as 3D UI and VR in general.

Federal State	**Democratic %**	**Republican %**	**Distance to 50**	**Gamepad**	**3D Gestural Input**	**Physical, Tracked Controller**	**Total**
Florida	48	49	3	6	1	5	12
Pennsylvania	48	48	4	3	4	5	13
North Carolina	47	51	4	0	1	0	1
Michigan	47	48	5	1	2	2	5
New Hampshire	48	47	5	1	1	2	4
Wisconsin	47	48	5	2	2	1	4
Arizona	45	50	5	0	0	0	0
Georgia	46	51	5	0	2	0	2
Nevada	48	46	6	1	1	0	2
Virginia	50	44	6	0	2	0	2
Maine	48	45	7	1	0	1	2
Minnesota	47	45	8	1	0	0	1

Table 5.5: The task answers of the participants in the evaluation of the first VE iteration, ranked according to their combined distance as close match. None of the participants provided *Arizona* as an answer. *Maine* and *Minnesota* are answers that are not considered to be within the top ten expected acceptable answers for this dataset (see Table 5.3).

Input Technology	Factor	MEAN	SD	MIN	MAX	MEDIAN
Gamepad	Nausea	1.50	1.41	0	4	1
	Oculomotor	2.00	2.73	0	7	0.5
	TOTAL	3.50	4.00	0	11	2
3D Gestural Input	Nausea	1.50	0.76	1	3	1
	Oculomotor	2.00	1.77	0	4	2.5
	TOTAL	3.50	2.20	1	7	3.5
Physical, Tracked Controller	Nausea	1.50	2.07	0	6	1
	Oculomotor	2.13	1.64	0	5	2
	TOTAL	3.63	3.58	1	11	2.5

Table 5.6: Results of the SSQ based on the participants' self-assessments in the evaluation of the first VE iteration, per input technology condition, and presented grouped according to nausea, oculomotor, and total based on the alternative score calculations proposed by Bouchard et al. (2007). **Note:** The scale maximums for nausea, oculomotor, and total scores are 27, 21, and 48 respectively.

5.3.4.3 Questionnaires

Table 5.6 presents the collected self-assessments with respect to VR sickness based on the SSQ. Figure 5.9 illustrates the workload results as inquired with the TLX. The collected measurements with regard to the experienced interaction flow using the FKS are presented in Figure 5.10 and Table 5.7.

5.3.4.4 Logging

Using the implemented logging system that was integrated as part of the VE, it is possible to analyze the participants' interactions in the VE with respect to various relevant matters. Although there was no time limit and the participants were encouraged to solve the task at their own pace, it is helpful to inspect the task completion times per input technology as presented in Figure 5.11. Similarly, Figure 5.12 presents the participants' interactions per minute across the three input modalities. Furthermore, it is possible to generate 2D pathway visualizations that illustrate the participants' travel interactions from data entity to data entity in the VE. Independent of the three input technology conditions, the participants followed different travel strategies to complete the task as the examples in Figure 5.13 illustrate. Some participants ($n = 10$) completed the task and named two data items as soon as they encountered two data entities they considered suitable as answers, exploring the different data entities in the VE to a comparatively minimal extent. In contrast, other participants ($n = 6$) explored the data entities to a greater extend, traveling back and forth multiple times between already visited ones.

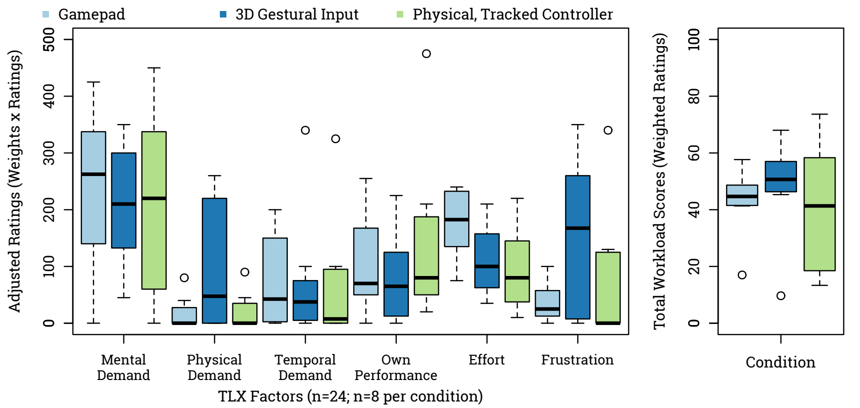


Figure 5.9: Results of the TLX based on the participants' self-assessments in the evaluation of the first VE iteration, per input technology condition. **Left**: Adjusted ratings ($weights \times ratings$) for each of the six TLX factors. **Right**: Calculated total workload scores ($weighted\ ratings$).

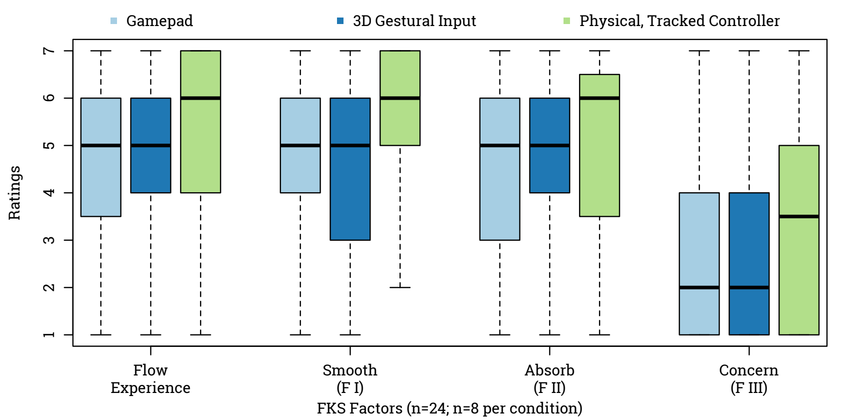


Figure 5.10: Results of the FKS based on the participants' self-assessments in the evaluation of the first VE iteration, per input technology condition, and presented grouped according to the overall flow experience as well as smooth automatized process (F I), ability to absorb (F II) and concern (F III).

Flow Short Scale	**Gamepad** **MEAN (SD)**	**3D Gestural Input** **MEAN (SD)**	**Physical, Tracked Controller** **MEAN (SD)**
F I - Smooth automatized process	**5.00 (0.33)**	**4.54 (0.33)**	**5.38 (0.52)**
8) I knew what I had to do for each step of the way.	4.63 (1.77)	4.88 (1.64)	5.38 (1.92)
7) The right thoughts/movements occur of their own accord.	4.75 (1.49)	4.38 (1.92)	4.38 (1.51)
9) I felt that I had everything under control.	4.88 (1.55)	4.88 (0.99)	5.50 (1.93)
4) I had no difficulty concentrating.	5.00 (1.93)	4.25 (2.49)	5.88 (1.46)
5) My mind is completely clear.	5.50 (1.20)	4.13 (1.73)	5.63 (0.74)
2) My thoughts/actions ran fluidly and smoothly.	5.25 (1.67)	4.75 (1.49)	5.50 (1.69)
F II - Ability to absorb	**4.47 (1.54)**	**4.94 (0.88)**	**4.91 (1.61)**
6) I was totally absorbed in what I was doing.	6.25 (0.71)	6.13 (1.13)	6.25 (1.04)
1) I felt the right amount of challenge.	4.50 (1.20)	4.75 (1.39)	4.63 (1.69)
10) I was completely lost in thought.	2.50 (1.20)	4.00 (1.60)	2.75 (1.49)
3) I did not notice time passing.	4.63 (1.85)	4.88 (1.13)	6.00 (1.31)
Flow experience (1-10)	**4.79 (0.96)**	**4.70 (0.60)**	**5.19 (1.03)**
F III - Concern	**2.54 (0.76)**	**2.92 (0.40)**	**3.17 (0.97)**
11) Something important to me was at stake here.	2.38 (1.41)	2.63 (2.26)	2.88 (1.81)
12) I did not make any mistake here.	3.38 (1.69)	2.75 (1.98)	4.25 (1.83)
13) I was worried about failing.	1.88 (1.64)	3.38 (2.39)	2.38 (1.92)

Table 5.7: Results of the FKS based on the $n = 24$ participants' self-assessments in the evaluation of the first VE iteration, per input technology condition ($n = 8$) as mean and standard deviation, and presented grouped according to the individual FKS factors and items, i.e., smooth automatized process (F I), ability to absorb (F II), flow experience, and concern (F III).

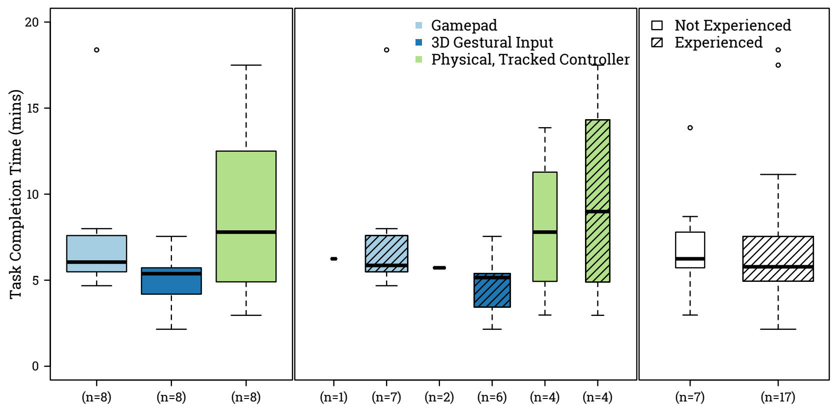


Figure 5.11: Task completion times of the $n = 24$ participants in the evaluation of the first VE iteration, per input technology condition ($n = 8$), and with respect to their self-assessed prior experiences with the assigned input technology.

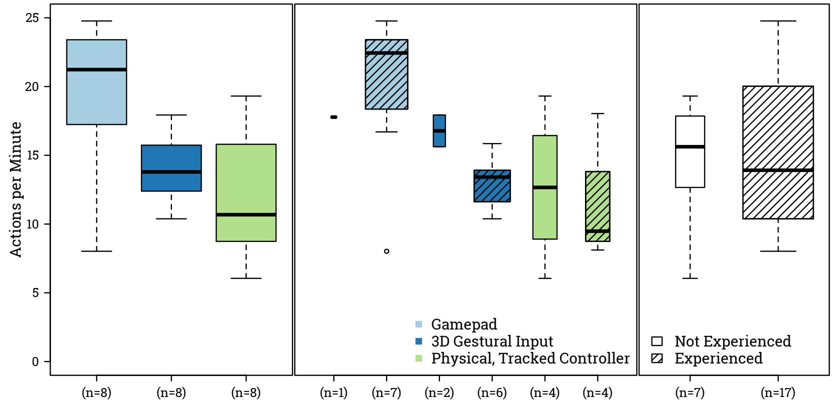


Figure 5.12: Interactions per minute of the $n = 24$ participants in the evaluation of the first VE iteration, per input technology condition ($n = 8$), and with respect to their self-assessed prior experiences with the assigned input technology.

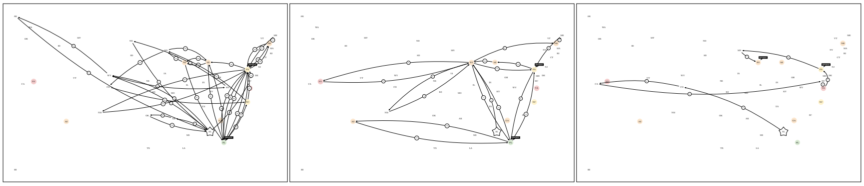


Figure 5.13: Examples of different data exploration strategies based on a participant's travel interactions in the VE, observed across all input technology conditions, compiled as pathway visualizations. **Left**: Highly complex and systematic. **Center**: Travel loops of lower complexity. **Right**: Straight walk. **Note**: An overview of all pathway visualizations is presented in Appendix C.

5.3.4.5 Observations and Informal Interview

16 participants were observed utilizing the provided filter options in a systematic manner to guide their immersive data analysis with the objective to identify suitable answers. They traveled to a data entity, toggled the information panels to inspect the voting results, and applied a filter option, for instance to find a state that would have slightly more votes for one of the parties, and then travel to that data entity accordingly. In most of these cases, this procedure was either repeated or the participant would travel back if they considered the prior data entity as a more suitable answer. It is noteworthy though that six participants asked for further clarifications about the functioning of the filter options after the initial VE introduction, while five users made only minimal use (if at all) of the filter functionality to complete the task, instead rather following a trial-and-error strategy to find suitable answers. Twelve participants actively emphasized the pleasantness of their experience in the VE, positively highlighting the target-based travel interaction for the automatic transition between different data entities. They were also excited about the various possibilities provided by the VE in general.

The participants also expressed some feedback in regard to further improvements. For instance, ten participants noticed occlusion issues, i.e., data entities in more densely populated spaces in the VE interfering with the ability to view the contents as displayed in the information panels. Five participants requested an additional feature to assist with their orientation in the VE, asking for some kind of *map* or *birds-eye view* in order to get an overview about all the data entities as well as their own position, for instance by having a more traditional view on the data entities from the top down. Four participants also suggested to "attach" information that is textually displayed in the information panels directly to the data entity spheres themselves, thus visually encoding data directly. They argued that this could prevent a frequent toggle of the information panels when searching for data entities that contain specific data variable values. Furthermore, five participants (3 gamepad, and 2 physical, tracked controller) had active troubles remembering the feature-to-button mapping on the respective physical controllers. They argued that the mapping did not feel intuitive, thus demanding mental efforts to learn and remember what button triggered what feature. The data entity selection through gaze-based input in the gamepad and 3D gestural input conditions was physically noticeable according to four participants, especially during attempts to select faraway spheres.

5.3.5 Discussion

As described in the introduction to Section 5.3, the development and evaluation of this first VE iteration is considered a proof-of-concept validation in general, investigating aspects of technical feasibility, the user's ability to successfully explore data, and the suitability of 3D gestural input compared to physical controllers. Based on the presented data entity visualization design, VE composition, interaction design, and the results of the empirical evaluation, it is possible to discuss the findings and obtained experiences so far. These serve as important foundations and impulses that subsequently inform the design of the next VE iterations. The discussion is organized around the examined input technologies, the participants' immersive data exploration, and finally the hypotheses assessments, addressing how visual representation and physicality aspects of the input technologies affected the participants during their time in the VE.

5.3.5.1 Input Technologies

One of the empirical evaluation's main purposes was to investigate the suitability of 3D gestural input, i.e., enabling the user to simply use their hands, for the interaction in an immersive VE within the context of IA. Therefore, 3D gestural input as one input technology was compared to modalities based on a gamepad and a physical, tracked controller. Table 5.2 highlighted the characteristical differences between these three conditions, particularly in regard to their visual representation in the VE and their physicality.

Task Assessment Generally, all three implemented interaction modalities enabled the participants to solve the given data analysis task in a satisfying manner. Based on the pathway visualizations created from the logging data, for instance as shown in Figure 5.13, it seems that every participant followed their own strategy to complete the data analysis task independent of the used input technology. Even though the task solving approach differed from participant to participant, no strong correlation (also under consideration of the sample size) between the input technologies and the participants' ability to solve the task could be identified. Examining Table 5.5, it becomes apparent that compared to gamepad and physical, tracked controller, all 3D gestural input participants provided answers within the predetermined acceptable task answers. An additional examination of the logging data reveals that the 3D gestural input participants had less travel interactions on average, but that did not prevent them from successfully completing the task. However, the participants that utilized the physical, tracked controller provided overall the more highly ranked answers.

VR sickness Based on the results of the SSQ that was administered only post-task to not disrupt the participants' task solving activity, the measured VR sickness was very low overall across all three input technology conditions, with only small variations between them (see Table 5.6). Consequently, no noteworthy differences could be identified, indicating a rather comfortable experience independent of the utilized input technology. In the presented evaluation and context, one can argue that the input technology affected the user's experienced VR sickness only minimal based on each individual's sensory perception in the VE. Furthermore, gamepad and 3D gestural input participants utilized an Oculus Rift CV1 HMD, while the physical, tracked controller participants used the corresponding HTC Vive HMD for their visual immersion in the VE (see Section 5.3.3.1). Naturally, the choice of display technology can impact the results of the SSQ. However, given the overall low scores and comparatively identical device specifications (see Table 2.1 in Chapter 2), it is hard to tell if either input technology or HMD had a noticeable impact on the measured VR sickness.

Workload Examining the results of the TLX workload assessments as illustrated in Figure 5.9, it is noticeable that the *mental demand* was reported comparatively high across all input technologies. Arguably, this is an outcome from interacting in the VE with the respective input technology on the one hand, while on the other attempting at the same time to complete the exploratory data analysis task. Although the individual workload assessments are overall rather unique to each participant, it is possible to observe certain trends. For instance, the gamepad condition scored a comparatively high adjusted rating for the *effort* factor. A closer examination of the collected logging data revealed that the participants that used the gamepad had the most contextually wrong interactions in the VE among the three conditions. This indicates that the gamepad participants had a harder time recalling the feature-to-button mapping, also as observed and pointed out by some participants during the informal interview. It is worth considering whether or not the fact that the gamepad controller itself has no visual representation in the VE may have contributed to this circumstance. Furthermore, the reported higher mental demand of the gamepad condition might also relate to the self-reported higher effort in this condition.

Although the medians for the perceived *physical demand* across the three conditions are relatively low in general, the reported higher physical demand by the 3D gestural input participants compared to the other two conditions is likely a result due to the increased hand, finger, and thumb movements to form the respective hand postures for the various gestural commands. Cardoso (2016) reported that interaction using the Leap Motion Controller for 3D gestural input required a considerably higher effort compared to gamepad and gaze-based interaction techniques (measured using the ISO 9241-9 questionnaire) to complete the path following tasks in their study. Their results correspond with the reported high physical demand factor (see Figure 5.9).[4]

The reported *frustration* by the 3D gestural input participants was considerably higher compared to the gamepad and physical, tracked controller conditions. An examination of the logging data shows that the 3D gestural input users had the least interactions in a wrong context on average. They were also all able to solve the task successfully (see Table 5.5), requiring the least amount of time to do so according to the calculated mean completion times (see Figure 5.11). In combination with the researcher's observations, a drawback of the logging system becomes apparent. The logging system only keeps tracked of *detected* interactions, such as pressing a button or in the case of 3D gestural input a detected gestural command. However, gestural command detection is not always reliable (Koutsabasis and Vogiatzidakis, 2019; Bachmann et al., 2018), either due to the sensory tracking or a user not performing the gestural command adequately, potentially still learning the 3D UI. In turn, some *undetected* interactions may have simply not been recorded for the 3D gestural input condition. Quantitative data, for instance based on video recording and subsequent analysis, could provide a better picture about this matter, but was not collected within the scope of this empirical evaluation. Nevertheless, participants utilizing the 3D gestural input were at times observed attempting repeatedly to successfully perform a respective gestural command in the VE. Despite their task success rate and the overall few contextually wrong interactions, having to attempt some gestural commands multiple times likely contributed to the reported higher frustration. Interestingly, despite these higher rated physical demand and frustration factors, the 3D gestural input participants assessed their *own performance* still competitive in comparison to the gamepad and physical, tracked controller conditions.

Interaction Flow Examining the *flow experience* results based on the collected FKS data as presented in Figure 5.10 and Table 5.7, it appears that the participants utilizing the physical, tracked controller felt slightly more "in the flow" compared to the other two conditions. However, the results are only marginally apart from each other. The physical, tracked controller users reported the best flow experience scores when interacting in the VE, which could relate to the lower workload compared to gamepad and 3D gestural input (see Figure 5.9).

3D gestural input users reported a slightly less *smooth automatized process* interacting in the VE. This arguably relates to the previously discussed frustration of having to attempt some gestural commands multiple times, disrupting their flow experience. In regard to the users' *ability to absorb*, the physical, tracked controller condition ranks first, followed by 3D gestural input, and gamepad last.

[4]Note that the effort TLX factor is not directly comparable with "effort" as defined in the ISO 9241-9 questionnaire, i.e., mental and physical (TLX) versus physical only (ISO).

Potentially, the visual representation of the input technology in the VE may have had an impact on these results, and should thus be taken into consideration. Since humans primarily use their visual senses under normal circumstances, the translation of their own movements, visually appropriately presented in the VE as in the cases of the physical, tracked controller and the 3D gestural input conditions, may have affected their perceived feeling of being absorbed in the computer-generated 3D space.

5.3.5.2 Immersive Data Exploration

The participants were tasked to identify two data items according to certain criteria (see Section 5.3.3.3). Independent of their data analysis and exploration strategy, the majority of the participants were able to identify not just one but two suitable answers, while only three of the 24 participants identified one data item that was just outside the range of pre-determined answers (see Section 5.3.4.2). Based on the provided dataset and task, the results indicate that the developed first VE iteration can be used for immersive data analysis and exploration, resembling a unified interface for data from multiple sources. Utilizing the provided features and displayed information in the VE, the participants were able to successfully solve a typical explorative data analysis task related to the dataset. This is particular encouraging in regard to the further development of visualization and interaction mechanisms as well as with respect to the display and exploration of other datasets in such a VE. Considering the overall positive response, enthusiasm, and ideas for future features from the participants in the evaluation, the presented VE iteration appears to be an exciting and fun approach to explore data in a computer-generated 3D virtual space.

It is also noteworthy that none of the participants quit or paused their immersion in the VE during their study session. On the contrary, twelve of the 24 participants actively acknowledged their overall experience as pleasant. This indicates the overall appeal of the VE iteration's visualization and interaction design. Both were, to a certain extent at this stage, held purposefully minimalistic and clear. Yet, the participants were able to interpret and make meaning of the data, as indicated by the successful task completion results. The implemented target-based travel mechanism of the provided travel and selection feature anchored the participants to one specific data entity at all times, preventing them from straying off into the empty 3D space. This has arguably two advantages. First, it focuses the user on the data, in particular one specific data item in the dataset, represented in the format of a sphere with its complementary information panels as details-on-demand. Thus, the user's attention is usually set on that data item, from where they can further explore and take the next steps in their analysis. Second, the target-based travel, implemented as automatic transition in linear motion from data entity to data entity, was perceived as pleasant, demonstrating that this type of movement technique in the VE worked well within the presented context. This is also supported by the overall low experienced VR sickness (see Table 5.6), particularly as the travel feature was one of the most frequently used actions the participants performed in the VE.

The visualization of the participants' pathways was particularly interesting, illustrating their step-by-step travels in the VE, representing their data entity exploration over time to solve the given task (see Figure 5.13). Although it was explained to them that time is not an important factor, encouraging them to analyze the data at their own pace and strategy, different participants ended up following different approaches. As the pathway visualizations clearly illustrate, some participants appeared ambitious to solve the task quickly, and thus only explored minimal aspects of the data. Other participants took more time and efforts to find a suitable solution. Some participants even explored the data to a great extent, appearing very adamant on finding the best possible solution, illustrating the general pleasantness of the VE. At the same time, those of them that explored the data to a more minimal extent seemed to be satisfied with their performance and exploration as well. With no given time limitation for the task completion, the exploratory analysis approaches of the participants is arguably dependent on multiple individual factors, such as their familiarity with the displayed data, their eagerness to gain new insights, or simply their mood, illustrating that users have their individual ambitions and task solving strategies.

Overall, based on the results of the empirical evaluation, the visualization and interaction design as well as the composition of this first VE iteration can be considered satisfactory, allowing a user to explore data utilizing immersive display and interaction technologies, even enabling them to come up and apply their own data analysis strategies. The results and the data-agnostic ability of the VE to display other datasets, as long as the data are transformed according to the respective *Data Structure Reference Model* as introduced in Section 4.2.1, encourage the utilization of the presented immersive data analysis environment in other contexts and scenarios. Under consideration of the presented results and discussion within the scope of this first VE iteration, it is possible to regard (1) the overall technological feasibility of utilizing off-the-shelf hardware and software technologies for the development of an immersive data analysis system using a VR approach as validated, and (2) the users' ability to make use of that system in order to explore data and successfully solve related tasks as validated.

5.3.5.3 Hypotheses Assessment

In regard to the investigation of the overall suitability of 3D gestural input in comparison to other input modalities for the respective interaction in an immersive data analysis environment, the presented empirical evaluation was centered around two hypotheses (see Section 5.3.3). The hypotheses focus on aspects of the input device's visual representation in the VE and its physicality. Based on the evaluation results, the hypotheses can be answered as follows.

H1: A visual representation tied to the input technology in the VE will have a positive impact on user experience and behavior. The physical, tracked controller scored the most positive results in regard to the TLX (lowest workload), arguably followed by the 3D gestural input (high physical demand and frustration, but good own performance), and then gamepad (high mental demand and effort). With respect to the reported workload assessments, this supports the *H1* hypothesis, as both the physical, tracked controller and 3D gestural input conditions feature a visual representation tied to the input technology. The physical, tracked controller arguably scored also better with respect to the interaction flow (FKS). Particularly with an emphasis on the users' ability to absorb and the overall flow experience, the same ranking as with the TLX can be observed, similarly supporting this hypothesis.

H2: A physical controller tied to the input technology in the VE will have a negative impact on user experience and behavior. Considering that 3D gestural input, in both TLX and FKS, ranks between the gamepad and physical, tracked controller conditions, it is difficult to argue for supporting or rejecting this hypothesis. While physical, tracked controller and gamepad participants interacted and explored the displayed data the most during the task, based on the amount of interactions and visited data entities, they also showed more contextually wrong interactions compared to the 3D gestural input participants. Prior experience with the physical controller's button layout and respective feature mapping is needed, as a few participants had noticeable challenges in these aspects. This could indicate an overall less intuitive interaction approach compared to the 3D gestural input condition, which allowed the participants to learn and recall the implemented gestural commands rather easily.

The results of the presented input technology comparison are in line with such findings as for instance presented by Gusai et al. (2017), indicating a slight preference for a physical, tracked controller over 3D gestural input due to its greater stability and accuracy. Aspects of performance and detection issues in regard to the Leap Motion Controller have been reported in the past, favoring the more stable sensory tracking of the HTC Vive system in comparison (Caggianese et al., 2018). Arguably, even though the reported hand posture and gestural command detection of the Leap Motion Controller led to a certain frustration with the participants in the presented VE evaluation, it interestingly seemed not to have led to a drastic decrease in their experienced interaction flow (see Figures 5.9 and 5.10) within the context of an exploratory data analysis task with no time limitations. It is interesting to observe that although users preform measurably better overall under one condition (HTC Vive), they still subjectively prefer another one (Leap Motion) for certain tasks, as reported by Figueiredo et al. (2018). Although presenting mostly qualitative results and subjective impressions, the findings by Streppel et al. (2018) are in line with the results of this first VE iteration, indicating a generally similar acceptance for both physical, tracked controller and 3D gestural input, utilizing the HTC Vive and Leap Motion Controller respectively, for the purpose of interacting in an immersive data analysis environment. Furthermore, one would arguably assume that the 3D gestural input and physical, tracked controller conditions are much more suited for the interaction in the VE due to their 3D spatial tracking capabilities and visual representation. Examining the overall results of all three input technologies, the comparatively close scores of the gamepad condition can be explained due to the general wide acceptance and familiarity with such devices (Lepouras, 2018).

Overall, based on the presented results and discussion, the choice of input technology was not decisive on the immersive data analysis activity considering various aspects such as workload, interaction flow, VR sickness, and task performance. However, some interesting trends could be highlighted throughout the discussion, for instance the indication of a preference in favor of an input device's visual representation in the VE, but not a clear trend towards the utilization of physical controllers within the presented IA context. As such, 3D gestural input, as one of the three conditions, can be considered suitable as an input modality to enable user interaction in immersive data analysis environments. Improved sensory tracking and input interpretation capabilities of 3D gestural input technologies, enabling a more robust hand posture and gesture detection, could arguably minimize experienced user frustration in the VE. At the same time, the implemented gestural commands appeared to be generally suitable for the utilization of the various features in the VE, and were overall swiftly learned by the participants.

5.4 VE Iteration 2: Data Analysis Using Stacked Cuboids

The design, development, and evaluation of the first VE iteration as described throughout Section 5.3 served as an important validation with respect to the presented immersive data analysis concept in general. While the first VE iteration focused on the analysis of a multivariate dataset in general, and among others with respect to spatial data variables, the objective with the design and development of the second VE iteration is to advance the immersive data analysis activity to support temporal contexts. Thus, two aspects are of main interest within the scope of the second VE iteration. First, how can the data entity visualization design be adapted to visually encode temporal data variables at a respective (spatial) location? And second, based on the updated data entity visualization design, what interactive features are required to support and facilitate the interaction with such temporal data visualizations, and thus support the interaction with spatio-temporal datasets in the VE?

Dataset: Nordic Tweet Stream (NTS) The main data context within the scope of the second VE iteration is centered around the Nordic Tweet Stream (NTS) dataset, a dynamic corpus of real-world Twitter messages (Laitinen et al., 2018). Besides the actual content of a *tweet*, usually in format of a text message, various metadata are attached, such as the geolocation from where a tweet was published and a language classifier of the tweet. Over 50 metadata variables for each tweet exist in the NTS, and thus qualifying it as a multivariate dataset. More specifically, the NTS corpus only includes tweets with an identified geolocation that originates in one of the five Nordic countries (Denmark, Finland, Iceland, Norway, and Sweden). As it is an ongoing project, new data items (tweets) are constantly captured and added to the dataset. The partial NTS dataset used within the scope of this thesis consists of overall 11.657.987 tweets, collected within the time period from November 6th, 2016, to February 26th, 2018. The number of tweets per day averages to 26.139. 1.452 unique locations and 188 unique languages (determined by each tweet's `lang` metadata value) have been identified.[5] The coordinates for each location are described as a rectangular bounding box, and the centroid geolocation coordinate as latitude and longitude for each rectangle was calculated.[6] Using these unique coordinates for each location, it is possible to bundle locations into clusters. In particular, using the R package *leaderCluster*,[7] the 1.452 unique locations were compiled into a total of 316 clusters with a radius of 60 km (utilizing the `haversine` distance metric). Through the additional utilization of a lookup table, containing each unique place and its assigned cluster, it is possible to allocate and aggregate the tweets that were published from each place and also represent the location clusters using their centroid latitude and longitude attributes. Due to the diversity of the different metadata attributes of each data item in the NTS dataset, various possibilities for data analysis arise. For instance, the analysis of tweets from a sociolinguistic perspective with respect to geolocation and language over time, i.e., a spatio-temporal context, is of particular relevance and interest to language researchers (Tyrkkö et al., 2021; Laitinen et al., 2018).

5.4.1 Visualization Design and VE Composition

Compared to the overall sphere approach in the first VE iteration, the data entity visualization design of the second VE iteration is centered around the concept and format of a *Stacked Cuboid*, as conceptually illustrated in Figure 5.14. Similar to the visualization approach of a stacked bar chart in 2D, a stacked cuboid can be

[5]The vast majority of tweets originated from the major urban areas through the Nordic region, and only in a small subset of these unique languages.

[6]This was done for all the unique locations, and to accommodate the fact that not every tweet in the corpus contained an exact latitude and longitude coordinate pair.

[7]Taylor B. Arnold. leaderCluster: Leader Clustering Algorithm. Retrieved June 1, 2022, from `https://cran.r-project.org/web/packages/leaderCluster/index.html`

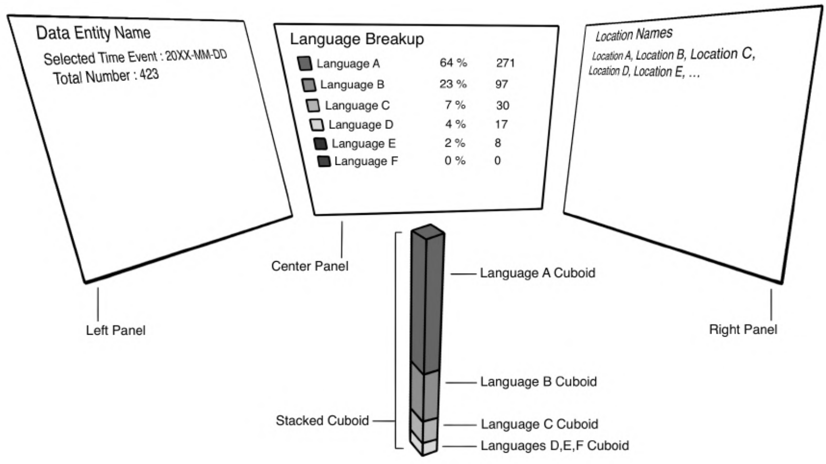


Figure 5.14: *Stacked Cuboid* visualization design of the second VE iteration.

regarded as its extension into the 3D space. It is composed of multiple different cuboids (or layers in 2D), where each cuboid may be associated with, and thus used to visually represent, a comparable data variable of the respective data item. Consequently, the individual cuboids of a stacked cuboid may vary in size and color, providing indications about the mapped data variables that the user can visually observe to detect patterns and data entities of interest that are worth of further exploration. Thus, the data entity itself can be used to make observations and infer first insights without the need to display further details. Within the presented context of spatio-temporal data visualization, stacked cuboids are utilized in the second VE iteration as follows. Similar to the approach in the first VE iteration, each stacked cuboid can be placed in the VE in accordance to a data item's geolocation values. Furthermore, the size and color of each cuboid is configured to represent individual comparable data variables for a given time domain relevant for the dataset. In the practical case of the NTS dataset, each data entity presents a cluster of tweets, where each tweet originated in close proximity to the identified cluster's geolocation on a per week basis. Each data entity consists of four cuboids, where the top three represent, in order, the amount of tweets of the three most frequently used languages, while the fourth one contains the remaining. All the individual cuboids are scaled in height, based on the amount of tweets they represent, and are color-coded, for instance yellow to represent the Swedish language. The overall height of each data entity is scaled logarithmically[8] to reflect the total amount of tweets, and thus providing an indication of the overall Twitter traffic. Thus, data entities with a high amount of Twitter traffic appear as taller stacked cuboids, while those with a low amount appear shorter. By utilizing the described approach to visualize the NTS dataset, the immersed user can obtain a visual impression of the Twitter social network activity according to location and language variability for a determined temporal context by looking around within the virtual 3D space. Due to the uniform color coding of the languages as well as the height scaling of tweet numbers, it is possible, among others, to spot data entities that are visually distinct compared to others, and are thus potentially interesting for further investigation.

To display more detailed information about a data entity in the VE (REQ 9 in Table 4.1), the same approach of utilizing a composite of three *information panels* is used as in the first VE iteration (see Section 5.3.1). In particular, the *center panel* features a detailed listing about the language distribution of the selected data entity, i.e., language name and its assigned color, share in percentage, and number of tweets. The *right panel* displays all unique location names contained within the respective data entity, i.e., the cluster of tweets. The *left panel* provides additional information, also in a listing format, among others, the total number of tweets and the selected time (REQ 11 in Table 4.1).

The second VE iteration features generally the same virtual floor setup, using a grid of solid and dashed lines as described in the first iteration (see Section 5.3.1). However, in alignment with the presented NTS dataset, and to provide visual guidance in the format of navigation and orientation cues to the immersed user as they move around and explore the data, the virtual floor additionally features the five Nordic countries rendered as color-coded extruded surfaces on the floor (REQ 11 in Table 4.1). The individual data entities are virtually placed at their appropriate locations on the virtual floor, enabling the user to get an overview about the data, particularly the more distance ones. However, with ergonomic considerations in mind, it would be rather cumbersome for the user in the VE to physically reach down on the virtual floor in order to select a data entity in close proximity. To address this matter, a mechanism is implemented in the VE that automatically lifts data entities in close proximity to the user's chest height. This mechanism aims to facilitate closer examination and interaction with each data entity, without losing the desired general perspective offered by neighbouring data entities. While a data entity may be raised for such facilitated ergonomic interaction, its shadow is projected on the virtual floor to maintain a visual indication of its exact location. This is particularly important for those that are located in border regions, for instance close to both Norway and Sweden. While one could argue that the list of unique real-world place names in the right

[8] A logarithmic scale has been chosen in order to deal with the wide range of tweet frequencies within the different clusters of this dataset, as more populated areas may show several thousand tweets per day, while clusters in rural areas rarely show more than single- or double-digits.

information panel provides enough indications in regard to whether or not the cluster is located in one country or the other, it is also important to keep novice users in mind. They might be unfamiliar with the dataset or context, and in turn wish to use the VE to learn about the dataset and its characteristics. Therefore, it is arguably important to provide a range of features and visual indications that facilitate the immersive data analysis activity.

5.4.2 Interaction Design

Aligned with the stacked cuboid visualization approach as described in Section 5.4.1, the VE provides several interactive features under utilization of 3D gestural input as exclusive input modality. Table 5.8 provides an overview of these features, their analysis task and interaction technique classifications, as well as describing how the interaction is performed in the VE using the 3D gestural input. Figure 5.15 illustrates the interactive features from the user's field of view in the VE. The remainder of this section provides additional descriptions about the implemented features.

Travel and Select The mechanisms that allow the user to *travel* and *select* data entities in the VE generally follow the same interaction concepts as implemented in the first VE iteration (see Section 5.4.2). In particular, two options are provided in the VE. Using *select-through-touch* the user can approach a data entity in close proximity by walking up to it and then simply touching it (REQ 7 in Table 4.1). Although the immersed user is theoretically able to walk near infinitely within the virtual space, the physical real-world boundaries limit their actual movement, thus requiring an alternative selection technique to select those data entities that are located outside the VR system's calibrated safe interaction area. Using the multimodal *select-through-point*, the user can gaze at a faraway data entity to highlight[9] it, and then, using an analogy of "*I want to go there*", make a *point forward* hand posture to initiate a *target-based travel* transition to the indicated data entity (REQ 8 in Table 4.1). Adopting the gaze-based interaction principle of *gaze suggests, touch confirms* as discussed by Stellmach and Dachselt (2012), this *select-through-point* mechanism implements the principle of *gaze suggests, point confirms*. Once the user arrived at their destination, the exploration of the new close-proximity data entities can continue. The selected data entity, i.e., the selected stacked cuboid, distinguishes itself from all others by featuring a red wireframe outline as opposed to a black one (REQ 12 in Table 4.1).

Information Panels The display of the three information panels as details-on-demand (REQ 9 in Table 4.1) follow generally a similar concept as implemented in the first VE iteration (see Section 5.4.2). The user in the VE may display and hide these panels by performing a "thumbs up" hand posture (thumb pointing

[9] A highlighted stacked cuboid is temporarily increased in its size (scale) to provide a preview as far away nodes appear smaller.

upwards, fingers not extended) with their right hand. The three information panels feature the same arrangement and are also anchored in accordance to the user's head position and rotation at the time of display initiation. However, compared to the first VE iteration and based on some of the received feedback, travel and select mechanisms continue to be available while the information panels are displayed. This enables the user to quickly switch between displaying details-on-demand for data entities in close proximity without the need to first dismiss and hide information panels, as these are automatically updated upon data entity selection. In case of a travel operation to a faraway data entity, the VE automatically dismisses the information panels.

Time Event Selection A stacked cuboid represents comparable data variables of a data item for a specific time domain and at their respective location. For instance, in the case of the NTS dataset, the data has been preprocessed for visualization in the VE on a per week basis with respect to the temporal context. Consequently, a mechanism to allow the user to change the temporal context and thus to explore the dataset with respect to its temporal characteristics is required (REQ 13 in Table 4.1). For this purpose, the VE provides an adapted 2D graphical menu (LaViola, Jr. et al., 2017, Chapter 9.5), attached and juxtaposed to the user's left hand, enabling them to adjust the currently select temporal context for all data entities. In particular, the hand menu features two virtual buttons, i.e., the right button moves the temporal context one step forward, while the left button moves it one step backward. In case of the NTS dataset, this enables to user to move step-wise to the next week, and to the previous one respectively. Additionally, the hand menu also features a textual label that displays the currently selected temporal context (REQ 11 in Table 4.1). With respect to its invocation and availability characteristics (Dachselt and Hübner, 2007), the display of the menu is user-dependent, only appearing when the user's left hand palm is facing towards them, and can be invoked freely at any given point in time during the data analysis activity, allowing the user to change the temporal context in the VE at any time. With a selected data entity and the displayed information panels, the changes in the data item's data variable values can be observed in the center panel. With the hidden panels, the user can instead focus on the observation of how the data entities, i.e., the stacked cuboids, visually change throughout the various locations in the VE.

Feature	**Interaction Description**	**Analysis Task**	**Interaction Technique**
Travel (incl. Select)	Look at a faraway Stacked Cuboid until it is visually highlighted (size scaled up), then point towards it (left/right hand index finger pointing forward) to initiate a position transition via target-based travel. The targeted Stacked Cuboid is automatically selected (see below).	Explore, Select	Selection-based Travel (via Multi-modal Technique: Gaze-based Input and Gestural Command)
Select	Touch the Stacked Cuboid (in close proximity). For faraway Stacked Cuboids, utilize the Travel feature (see above). The selected Stacked Cuboid features a red colored wireframe, as opposed to a default black wireframe of all other Stacked Cuboids.	Select	Hand-Based Grasping
Toggle Information Panels	With the Stacked Cuboid selected, make a “thumbs up” posture (thumb pointing upward, fingers not extended) with the right hand to toggle between displaying and hiding the Information Panels with respect to the currently selected Stacked Cuboid.	Abstract/Elaborate	Gestural Command
Time Event Selection	A two button graphical menu is attached to the user’s left hand, displayed when the left hand palm is facing towards the user. With the graphical menu displayed, the right hand can be utilized to press either of the two buttons to select a new time event, i.e., press the left button to navigate one time event backward in time or press the right button to navigate one time event forward in time respectively. Time Event Selection is enabled at all times, i.e., with displayed and hidden Information Panels, allowing for time navigation at any time.	Select	Adapted 2D Menu (via Graphical Menu)

Table 5.8: Overview of the various features available in the second VE iteration, including descriptions on how to perform them via 3D gestural input, as well as their respective data analysis task and interaction technique classifications.

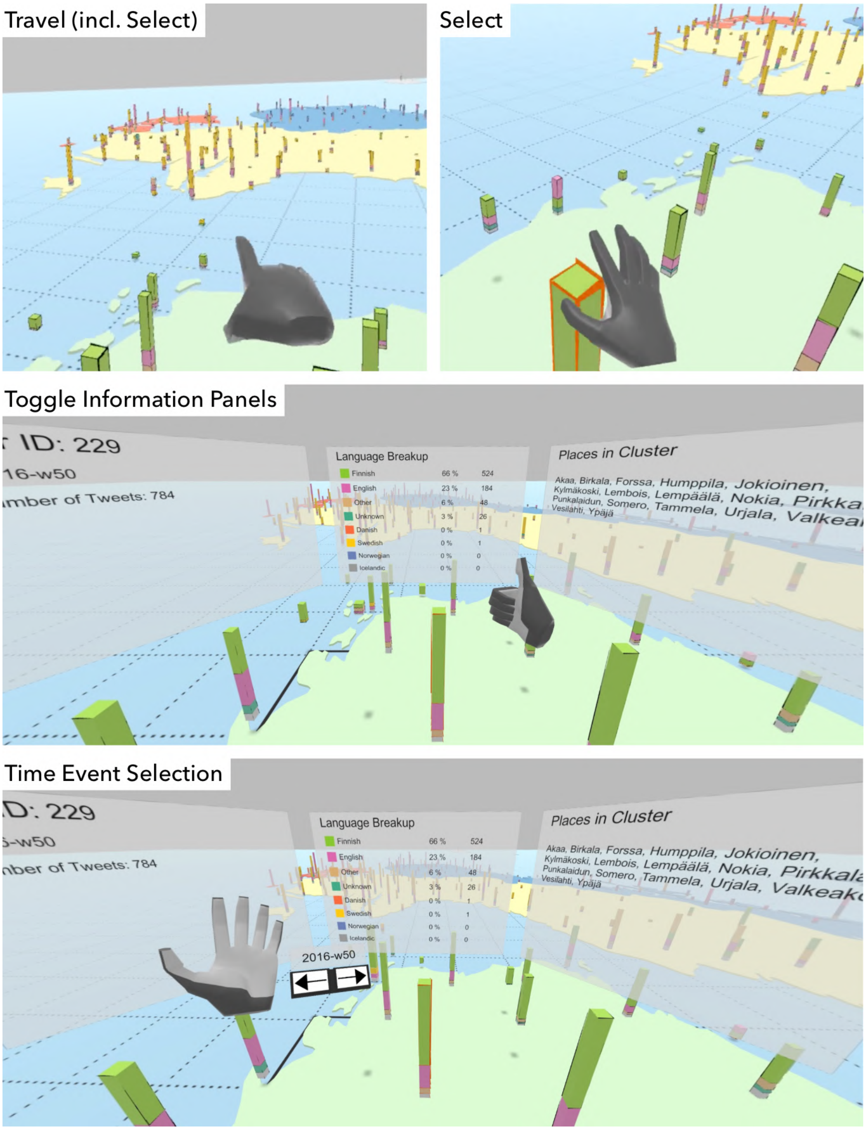


Figure 5.15: Impressions of the various features (see Table 5.8) available in the second VE iteration, from the immersed user's field of view, and utilizing the 3D gestural input modality.

ID	When	What	n	Participants
A	May 2018	lab experiment	7	first-year linguistics students
B	May 2018	demonstration	26	linguistics researchers at ICAME 39
C	Nov 2018	demonstration	11	attendees at LNU Big Data 4
D	May 2019	lab experiment	15	first-year linguistics students
E	May 2019	demonstration	26	researchers at ADDA 2
F	Dec 2019	demonstration	12	attendees at LNU Big Data 5

Table 5.9: Overview of the various lab experiments and hands-on demonstrations, allowing participants to operate the second VE iteration. The students in A and D were enrolled in the *Sociolinguistics* module of the English language B.A. program at Linnæus University. **Notes**: An initial work-in-progress prototype of the developed VE as presented in A and B featured a physical, tracked controller as input modality, based on the HTC Vive, similar as implemented in the first VE iteration (see Section 5.3.2). The VE presented in C–F utilized the 3D gestural input modality as described throughout Section 5.4.2. D is included here as part of the collaborative system evaluation, and described in detail in Section 6.3.3 as part of Chapter 6. The attendees at the LNU Big Data conferences were comprised of students, researchers, and practitioners.

5.4.3 Case Study: Linguistics

At multiple instances throughout the development of the second VE iteration the immersive data analysis environment was presented to the respective target group of linguists in the format of lab experiments with students as well as hands-on demonstrations with both students, researchers, and practitioners (see Table 5.9). These served as an opportunity to gather empirical feedback in regard to the visualization and interaction design as well as with respect to the comparatively novel approach of exploring Twitter data from a sociolinguistic perspective in an immersive VE. In contrast to the interpretation of static visualizations with a predetermined view on the data, the immersive VE enables linguists to be placed in-situ into a virtual landscape that they recognize, and are free to explore places of interest in the 3D space, to extract insights. Naturally, the NTS dataset and its exploration from a sociolinguistic perspective, as introduced in Section 5.4, served as the foundational scenario for the experiments and demonstrations. Users were able to explore the Twitter traffic with respect to the distribution of the main Nordic languages (Danish, Finnish, Icelandic, Norwegian, and Swedish) and English across the Nordic region (spatial) and over time (temporal). Figure 5.16 provides some impressions of various participants engaging with the developed second VE iteration during the public demonstrations.

Within the scope of the various hands-on demonstrations, the overall objective was to gather feedback and thoughts from the linguistics researchers as experts.

Figure 5.16: Immersed participants trying hands-on the second VE iteration at various public demonstrations, i.e., at *The 39th Annual Conference of the International Computer Archive for Modern and Medieval English* (ICAME 39 in 2018), *The 2nd International Conference: Approaches to Digital Discourse Analysis* (ADDA 2 in 2019), and *The 4th and 5th Big Data Conference at Linnæus University* (LNU Big Data 4 and 5 in 2018 and 2019).

The first lab experiment (see Table 5.9 A) was organized as a small pilot study to collect feedback directly from linguistics students. Additional feedback from linguistics students was acquired within the scope an empirical evaluation (see Table 5.9 D) that involved the presented second VE iteration and the stacked cuboid visualization approach in a collaborative data analysis setting. The students were introduced to the concept of multilingualism as part of the course work before the respective lab experiments. Naturally, throughout all experiments and demonstrations, the participants were briefly introduced to the data context as well as the VE and its features, allowing them thereafter to immerse themselves in the data analysis environment. This enabled them to provide hands-on feedback as they were interacting in the VE,[10] and of course post-VE immersion, continuing their conversation with the respective researchers.

As a practical task reference, the participants were generally encouraged to explore the "data landscape" freely to identify a location that seemed of interest to them. Additionally, a second alternative task was posed to the participants in some instances, asking them to identify regional patterns, for instance analyzing the distribution of English in Nordic tweets or examining where people tweet most in a specific language. It was interesting to observe that similar data exploration strategies emerged as observed in the evaluation of the first VE iteration (see Section 5.3.5.2). Some would travel from data entity to data entity, guided by the temporal data encoded directly in the stacked cuboid, previewing the most dominant languages at a location. Naturally, locations with many tweets, indicated by an increased height in the VE compared to stacked cuboids with fewer tweets, often caught the participants' attention. At the same time, they were also able to detect anomalies, for instance a location that featured a different dominant language compared to most of the others in their spatial surroundings. Some participants initiated their data exploration by first moving to a peripheral location, allowing them to get an overview by having most of the data entities in their field of view, followed by a subsequent point of interest identification. However, there were also participants that utilized previous knowledge or a personal connection to a location to guide their data analysis. For instance, they either traveled to a location they knew would be interesting based on existing background information, such as the demographics in a specific region, or attempted to find a specific data entity that included data about the place where they grew up, to name just two examples. Seemingly, most of the participants attempted first to find an interesting location based on the selected temporal context at the start of their data analysis activity, and only engaged with the time event selection feature in-situ when they were at an interesting location, already examining details-on-demand. The utilization of the time event selection, globally changing all data entities in the VE, with the objective to identify an

[10] To provide a general reference, most of the participants spent 10 to 15 minutes in the VE, independent of lab experiment or demonstration. The exact duration was not measured.

interesting location to begin with, was a mechanism the participants made use of as they became more familiar in the VE later during their data analysis activity. Even though they were visually immersed in the virtual 3D space, their interactions generally led to enthusiastic conversations with the researchers, and other bystanders in case of the public demonstrations, speculating about possible reasons behind the observed phenomena.

None of the participants had prior experiences with the developed VE iteration. In fact, only few of them had even tried VR technologies beforehand. Nonetheless, all of the participants were able to learn the 3D UI and adopt the features provided in the VE comparatively swiftly – naturally some slightly quicker than others. Furthermore, all of them were able to make meaningful assessments. In the case of the students, their assessments were adequate with respect to what is expected of them at that stage in their education according to a respective expert, i.e., a professor of English linguistics.

The participants were generally able to orient themselves in the Nordic region, based on the Nordic countries that were displayed as color-coded extruded surfaces on the virtual floor (see Section 5.4.1). Nevertheless, some of them thought that it would be worth considering to have additional landmarks on the map for further guidance, for instance with respect to the capital cities. The participants were able to quickly understand the visualization concept of the stacked cuboids. Arguably to be expected, some of the participants commented that it would be interesting to read the actual tweets while being immersed in the VE as part of the displayed interface. During their interaction in the VE, particularly when inspecting the data presented in the information panels, several participants became confused with the language entry *Unknown*. In these cases, the researchers explained that such a language classification in the NTS dataset is commonly related to tweets that only contain a hyperlink or multimedia content. Some of the participants also observed and actively discussed the distributional differences in the amounts of tweets from region to region that appeared to vary significantly between different time event (week) selections. Such observations were particularly made for regions with a comparatively low population in general.

5.4.4 Use Case: Swedish Election

Following the overall concept of the presented stacked cuboid data entity visualization and interaction design of the second VE iteration as presented in Sections 5.4.1 and 5.4.2, a proof-of-concept prototype was developed with two intentions in mind. First, the presented immersive data analysis environment is demonstrated in the context of a dataset and scenario that differs from the NTS one, which is used as the main data context through the presented second VE iteration. And second, to illustrate the overall data-agnostic capabilities of the developed system (REQ 5 in Table 4.1), enabling the display of data across various contexts with comparatively minimal additional implementation efforts.

Inspired by the election data context used in the first VE iteration, the developed proof-of-concept prototype is centered around the immersive data analysis of the Swedish parliament (Sveriges riksdag) election results. The election results for each political party on a per municipality basis for the parliament elections in the years of 2010, 2014, and 2018 are provided by the Swedish Election Authority.[11] Additional information about all the municipalities in Sweden are available through Statistics Sweden.[12] Based on these two data sources, a dataset was compiled under consideration of the system architecture's data structure reference model (see Section 4.2.1). Importing the dataset into the developed second VE iteration, it is possible for the immersed user to explore the election results in regard to the various municipalities (spatial) and across the three included election years (temporal).

Figure 5.17 provides some impressions of the presented use case. Compared to the NTS main data context and VE composition as described in Section 5.4.1, the Swedish election use case features a VE composition as follows. All 290 municipalities are displayed as as colored extruded surfaces on the virtual floor. The brightness of a municipality's color encodes the respective voter participation, i.e., a brighter color value represents high voter participation, and a lower brightness low voter participation. Each stacked cuboid data entity visualization represents an individual municipality, placed accordingly at the respective centroid of an extruded surface. Each stacked cuboid is composed of four cuboids, i.e., the three top ones representing, in order, those political parties that received the most votes, and the forth cuboid accumulating the share of all remaining parties. The immersed user can display details-on-demand in the format of the three information panels. The center panel features a detailed listing about the voting results distribution of the selected municipality, i.e., political party name, assigned color coding, voting results in percentage, and absolute received voting numbers. The left panel displays more detailed information about the municipality within the context of the election, among others the total numbers of valid votes, the total numbers of votes, the total number of eligible voters, and the final voter participation in percentage. The right panel displays some information about the use case and the utilized data sources. Interaction in the VE is possible as described in Section 5.4.2. The immersed user can travel around in the 3D space, make selections, display details-on-demand, and change the temporal context by selecting a new time event with respect to the three available election years.

---

[11] Valmyndigheten. Valresultat. Retrieved June 1, 2022, from `https://www.val.se/valresultat.html`

[12] Statistikmyndigheten SCB. Digitala gränser. Retrieved June 1, 2022, from `https://www.scb.se/hitta-statistik/regional-statistik-och-kartor/regionala-indelningar/digitala-granser/`

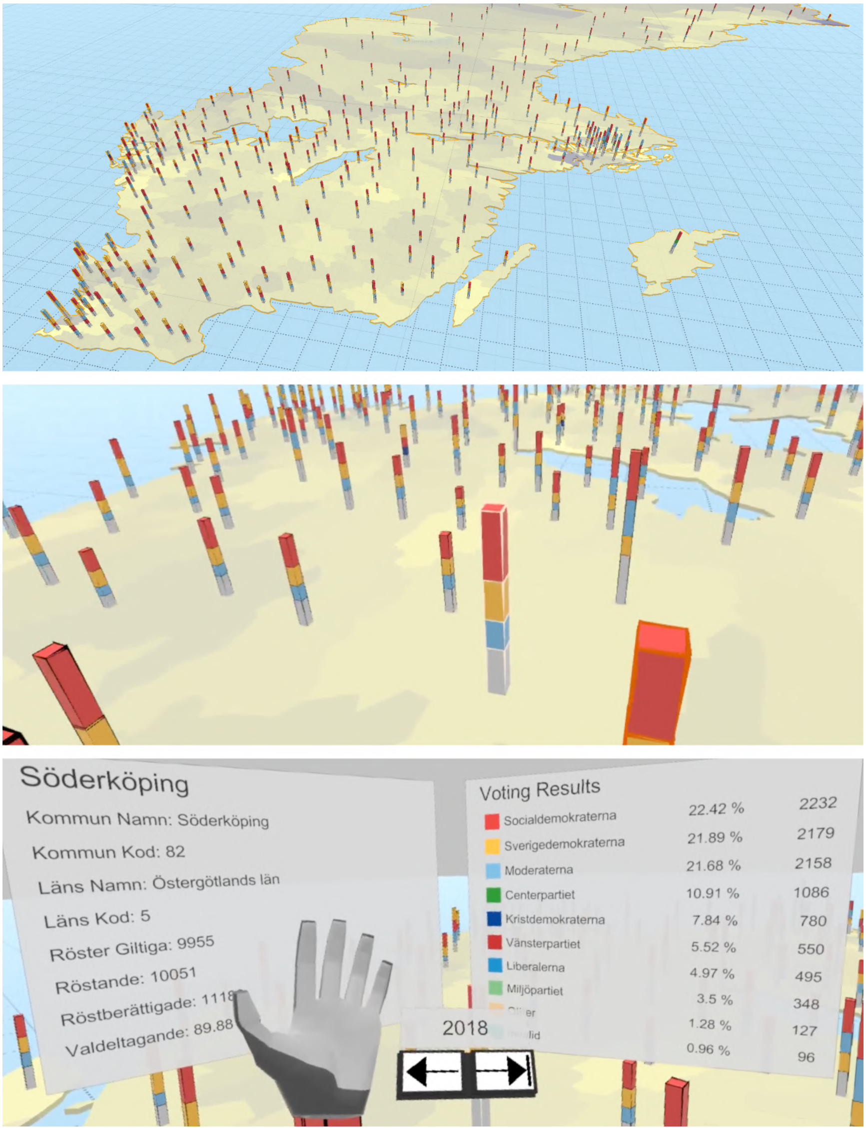


Figure 5.17: Impressions of the second VE iteration, demonstrating the immersive data analysis use case for the Swedish Election. **Top**: The VE composition, from an angled top down view to provide an overall impression. **Middle**: The immersed user's field of view during an overview-like spatial exploration. **Bottom**: The immersed user's field of view during a details-on-demand temporal exploration with the data entity selected that represents Söderköping Municipality, displaying the information panels and the time event selection hand menu.

5.4.5 Discussion

The response from the participants in regard to the presented second VE iteration was overall very positive, with several of them stating that the immersion allowed them to experience the displayed data in a subjectively "*real*" and memorable manner. The majority of the positive feedback from the students, researchers, and practitioners was related on the immersive characteristics of the VE, enabling them to focus their attention on the displayed data – due to the nature of the involved technologies, arguably even requiring them to do so. Compared to a visualization shown on a TV, projected against the wall, or printed in a book, the immersive data visualization cannot just be casually glanced over or looked at quickly. Based on the hands-on experiences of the students, it appears that such an immersive data analysis approach can be a particularly useful teaching tool, allowing the students themselves to discover relations and patterns in the data in an interactive and engaging way.

Compared to the first VE iteration as presented in Section 5.3, the second one utilized a stacked cuboid visualization design to encode temporal information directly in the data entity, allowing for spatio-temporal data analysis in the VE. In that regard, it was particularly interesting to observe how the participants explored the temporal data context and utilized the respective time selection feature. The majority started their data analysis under consideration of the spatial context first, and engaged more deeply with the temporal contextual exploration only later in their activity. Arguably, a spatial location, i.e., a data entity in the VE, provided them with an important reference point at the start of the data analysis. Such a reoccurring data analysis behavior is something to take into consideration, especially with respect to the design of a VE. For instance, what changes could be made to the VE design to facilitate the user's ability to first analyze the displayed data in regard to a temporal context, and only afterwards with respect to the spatial data variables? The time event encoding and related interactive features worked well in general. However, the immersed users also noticed that it was somewhat difficult for them to get an overview about the temporal data context, for instance in regard to their ability to identify trends over time. In the case of the presented NTS dataset and the sociolinguistic analysis scenario, such diachronic trends can be of particular interest to the linguists. In that regard, the visual encoding of only a single time event in a data entity was not sufficient enough for the participants to properly grasp the overall trend of a data variable over time at a specific location.

Some of the researcher participants at the demonstrations also expressed a desire to take notes or read textual data, such as respective tweets, directly in the VE. According to them, the inability to do so could limit certain research tasks. The export of their findings for use after the immersive data analysis activity seemed of particular importance and interest to them, allowing the researcher to continue their data analysis workflow using other tools and methods.

To summarize, the positive and constructive empirical feedback, obtained from various linguistics researchers and students with respect to the developed second VE iteration, highlights the overall potential that VR-based tools can offer to data-rich disciplines in the humanities. Immersive display and interaction technologies may support and facilitate data analysis as part of the workflow alongside other non-immersive programming environments and statistical tools. Particularly with respect to the visual stimuli, explorative data analysis in an immersive VE can provide important impulses to detect patterns and points of interest in the data that can guide and inform the subsequent data analysis. Furthermore, the interactive and immersive VE represents a tool that can be utilized by students to explore large multivariate datasets that highlight linguistic diversity in an engaging manner that is arguably often lacking in the existing tools in corpus linguistics. As such, immersive data analysis tools have the potential to synergize with existing tools and practices, complementing data exploration and result visualization in the presented context.

5.5 VE Iteration 3: Data Analysis Using 3D Radar Charts

Based on the insights gained throughout the design, development, and evaluation of the second VE iteration as described throughout Section 5.4, a third and, within the scope of this thesis, final VE iteration was approached. The main objective of the third VE iteration is to further advance the data analysis support with respect to the temporal context. Compared to the second VE iteration, the third one focuses on two aspects in particular. First, while the stacked cuboid visualization design was used to visually encode data with respect to a single time event at a time, the data entity visualization design in the third VE iteration should support the encoding of a time series that consists of two or more time events. And second, inherent from such a more complex data entity visualization design, the support for features to accommodate interactive data analysis of the time-series visualization should be investigated.

Dataset: Plant-Weather timelines (PWt) After the utilization of the NTS corpus as a real-world dataset throughout the work on the second VE iteration (see Section 5.4) and mainly used for explorative analysis tasks due to the corpus' characteristics, a baseline dataset should aid the empirical evaluation of the work within the scope of third VE iteration, for instance to appropriately evaluate user performance for confirmative analysis tasks. A variety of open data sources for

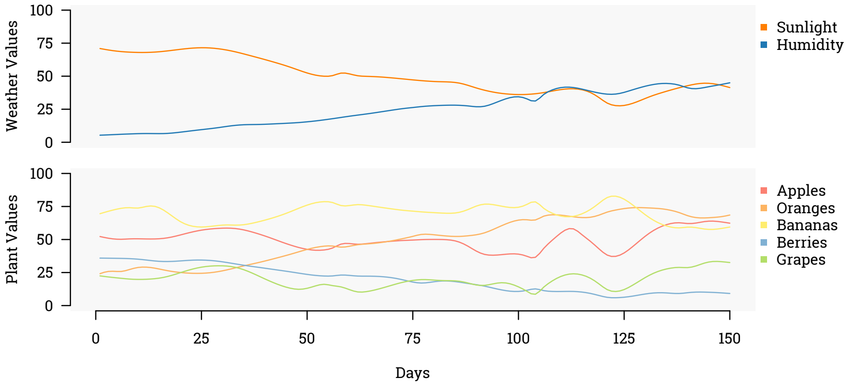


Figure 5.18: An example of the generated Plant-Weather timeline data for one location. Each of the plant (fruit) variables is either positively or negatively correlated to each of the two weather (sunlight and humidity) variables. **Note**: Within the scope of this thesis, the plant variables are utilized in the third VE iteration, while the weather variables are utilized in the non-immersive desktop terminal as part of the collaborative data analysis scenario (see Section 6.4).

real-world inspiration and potential use were considered.[13] Also in anticipation of potential collaborative data analysis tasks, later described in Section 6.4 as part of Chapter 6, a spatio-temporal dataset was needed that allowed the convenient evaluation of the users' data analysis activity with respect to their performance, not just for open-ended explorative analysis tasks, but also for confirmative analysis tasks, where they examine the dataset to confirm or reject hypotheses that they were given beforehand (see Section 5.2). In particular, a confirmative analysis task allows for a more direct task performance comparison among different study sessions. Thus, a baseline or "benchmark"-like dataset was needed that would allow to define representative analysis tasks that could be used to asses in a comparative way the participants' ability to complete a specific task using the interfaces provided as part of the third VE iteration. Unfortunately, to the best assessment, none of the referred real-world datasets would have allowed to easily achieve this.

[13](1) Global Change Data Lab. Our World in Data. Retrieved June 1, 2022, from `https://ourworldindata.org/`; (2) Christopher K. Wikle, Andrew Zammit-Mangion, and Noel Cressie. Spatio-Temporal Statistics in R. Retrieved June 1, 2022, from `https://spacetimewithr.org`; (3) Edzer Pebesma. Time, Space, Spacetime in R. Retrieved June 1, 2022, from `http://pebesma.staff.ifgi.de/R/Lancaster.html`; (4) The Agency for Digital Government (DIGG). The Swedish dataportal. Retrieved June 1, 2022, from `https://www.dataportal.se/en`; (5) Swedish Meteorological and Hydrological Institute. Official Website. Retrieved June 1, 2022, from `https://www.smhi.se/en/`

Consequently, the *correlated-timelines* project, developed by Aris Alissandrakis,[14] was utilized to create the Plant-Weather timelines (PWt) dataset for use within the scope of the third VE iteration, i.e., a custom, representative multivariate dataset that features artificially generated data. The data context is held purposefully simple to understand, allowing to be as inclusive as possible with respect to the recruitment of participants for any related empirical evaluations, as no expert knowledge is required. With the focus on spatio-temporal data, time-series data of plant and weather data variables were generated for 39 (spatial) locations, i.e., 39 countries in Europe. Each location features five *plant* data variables (different types of fruits or vegetables) and two *weather* data variables (sunlight and humidity). Each of these seven data variables per location features a time series of 150 consecutive time events on a per day basis. Thus, there is a total amount of 40.950 data variable values in the generated dataset.[15] The special property of this artificially generated dataset is that each of five plant variables features either a *positive* or a *negative correlation* to each of the two weather variables. While the values for all of the variables are diverse across the different locations, the correlations are coherent with a defined model, i.e., the correlations between the five plant and two weather variables are the same independent of the location. Figure 5.18 presents an example of a generated plant and weather timeline for one location. Within the scope of the the third VE iteration as described throughout this section, the focus is on the utilization of the five plant variables as the scenario for the immersive data analysis activity. The contextual interplay between these and the two correlated weather variables is relevant within the scope of the collaborative data analysis scenario (see Section 6.4).

For each location, the two *sunlight* and *humidity* timelines were generated using a R function.[16] Each of the five plant[17] timelines were generated by adding the humidity and sunlight timelines, multiplied by the weights as dictated by the model (either one or minus one, to indicate a positive or negative correlation respectively). These timeline data were further validated to confirm the compliance to the applied model.[18]

[14] Aris Alissandrakis. correlated-timelines. Retrieved June 1, 2022, from `https://github.com/arisalissandrakis/correlated-timelines`

[15] $39\ locations \times 7\ data\ variables \times 150\ time\ events = 40.950\ data\ variable\ values$

[16] Each timeline was generated taking into account length (number of time events), minimum and maximum values, a regression slope, amount of noise, and a series of normal distributions that could be added at different places along the timeline. The function output was further smoothed as a spline, and vertically scaled and/or re-positioned. For more details, see the GitHub repository referred in Footnote 14 and Appendix D.

[17] Apples, Oranges, Bananas, Berries, and Grapes for the fruits scenario; Tomatoes, Carrots, Potatoes, Cabbages, and Lettuces for the vegetables scenario.

[18] Sign and value of Pearson's correlation coefficients agreed to the defined model, and p-values were below significance level for the majority of plant-weather pairs, allowing the model to be applied as the base truth to measure the participants' observations against.

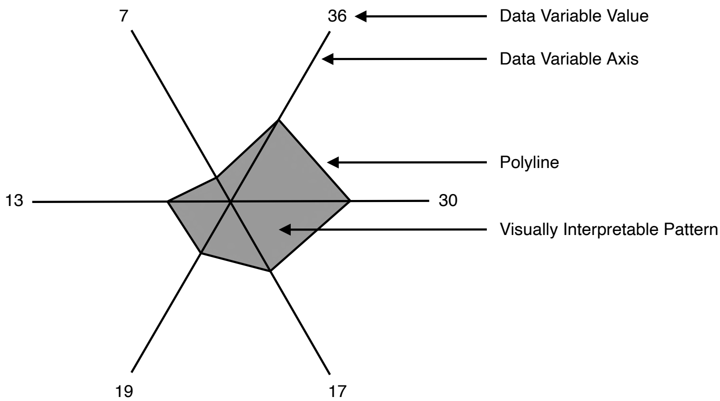


Figure 5.19: An example of a radar chart and its components, adapted from Kolence and Kiviat (1973).

Radar Chart: Foundations The idea of utilizing a *radar chart* approach, among others also known as *Kiviat figures* or *star plots*, for the purpose of software unit visualization has been described in 1973 by Kolence and Kiviat (Kolence, 1973; Kolence and Kiviat, 1973). Rather than presenting values of individual data variables perpendicular to one another, as for instance in the format of a histogram or bar chart, they are instead radially arranged as data variable axes (Kolence, 1973). The values for each data variable along the different adjacent axes can then be connected by a polyline, resulting in a visually interpretable pattern (Kolence and Kiviat, 1973). Figure 5.19 provides an example of a radar chart and its components.

The radar chart has become an established method to visualize multivariate data in 2D across various contexts and scenarios. Within modern 3D graphics computing over the years however, various attempts to transfer the concept of the original 2D radar chart into the 3D space have been made, often with the aim to utilize the additional graphical dimension to visualize further information. A common use case is to utilize that third additional dimension to visually encode time events, visualizing changes in the data over time by stringing together multiple 2D radar charts in 3D. The idea of using a 3D volumetric approach to generate a *Kiviat tube* within the context of visualizing parallel computing processes has been demonstrated by Hackstadt and Malony (1995) as well as Heath et al. (1995). Akaishi and Okada (2004) describe their *Time-tunnel* approach, a 3D presentation tool viewed through a normal computer monitor, placing individual time-series data variables in the format of 2D line charts as *data-wings* along the third dimension in radial arrangement. Individual axes can then be rotated and overlapped with other axes in 3D, allowing comparison between different data variables (Akaishi and Okada, 2004). A series of different 3D axes-based visualization approaches for temporal data contexts has also been explored by Tominski et al. (2005), for instance as *3D Time Wheel*, *3D Multi Comb* as a similar approach to the one presented by Akaishi and Okada (2004), and *3D Kiviat tube* that is inspired based on the work by Hackstadt and Malony (1995). A slightly different approach is described by Kerren and Jusufi (2009), using the 3D space and a *fanning out* metaphor to create interactive visualizations of software metrics, allowing users to interact with the individual axes to examine the data in different spatial configurations with the aim to overcome occlusion problems. Draper et al. (2009) surveyed radial methods within the context of InfoVis, including *star plots*. Forlines and Wittenburg (2010) explored an approach of visualizing radar charts in 3D as *Wakame*, stacking 2D radar charts in 3D, creating a hollow tube-like shape that encodes time along the third dimension. The application of temporal radar plots within the context of VA has also been explored by Peters (2014). Aiello et al. (2015) investigated the placement of individual 2D radar charts in the 3D space with the specific goal to highlight trends in the data variables' values over time. The use of 3D Kiviat plots, similar to the Wakame approach described by Forlines and Wittenburg (2010), was extensively explored by Wang (2017) in a scenario aimed towards fault detection and process monitoring.

Such approaches of transferring 2D visualizations into the 3D space are promising and invite for further investigations. It is noteworthy that all the work described throughout the previous paragraph utilized 3D graphics, but were displayed through a non-immersive 2D visual display. Within the scope of this thesis and the analysis of data in an immersive VE, some exciting possibilities arise with respect to the utilization of a 3D radar chart as data entity visualization, for instance allowing the encoding of not just one time event of the dataset, as presented with the stacked cuboid approach in Section 5.4.1, but instead the encoding of a time series that consists of multiple consecutive time events. Considering the advantages that immersive technologies can provide, among others a better spatial understanding through stereoscopic depth cues as described in Section 2.2.1, revisit and reiteration of 3D visualization approaches and their respective interactive features are subjects that are worth investigating further. Consequently, the design and development of the third VE iteration for spatio-temporal data analysis is centered around a 3D radar chart approach that is inspired by the overall concept of visualizing data as *Kiviat figures* (Kolence and Kiviat, 1973) and has similarities to the *Time-tunnel* approach originally presented by Akaishi and Okada (2004).

5.5.1 Visualization Design and VE Composition

The visualization design to represent individual data entities as *3D Radar Charts* in the third VE iteration is, as previously stated, similar to the *Time-tunnel* approach

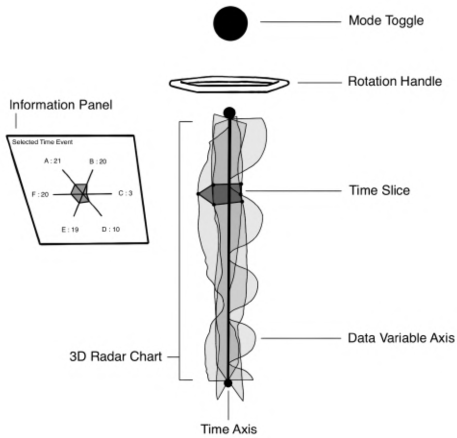


Figure 5.20: *3D Radar Chart* visualization design of the third VE iteration.

(Akaishi and Okada, 2004). Figure 5.20 conceptually illustrates the 3D radar chart approach of the third VE iteration. In 3D, the vertical dimension represents the time domain and is visualized through a black axis with start and end points. Following the conceptual approach of radar charts in general, individual data variables are organized as individual spokes in a radial arrangement around the time axis, i.e., the time-series data for each data variable is visualized as a 2D frequency polygon. The angular rotation for each spoke, or *data variable axis*, is based on the overall amount of displayed data variables. As the *time axis* represents the origin for each data variable axis, naturally if a data variable value (magnitude or frequency) is closer to zero, it is located closer to the time axis, while higher values are farther away. Each data variable axis is color-coded and semitransparent. Such an arrangement in 3D, and facilitated through the stereoscopic capabilities of a HMD, should allow the user to get a spatial visual impression of the data over time. Rather than creating an occlusive 3D tube, the visualization of each data variable axis as a semitransparent 2D frequency polygon is intended to prevent occlusion, independent of the user's viewpoint in the VE. Thus, the user should be able to visually perceive, at the least, a preview of those data variable axes that are located behind the ones that are currently in the front. This overall visualization design is intended to provide the user with the ability to visually detect patterns and identify time events and ranges deemed interesting, both per individual data variable as well as in relation to all others in a 3D radar chart.

Conceptually following the features provided throughout the first and second VE iterations (see Section 5.3.1 and 5.4.1), the user can engage in a closer in-situ interaction with an individual 3D radar chart, among others to display details-on-demand (REQ 9 in Table 4.1). In the case of the 3D radar chart visualization, the display of details-on-demand is implemented through (1) a 2D radar chart visualization that is directly integrated into the 3D radar chart itself, and (2) one juxtaposed *information panel*. In particular, the integrated 2D radar chart, subsequently referred to as *time slice*, is a semitransparent 2D mesh created from the 3D vertices based on the individual values in all the data variable axes at a given point in time. Thus, the time slice represents the traditional interpretable pattern of a radar chart, facilitating the examination and comparison of all data variables and their values at a specific point in time (REQ 11 in Table 4.1). Naturally, the time slice updates its shape in accordance to the selection of new events along the displayed time-series data of the 3D radar chart. The vertices of the time slice are additionally highlighted as small spheres to provide visual guidance, and are color-coded based on the respective data variable axes. To provide textual and numerical information about the data using the 3D radar chart visualization design, the VE displays a complementary information panel that is juxtaposed to the time slice. This approach is similar to the information panels utilized in the first and second VE iterations, but differs in a few aspects. Arguably most noticeably, it consists of only one information panel instead of three. Furthermore, rather than aligning it with the user's head position and rotation in the VE upon display, it is anchored next to the time slice of a 3D radar chart. Once the position of the time slice is updated along the 3D radar chart's vertical time axis, the position of the information panel is updated accordingly. As for the composition of the information panel's contents, it features a traditional 2D radar chart presentation of the currently selected time event (REQ 11 in Table 4.1). Each data variable axis displays a caption that consists of the axis' respective color coding as well as the data variable's name and value. Additionally to the 2D mesh representing the current time event, the information panel displays a radar chart outline that represents the calculated average values across the data variables for the entire time series.

Based on the presented data entity visualization design as 3D radar charts, the composition of the VE follows generally the same approach as presented in the second VE iteration (see Section 5.4.1). The grid on the virtual floor remains to indicate the immersed user's safe interaction area and its boundaries. Additionally, the floor is populated with the respective visual representation of geographic features that are relevant within the context of the dataset (REQ 11 in Table 4.1). For instance, in case of the presented PWt dataset, not just the Nordic countries are displayed as extruded surfaces on the virtual floor, but the majority of countries across the European landmass. Individual 3D radar charts are placed in accordance to their geolocation in the VE with the time axis at the respective center, floating 40 cm above the floor, and casting a shadow indicating its exact location. The technological implementation of the 3D radar chart (see

Appendix B) enables various options for configuration. For instance, the vertical length (or height) can be predetermined and set to a static value independent of the amount of displayed time events. Alternatively, that length may also be dynamically derived from the amount of displayed time events and based on a fixed distance descriptor that specifies the vertical distance between two time events in the data variable axis. Throughout the application of the 3D radar charts within the scope of this thesis and the presented empirical evaluations, unless otherwise noted, the 3D radar charts have been configured to feature a fixed vertical length that corresponds to 100 cm. Thus, a 3D radar chart's top is located approximately in the upper chest region, depending on the height of the user and their position in the VE. Through the encoding of the time-series data directly in the 3D radar chart as described throughout this section, it is possible to simultaneously display many temporal data values in the VE. A detailed example for this is described as part of the VE setup of the second evaluation (see Section 5.5.5), displaying 39 3D radar charts, each with five data variable axes that are composed of 150 time events, resulting in a total of 29.250 time event data values displayed in the VE.

5.5.2 Interaction Design

To facilitate the analysis of spatio-temporal data in an immersive VE, and aligned with the 3D radar chart visualization design presented in Section 5.5.1, various interactive features are available in the VE. Following a prototypical approach as well as naturally considering the various theoretical aspects (see Sections 2.2.4 and 5.2) and the obtained insights from the first and second VE iterations, several features to interact with 3D radar charts via 3D gestural input were designed.

Naturally, during the immersive data analysis activity, the user is arguably going to perform certain tasks more frequently than others. This requires to keep in mind hand comfort recommendations, for instance as reported by Rempel et al. (2014), to avoid the use of uncomfortable hand configurations for anticipated frequent interactions. Assuming a general VE composition as described in Section 5.5.1, i.e., the VE being populated with various data entities, each representing different time-series data according to their respective location, a *travel* feature as presented in the first and second VE iterations remains necessary to enable user movement beyond the physical space limitations of the VR system's calibrated safe interaction area. Consequently, the virtual 3D space can be utilized in order to allow the user to explore the data in a more overview-like manner (Shneiderman, 1996), conceptually similar to *walking among the data* (Ivanov et al., 2019; Streppel et al., 2018). Naturally, when discovering something of interest, the user is expected to engage in-situ with the data to display *details-on-demand* (Shneiderman, 1996), thus entering a closer contextual interaction (Nehaniv et al., 2005). At this stage, the user is expected to *select time events* and *time ranges* as well as to potentially *reconfigure (sort)* the order and *filter* out individual data variables. Besides these envisioned frequent tasks, the user may also use other, arguably, more infrequent ones, such as to *zoom in* and *out* of the temporal data context or to *reset* any prior interactions and manipulations with a 3D radar chart.

Table 5.10 provides a comprehensive overview of these features and their descriptions, the analysis task and interaction technique classifications, as well as with respect to hand posture comfort based on the assessments provided by Rempel et al. (2014). Figures 5.21, 5.22, and 5.23 provide an overview of the implemented interactive features from the user's field of view in the VE. Additionally, Figure 5.24 illustrates the applied hand posture comfort configurations. The remainder of this section provides detailed descriptions about the implemented interactive features within the scope of the third VE iteration.

Travel To virtually move around to data entities outside the immersed user's safe interaction area, the same multimodal approach of gaze-based input and gestural command is followed as in the prior VE iterations (REQ 8 and REQ 12 in Table 4.1). Once the user centers their gaze on a faraway 3D radar chart, a black frame is temporally displayed to provide visual feedback with the aim to clearly indicate which 3D radar chart is targeted. Applying the same "*I want to go there*" analogy and *gaze suggests, point confirms* principle, the user can then point towards the targeted 3D radar chart to initiate an automatic transition via *target-based travel*. As in the second VE iteration, the destination data entity is always placed to be located in the center of the immersed user's safe interaction area, ensuring that they can freely move around the 3D radar chart, without obstacles, to observe and inspect all its data variable axes – a mechanism that is arguably even more important in this VE iteration compared to the second one, as the 3D radar chart approach encodes a considerable larger amount of temporal data values and in a radial arrangement than in comparison to the stacked cuboid approach (see Sections 5.4.1 and 5.5.1).

Selection through Mode Toggle To engage into a closer in-situ and details-on-demand interaction with a 3D radar chart, a *Mode Toggle* feature is provided (REQ 7 and REQ 9 in Table 4.1). In particular, a mode toggle widget in the format of sphere is placed above a 3D radar chart. By directly touching the widget, the user can iterate between three overall interaction states, i.e., (1) Activate/Rotate, (2) Reconfigure/Filter, and (3) Deactivate. The widget is color-coded to provide visual feedback about its state, i.e., red when the user is not engaged with the 3D radar chart (state 3), and green when the user is engaged (states 1 and 2). Once activated, the 3D radar chart displays the time slice to indicate what time event is currently selected, subsequently enabling the selection of a new time event, as well as the juxtaposed information panel to show details-on-demand. Additionally, the mode toggle controls what other indirect widget is currently displayed, i.e., either the rotation handle or the reconfiguration and filter handle widget. While the user is engaged in such a close interaction with an activated

3D radar chart (REQ 12 in Table 4.1), the travel feature is temporally deactivated to prevent accidental travel operations to other data entities.

Rotation, Sort, and Filter Once the user engaged into close interaction with a 3D radar chart via respective mode toggle interaction, and thus based on its interaction state, one of two possible 3D widgets are available that allow indirect interaction (LaViola, Jr. et al., 2017, Chapter 7.7.3) with the 3D radar chart. The *Rotation Handle* is composed to represent the 3D outline of a geometrical shape that is aligned with the total amount of data variable axes in the 3D Radar Chart, for instance the shape of a pentagon in the case of five data variable axes. The handle's rotation in the VE is directly linked with the rotation of the 3D radar chart and thus with the data variable axes. Thus, the user may grab and hold the rotation handle, and then drag it left or right to *rotate* the 3D radar chart in place. Alternatively, they can also give it a little left or right "flick" with their hand, to initiate a physics-based rotatory movement, similar to a carousel in the real world. The *Reconfigure and Filter Handle* is composed of various *data variable axis spheres*, equivalent to the amount of data variable axes displayed in the 3D radar chart. Each data variable axis sphere is color-coded according the data variable they represent, and visually connected through a line to the time axis origin, resulting in a star-like shape. Data variable axis spheres are interactive in order to implement two fundamental features. First, by grabbing and holding a data variable axis sphere, the user can manipulate the radial arrangement of the displayed data variable axes by moving the held sphere around, essentially changing their order and thus *sort* the data variables (REQ 14 in Table 4.1). And second, by grabbing and dragging a data variable axis sphere away from the time axis origin, i.e., until its visual connection disappears (or "snaps"), the user may *filter* out undesired data variables that will be removed from the 3D radar chart[19] (REQ 14 in Table 4.1). When available, either of the two widgets, i.e., the rotation handle or the reconfigure and filter handle, are placed in between the space above the 3D radar chart and beneath the spherical mode toggle widget.

Time Event Selection The *selection* of a new *time event* is possible by utilizing the time slice that is integrated in the 3D radar chart (REQ 13 in Table 4.1). The user can directly interact with the time slice and manipulate its position inside the 3D radar chart through a hand-based grasping technique. More specifically, they can reach out with their hand, grab and hold the time slice, and consecutively drag it up and down along the 3D radar chart's time axis to a desired point in time by naturally moving up and down their hand. The time slice automatically updates its shape based on the newly selected time event along the way.

Time Range Selection With respect to potential interactions with the time-series data, the VE does not just provide means to select a single time event, but

[19]Based on the overall concept of a radar chart, at least three data variables must remain at any given point in time.

also to *select* a *time range* that is composed of multiple consecutive time events (REQ 13 in Table 4.1). Using a gestural command with both hands, each being held in a pinch posture (thumb and index finger held together), the user can vertically unfold and "sculpt" a desired range of time events. Those included remain color-coded, while the time events outside the selected time range are displayed in a neutral gray color. This allows the user to visually identify the selected time range on the one hand, while on the other still retaining some visual indications about the temporal context outside the selected time range. The user is restricted to select time events only within the selected time range.

Zoom Depending on the amount of time events in the time-series encoded over the length of a 3D radar chart, the VE provides a feature to *zoom* with respect to the temporal data context (REQ 14 in Table 4.1). With a time range selected, the user may *zoom in* by temporary "stretching" their time range selection over the entire virtual length of the 3D radar chart, visually cutting off any time events outside that range. Reversely, assuming the entire time-series is not already displayed, the user may also *zoom out* from previous zoom in interactions. Aligned with such zoom interactions, an underlying *history* functionality allows for step-wise zoom out based on multiple prior zoom in operations. These interactions are implemented through respective gestural commands with both hands, moving them either apart to zoom in or towards each other to zoom out.

Reset Due to the comparatively more complex and advanced interaction with 3D radar charts as data entity visualizations than compared to the sphere and stacked cuboid approaches used in the first and second VE iterations (see Section 5.3.2 and 5.4.2), there is consequently also a higher need to provide the user with means to reverse selections and manipulations. Thus, the VE provides a *reset* feature that conveniently reconfigures a 3D radar chart back to its original state, i.e., displaying all available time-series data as well as all data variable axes arranged in their original order (REQ 14 in Table 4.1). Naturally, such a reset operation is a comparatively drastic one, that in turn should be performed with caution. To prevent the unintentional operation of the reset feature in the VE through the 3D gestural interface, it has been mapped onto a respective gestural command with both hands that is arguably rather unlikely to be performed by accident under normal circumstances. In particular, the user is required to hold up and cross their index fingers, composing a "X"-like posture accordingly.

Pause / Resume The VE allows the user to temporarily *pause* (and *resume*) any kind of interaction, by briefly holding their hands in front of them in a "stop"-like posture. The intention is to provide a mechanism that actively prevents feature execution due to unintentional hand movements (Pavlovic et al., 1997) during periods when the user desires to make observations in the VE more passively.

Feature	Interaction Description	Analysis Task	Interaction Technique	Comfort
Travel	Look at a faraway 3D Radar Chart until its outline is displayed, then point towards it (left/right hand index finger pointing forward) to initiate position transition via target-based travel.	Explore	Selection-based Travel (via Multi-modal Technique: Gaze-based Input and Gestural Command)	5c
Mode Toggle	Touch virtual *Mode Toggle* widget to iterate between three states: (1) Activate/Rotate, (2) Reconfigure/Filter, and (3) Deactivate.	Abstract/Elaborate, Change configuration	Hand-Based Grasping	1c, 5c, 9c (either)
Rotation	Grab *Rotation Handle* widget and drag it left/right to rotate around its Time Axis <u>or</u> give the *Rotation Handle* widget a little left/right flick with the whole hand.	Change Configuration	Indirect Widget (via Hand-Based Grasping)	2c, 10c (either)
Data Variable Sort	Grab *Data Variable Axis Sphere* widget, drag it around the Time Axis, and release it to apply the new radial arrangement.	Reconfigure	Indirect Widget (via Hand-Based Grasping)	2c
Data Variable Filter	Grab *Data Variable Axis Sphere*, drag it away from the Time Axis until its visual connection disappears ("snaps"), and release it to remove the associated *Data Variable Axis*.	Filter	Indirect Widget (via Hand-Based Grasping)	2c
Time Event Selection	Grab *Time Slice*, and drag it up (forward in time) or down (backward in time) to select a new time event.	Select	Hand-Based Grasping	2c
Time Range Selection	Pinch (index finger and thumb held together) with each hand to unfold a highlighted time range. As long as the hands remain in that posture, the selected time range is updated, allowing to move the hands closer together/further apart for preview. Releasing the pinch applies the time range selection.	Select	Symmetric Bimanual (via Gestural Command)	2c + 2c (together)
Zoom in	With a time range selected, hold both hands with their palms facing each other, and move them apart, "stretching" the selected time range over the entire length of the 3D Radar Chart.	Elaborate	Symmetric Bimanual (via Gestural Command)	11u + 12u (together)
Zoom out	Hold both hands with their palms facing each other, and move them towards each other ("clapping") to apply the previous time range over the entire length of the 3D Radar Chart.	Abstract	Symmetric Bimanual (via Gestural Command)	11u + 12u (together)
Reset	Hold both hands with index fingers pointing upwards, then move index fingers to cross each other ("X"-like posture) to reset the state of the 3D Radar Chart (display entire time-series and all data variables in original arrangement).	Undo	Symmetric Bimanual (via Gestural Command)	5c and 3u (composite)
Pause/ Resume	Hold both hands stretched out in front of the torso in a "stop"-like posture for 1.5 seconds to iterate between two states: (1) Paused (hands semitransparent, no interactions available), and (2) Resumed (hands opaque, all interactions available).	Change Configuration	Symmetric Bimanual (via Gestural Command)	3u + 3u (together)

Table 5.10: Overview of the interaction features available in the third VE iteration, including descriptions on how to perform them via 3D gestural input, as well as their respective data analysis task, interaction technique, and hand posture comfort classifications.

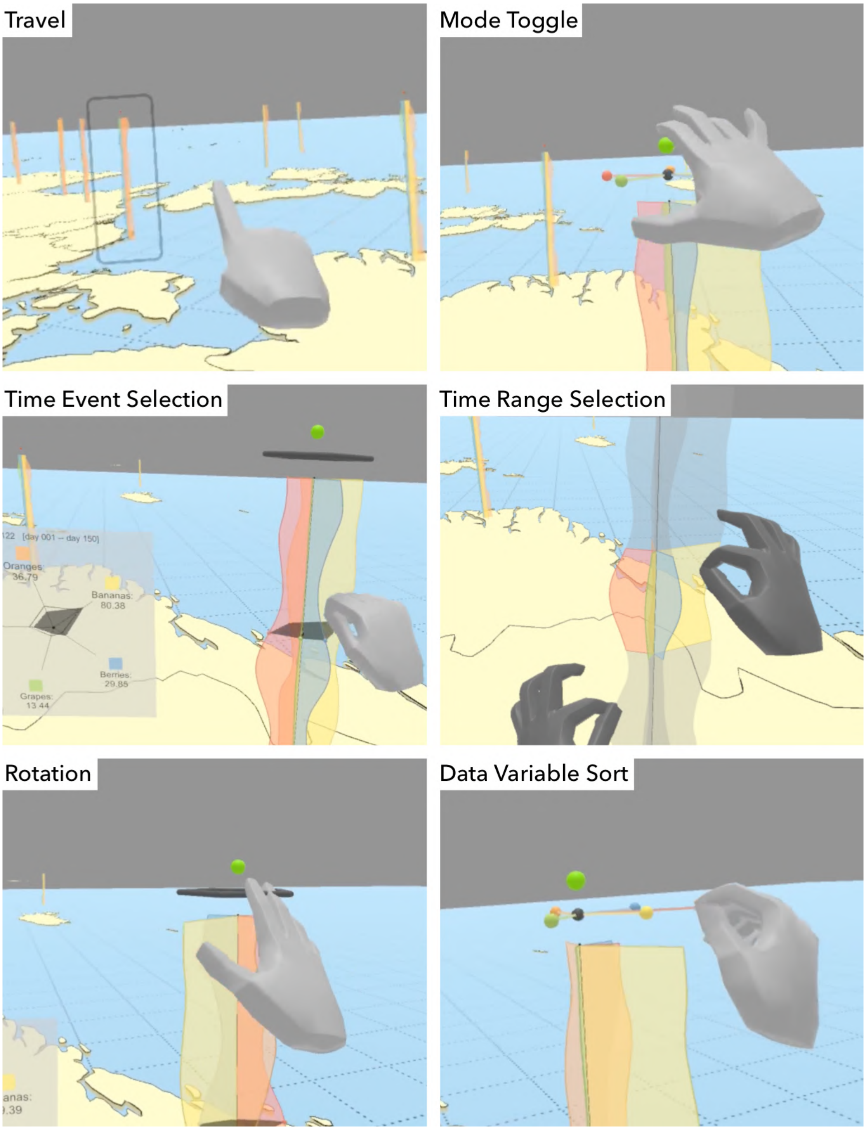


Figure 5.21: Impressions of the various features (see Table 5.10) available in the third VE iteration, from the immersed user's field of view, and utilizing the 3D gestural input modality. **Note**: Part 1 of 3.

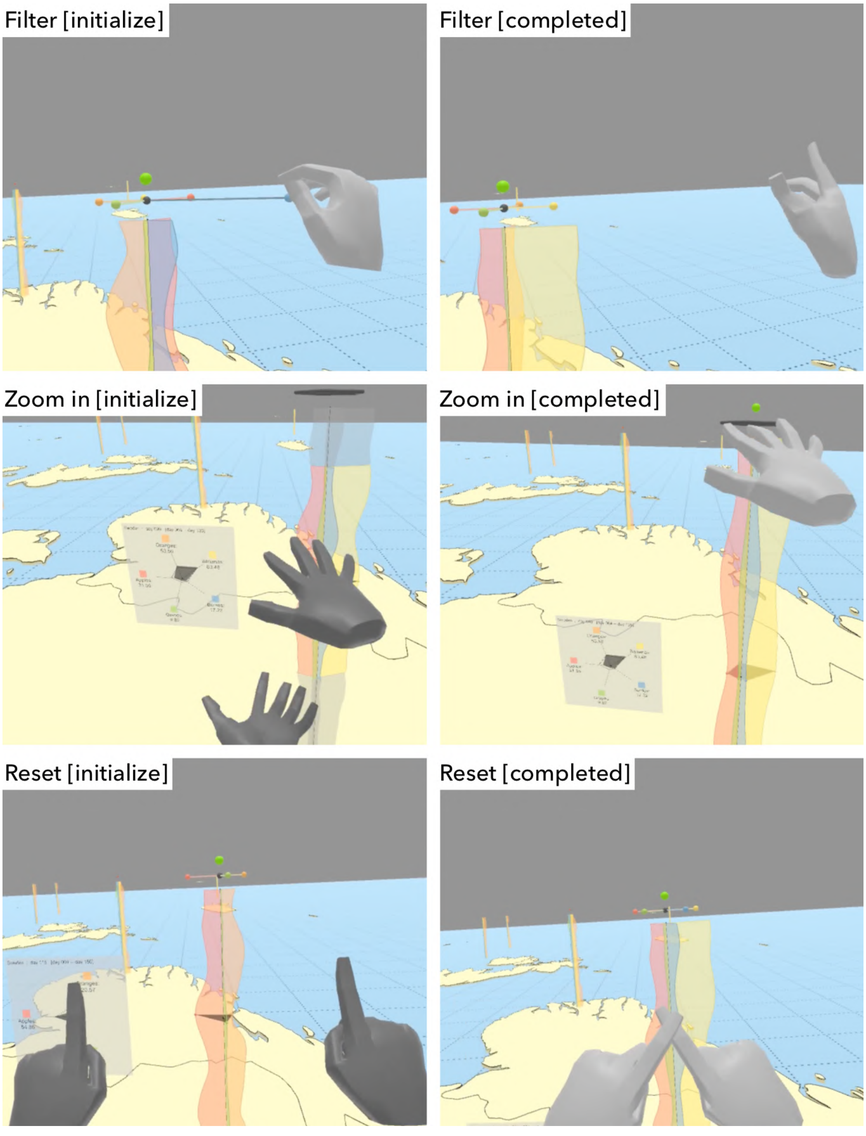


Figure 5.22: Impressions of the various features (see Table 5.10) available in the third VE iteration, from the immersed user's field of view, and utilizing the 3D gestural input modality. **Note**: Part 2 of 3.

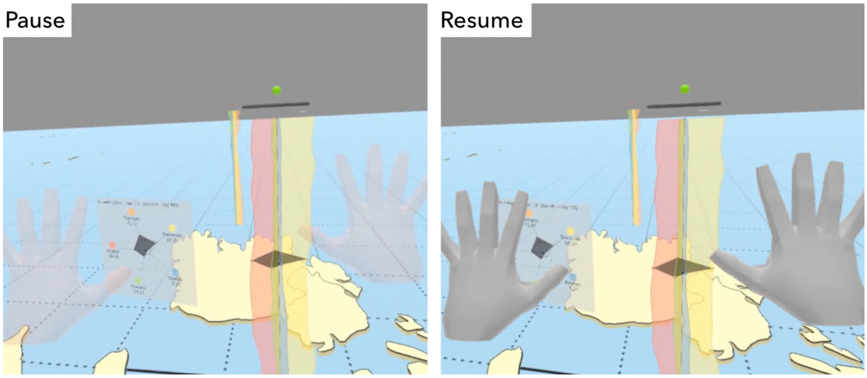


Figure 5.23: Impressions of the various features (see Table 5.10) available in the third VE iteration, from the immersed user's field of view, and utilizing the 3D gestural input modality. **Note**: Part 3 of 3.

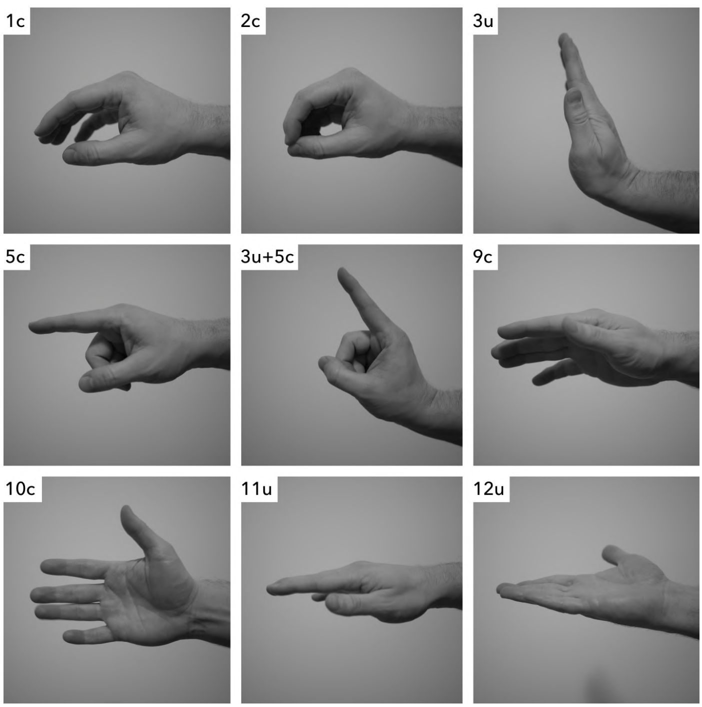


Figure 5.24: An overview of the applied hand posture comfort configurations, in accordance to the report and recommendations by Rempel et al. (2014). **Note**: The applied label coding corresponds to the *comfortable* and *uncomfortable* hand posture configurations as presented in the manuscript by Rempel et al. (2014).

5.5.3 Evaluation 1: Visualization Design Validation

Within the scope of the third VE iteration, this first evaluation is dedicated to the overall validation of the proposed 3D radar chart data entity visualization design as well as its initial interaction design with the aim to confirm the immersed user's ability to make analytical assessments. The results of this evaluation are important to provide impulses and to determine the future direction of this VE iteration, for instance with respect to the visualization design in general as well as through the extension of additional analytical features. The evaluation methodology was centered around a series of tasks that required the participants to make data assessments with respect to the different data variables and their values, for an individual data entity as well as for a comparison between two data entities. Based on the obtained insights from the prior VE iterations (see Sections 5.3 and 5.4) and at this stage in the development of the third VE iteration, only a first set of basic features were provided, implemented through a mixture of various interaction techniques, i.e., hand-based grasping, gestural command, and graphical menus. Measurements with respect to usability, engagement, and task performance allow for the evaluation of the presented visualization and interaction design. The VE setup was the same across all participants. Each study session was conducted in an one-on-one scenario between one participant and one researcher at a time. One study session was aimed to take approximately 45 to 60 minutes to conduct, whereof the participant would spend approximately 25 to 30 minutes immersed in the VE and wearing the HMD. All study sessions were conducted at the VRxAR Labs research group lab at Linnæus University.

5.5.3.1 Physical Study Space

The physical study space generally featured the same setup as in the prior evaluations, for instance as described in Section 5.3.3.1. With respect to the utilized display and input devices, each participant wore a HTC Vive HMD with a Leap Motion Controller attached in front of it to enable 3D gestural input.

5.5.3.2 VE Setup

Within the scope of this first evaluation, the overall data context and scenario were inspired by the NTS corpus, as used in the second VE iteration (see Section 5.4), and thus related to the investigation of language variability on social media in the Nordic region over time. However, with respect to the anticipated task design, a baseline dataset was required to allow for a more accurate task performance assessment compared to using more noisy, real-world data. Thus, time-series data for different data items were generated using the same approach as utilized for the PWt dataset (see Section 5.5). More specifically, the dataset utilized in this evaluation featured (1) two data items, i.e., one representing Sweden and one Denmark, for the subsequent creation of two data entities as 3D radar charts in the VE, (2) six language identifiers as data variables (Swedish, Danish, Norwegian, Finnish, Icelandic, and English) in each data item, thus each 3D radar chart was composed of six data variable axes, (3) time-series data for a total of 50 events on a *per day* basis for each data variable, and (4) frequencies representing the amount of social media posts for each language as data variable values for each time event. Consequently, the generated dataset contained a total of 600 data variable values,[20] with 300 data variables values encoded in each 3D radar chart.

Furthermore, at this stage in the development, a basic set of initial features for the interaction with 3D radar charts in the VE was provided. In particular, and with reference to the feature overview presented in Table 5.10, the following features were available to the participants: *Travel*, *Mode Toggle (Activate/Rotate and Deactivate)*, *Rotation*, *Time Event Selection*, and *Time Range Selection*.[21] Additionally to the presented 3D UI design, some of these features were also available as adapted 2D menus (via graphical menu), following a generally similar approach as applied in the second VE iteration (see Section 5.4.2). The VE provided an adapted 2D graphical menu, attached and juxtaposed to the user's left hand, featuring two virtual buttons for the step-wise selection of the next, respectively previous, time event by remotely moving the 3D radar chart's time slice accordingly. A second adapted 2D graphical menu is attached and juxtaposed to the user's right hand, featuring a virtual button that is dedicated to the time range selection. Pressing the button iterates through three states, the first two of the three being in correspondence with the current position of the time slice: (1) Select the start point for the time range selection, (2) select the end point and thus apply the time range selection, and (3) reset to show the entire dataset. Once a start point is selected, a visual highlight provides a preview of the to be selected time range as user feedback. At this stage in the development, rather than displaying the values of the data variables axes outside of the selected time range in a neutral gray color (see Section 5.5.2), they were "cut off" to be visually excluded.[22] Figure 5.25 presents the described additional graphical menu-based interaction techniques that were available to the participants in this first evaluation.

Finally, at this stage in the VE's development and in regard to the overall system architecture, the support for *User Session Data Transfer* according to the descriptions in Section 4.2.3 was implemented as a proof-of-concept feature. The purpose of this feature is to allow the immersed user in the VE to capture their observations and thoughts, similar to a general note taking process, and later revisit those for the use in other, different data analysis tools (REQ 10

[20] $2\ locations \times 6\ data\ variables \times 50\ time\ events = 600\ data\ variable\ values$

[21] The following interaction features were not available as part of the visualization design validation presented here: *Mode Toggle (Reconfigure/Filter)*, *Data Variable Sort*, *Data Variable Filter*, *Zoom (in/out)*, *Reset*, and *Pause/Resume*. These features were later added and evaluated within the scope of the second evaluation presented in Section 5.5.5.

[22] The neutral gray color data variable axes design, previewing the data values outside of the time range selection, was adopted later within the scope of the second evaluation presented in Section 5.5.5.

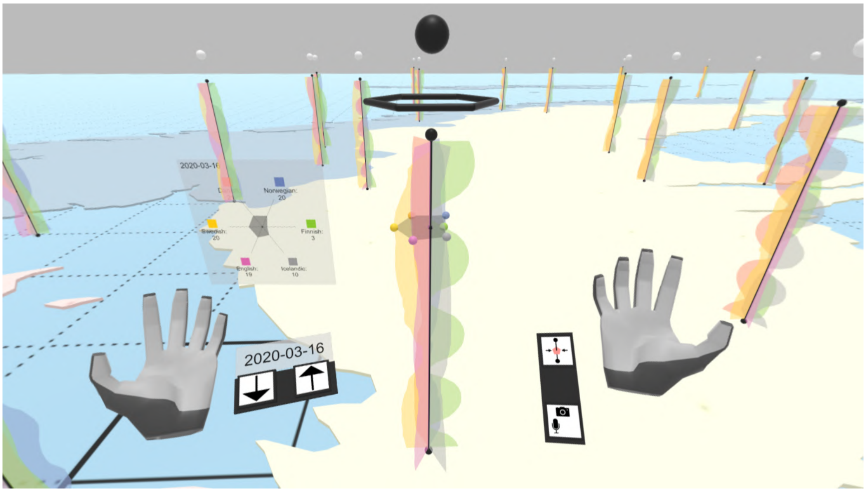


Figure 5.25: An impression of the additional graphical menu interaction techniques available in the first evaluation of the third VE iteration, from the immersed user's field of view, and utilizing the 3D gestural input modality. **Left**: The user's left hand features a juxtaposed adapted 2D graphical menu to select a new time event. **Right**: The user's right hand features a juxtaposed adapted 2D graphical menu to (1) select a new time range (top), and (2) record a virtual note (bottom).

in Table 4.1). The main intent with this experimental feature is to provide a graspable impression of how such a feature could look like, establishing a basis for further discussion and investigation in the future. Within the scope of the presented third VE iteration, the *User Session Data Note Taking and Report* interfaces have been implemented as follows. In the VE, an additional second virtual button is integrated in the adapted 2D graphical menu that is attached and juxtaposed to the user's right hand. Pressing that button triggers two system events in iteration, i.e., (1) *initiate a virtual note*, and (2) *complete a virtual note*. While the virtual note taking process is ongoing, the user can simply *speak aloud* noteworthy observations and thoughts, which are recorded by a microphone connected to the computer system and the VE. Two screenshots are captured based on the user's field of view at the start and the end of the note taking. Once the note taking process has been completed, the recorded audio and image data are transferred as user session data to the respective repository. After the immersive data analysis activity has finished, the user can play back and view their notes in the format of an illustrative report outside the VE on another computer system, for instance via normal web browser. Figure 5.26 illustrates the implemented user session data report interface. Finally, Figure 5.27 provides a representative overview of the system architecture as set up within the scope of this evaluation.

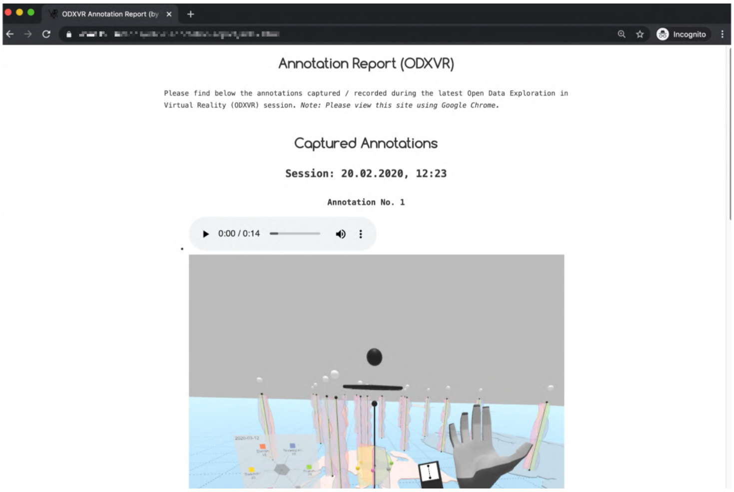


Figure 5.26: A screenshot of the implemented user session data report interface, displaying the virtual notes, in the format of audio and image recordings, taken during the immersive data analysis activity as part of the first evaluation of the third VE iteration. **Note**: The virtual note taking interface implemented in the VE is presented in Figure 5.25.

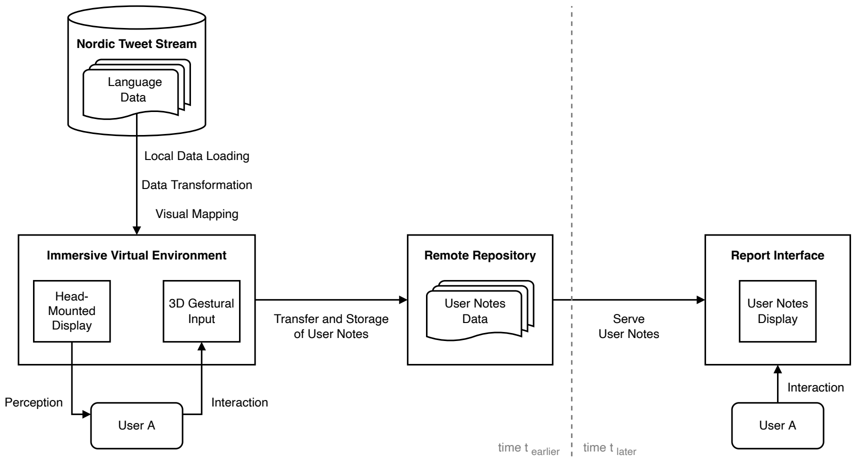


Figure 5.27: Conceptual system overview of the third VE iteration in its first evaluation. Detailed descriptions about the various system components are provided in Section 4.2.

5.5.3.3 Task

Each participant was asked to complete the same series of tasks, involving making observations and interacting with the two 3D radar charts – initially individual, and then together for some comparative tasks. The tasks were designed with the objective to evaluate a participant's ability to make sense of the 3D radar chart visualization design insofar to assess its underlying data. To complete the individual tasks, they had to state their answers spoken aloud to the researcher, who made the respective documentations accordingly. Based on the task's type, an answer corresponds either to a data variable value, the time event as a date (internally corresponding to the *time index* of the time slice), or a time range composed of a start and an end date. In particular, the participants had to complete the following tasks:

- T1: For the first 3D Radar Chart, determine the minimum and maximum values for all data variables.
- T2: For the first 3D Radar Chart, determine the date when all data variables are minimized/maximized simultaneously as much as possible.
- T3: For the first 3D Radar Chart, determine the date when the *Swedish* data variable has the highest value and *Finnish* has the lowest, and vice versa.
- T4a: For the first 3D Radar Chart, determine a time range that contains the most low/high data variable values.
- T4b: For the second 3D Radar Chart, determine a time range that contains the most low/high data variable values.
- T5: Considering both 3D Radar Charts, determine a time range in each that contains the most low/high data variable values and make a comparative assessment.

Additionally, after each time range selection in tasks T4a, T4b, and T5, the participants were asked to use the implemented proof-of-concept note taking feature in the VE to capture some observations. Each participant was encouraged to solve the tasks as best to their ability using the provided interactive features in the VE, using their own strategy and pace with no time constraints.

5.5.3.4 Measures

To obtain an understanding of the participant sample in general, a custom pre-task questionnaire was applied to inquire some demographic information (educational/professional background) and their self-assessed prior experiences with VR technologies in general. The participants' task performance was measured with respect to their ability to make accurate assessments, based on their provided answers to all the individual tasks, independent of the time needed to make these assessments. Furthermore, the System Usability Scale (SUS) and User

Engagement Scale - Short Form (UES-SF) questionnaires were utilized to measure the participants' subjective self-assessment with respect to usability and user engagement. The researcher made observations during the participants' task completion, taking notes accordingly. An informal interview allowed each participant to freely provide some additional feedback based on their own impressions of interacting in the VE, as well as the researcher to pose questions as a follow-up based on the made observations. Foundational aspects of these evaluation methods are described in Section 2.5.1.

5.5.3.5 Study Procedure

The format of the overall study procedure follows conceptually the setup of prior study procedures, for instance as described in Section 5.3.3.5, anticipating an overall study duration of approximately 45 to 60 minutes per study session, dependent on the participant. Each individual study session followed the same procedure of five stages:

1. Introduction (10 min);
2. Warm-up (5 min in the VE);
3. Task (25 min in the VE);
4. Questionnaires (5 min);
5. Interview (10 min).

With the objective to validate the presented visualization design and based on an overall easily understandable data context, i.e., language variability on social media over time, no specific user target group was defined as no specific prior knowledge was required for the participation. The researcher welcomed the participant in the *introduction*, outlined the study procedure, asked them to complete an informed user consent, administered the pre-task questionnaire, and introduced the data context as well as the developed VE. For the latter, the researcher and the participant watched together a brief video demonstration (see Appendix A). After watching the video, each participant was given some *warm-up* time to become familiar with the VE and the involved technologies, i.e., wearing the HMD, walking around in the VR system's calibrated safe interaction area, and interacting in the VE using the provided features via 3D gestural input. Once the participant felt comfortable in the VE, the *task* completion stage was initiated through the researcher, presenting the participants with the tasks as described in Section 5.5.3.3. As in the prior study procedures, it is noteworthy that the warm-up stage used a different representative dummy dataset to prevent a potential insights transfer from the warm-up to the task stage. The researcher in their role as moderator spoke aloud the individual tasks in order, and documented the respective answers from the participants. Additionally, the

Figure 5.28: Immersed participants during their task completion in the first evaluation of the third VE iteration, wearing a HMD and interacting in the VE using 3D gestural input.

researcher also observed the participants during the task completion and took notes accordingly. Once all tasks were completed, the participants were asked to complete, in order, the SUS and the UES-SF *questionnaires*. Finally, a short informal *interview* concluded the study, starting with a brief examination of each respective participant's recorded virtual notes displayed in the format of the user session data report interface in a normal web browser.

5.5.4 Results of Evaluation 1

5.5.4.1 Participants

A total of $n = 15$ participants were recruited for the evaluation to validate the 3D radar chart data entity visualization design. With respect to the sample's demographics, twelve participants categorized their background as *technical*, one as *design* related, one as *pedagogy* related, and one as *humanities* related. Furthermore, seven participants reported to have *no* prior VR experiences, six *a little*, and two *a lot*. Figure 5.28 provides some impressions of several participants during their task completion, immersed and interacting in the VE.

5.5.4.2 Task Assessment

Task T1 Seven participants were able to determine the lowest and highest values for each language data variable, with no mistakes. Five participants made a single mistake, as shown in Figure 5.29 (left). Figure 5.29 (right) indicates that the mean value error of those mistakes was less than one percent.

Task T2 Figure 5.30 (right) indicates that although only a few participants managed to determine the *exact* dates when all data variables were simultaneously minimized or maximized, most were able to come rather close (with two notable outliers for the maximized case). Figure 5.30 (left) indicates that even when not determining the correct date, the sum of the data variable values was appropriately minimized or maximized.

Task T3 It was easier for the participants to determine the date that the Finnish data variable was at its lowest while the Swedish one was at its highest, given that the lowest Finnish value was the same as the overall minimum. It was more difficult for the reverse case, as the lowest Swedish value was much higher than the overall minimum. These results are illustrated in Figure 5.31 (left). However, Figure 5.31 (right) shows that for both cases the participants were able to determine a date very close to the correct one.

Tasks T4a and T4b Given that the participants were not restricted regarding the length of the selected time ranges (neither in T4a/b nor T5), a precise solution could not be pre-determined. However, the individual *dates* (not time ranges) where the sum of all data variables is minimized/maximized, are indicated in all relevant following figures. Figure 5.32 (top) illustrates the time ranges determined initially (T4a/b) for the two data entities, containing the most low and high data variable values. Some participants selected short time ranges, while others selected relatively long ones. Nevertheless, some overall consensus can be identified. Figure 5.32 (bottom) shows the mean sum of the data variable values from the selections in Figure 5.32 (top), indicating that despite the variance in the time range selections themselves, these mean sums satisfy the task instructions, i.e., the value medians are not far from the targets, and the low value selections were smaller than the high value selections.

Task T5 Comparing the latter time range selections in Figure 5.33 (top) with the initial selections in Figure 5.32 (top) indicates similar patterns with less variance around the corresponding dates. Similarly, comparing Figure 5.33 (bottom) with Figure 5.32 (bottom) also indicates overall consistency with less variance.

5.5.4.3 Questionnaires

Figure 5.34 presents the collected self-assessments in regard to user engagement (left) and system usability (right).

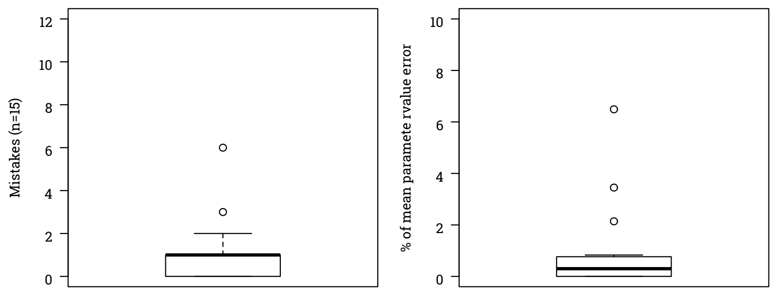


Figure 5.29: Results of task T1, determining the min and max values for each of the six language parameters in the first 3D radar chart. **Left**: Most participants made zero to one mistakes over the twelve questions. **Right**: For those mistakes, the average value error was less than one percent.

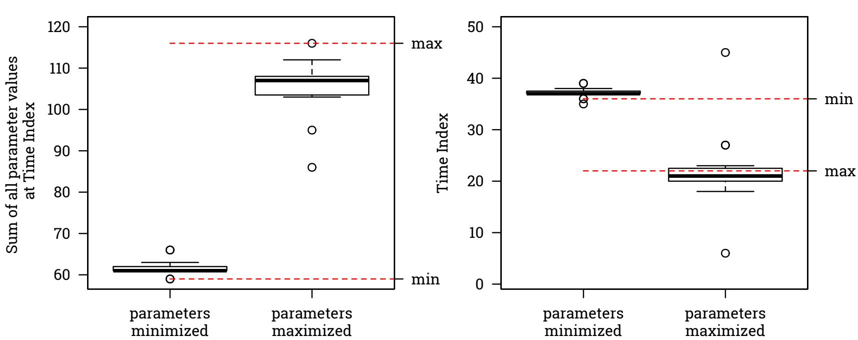


Figure 5.30: Results of task T2, determining the time indexes when all language parameters are minimized or maximized in the first 3D radar chart. **Left**: Sum of all parameters at the time index the participants selected. **Right**: The time indexes that the participants selected. **Note**: For both figures, the correct target values for the min and max cases are indicated in red. In both cases (although there were some outliers for the max case), the participants chose a time index very close to the target, which also closely satisfied the task goal.

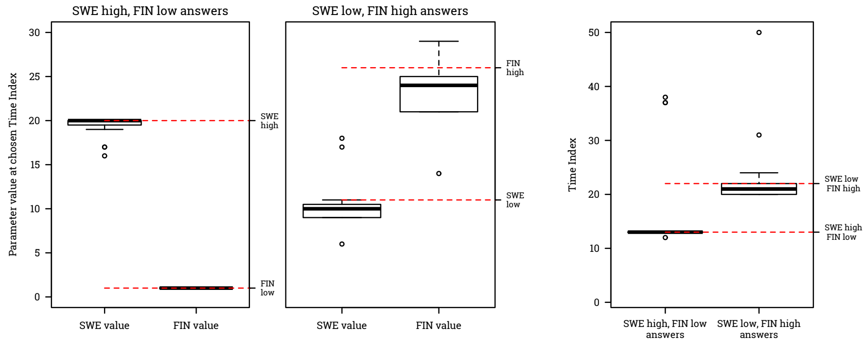


Figure 5.31: Results of task T3, determining the time indexes when the value for Swedish (SWE) is at its lowest while the value for Finnish (FIN) is at the same time at its highest, and vise versa, in the first 3D radar chart. **Left**: Looking at the parameter values, it was easier for the participants to find a close answer for SWE high, FIN low. **Right**: Except a couple of outliers (that still were close to satisfying the task), most participants chose a time index very close to the pre-determined correct answers (indicated by red lines).

5.5.4.4 Observations and Informal Interview

The researcher's notes from the participant observations and the informal interviews were combined and compiled into reoccurring themes. These are presented throughout the following paragraphs.

Interaction Eleven participants were observed to approach the task completion in a noticeable structured and strategic manner, i.e., first examining a 3D radar chart in general by walking around and rotating it in place to get an overview of the time-series data, and then examining specific events (or ranges) in time by strategically manipulating the position of the time slice to obtain further information. One participant emphasized the stereoscopic capabilities of the HMD used to visually perceive the contents in the VE, stating that the *"3D [effect] is actually really good, because you can really see it and get an impression without turning."* Another participant positively stated, *"It is quite interesting to look at the graph like this. I have this negative impression from [3D UIs in] Science Fiction movies."* Two participants seemed to make not much use of the visualization overview, but rather focused on the step-wise movement from time event to time event to extract insights from the information panel. Eight participants explicitly mentioned that the interaction using the 3D gestural input felt *"very natural."* Comments of the participants included, *"I felt like I am already used to it."* (first time VR user), *"It felt very intuitive, very logically, and easy to learn."*, and *"Once I learned the pinching [for the time range selection feature], I felt fairly fluent."*

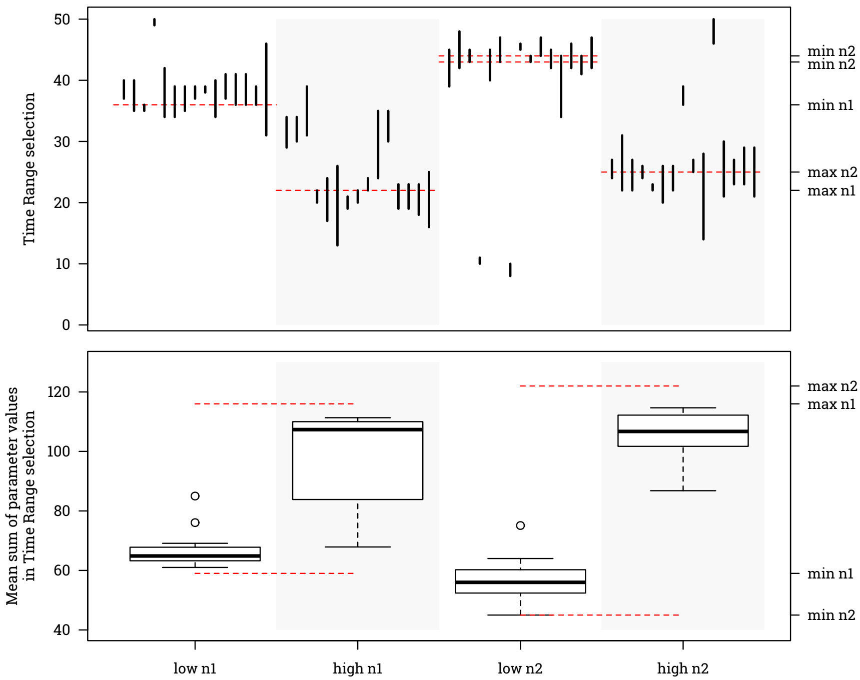


Figure 5.32: Results of tasks T4a and T4b, determining the time ranges that contain most low and high languages parameter values in the first (n1) and second (n2) 3D radar chart. **Top**: The time ranges that the participants selected. The participant order is the same to allow comparisons. The red lines indicate the 3D radar chart's min and max time indexes (n2 features two very close minimums). **Bottom**: The mean sum of parameter values for the time period selections shown in *Top*. The red lines indicate the theoretical possible min and max values for each 3D radar chart.

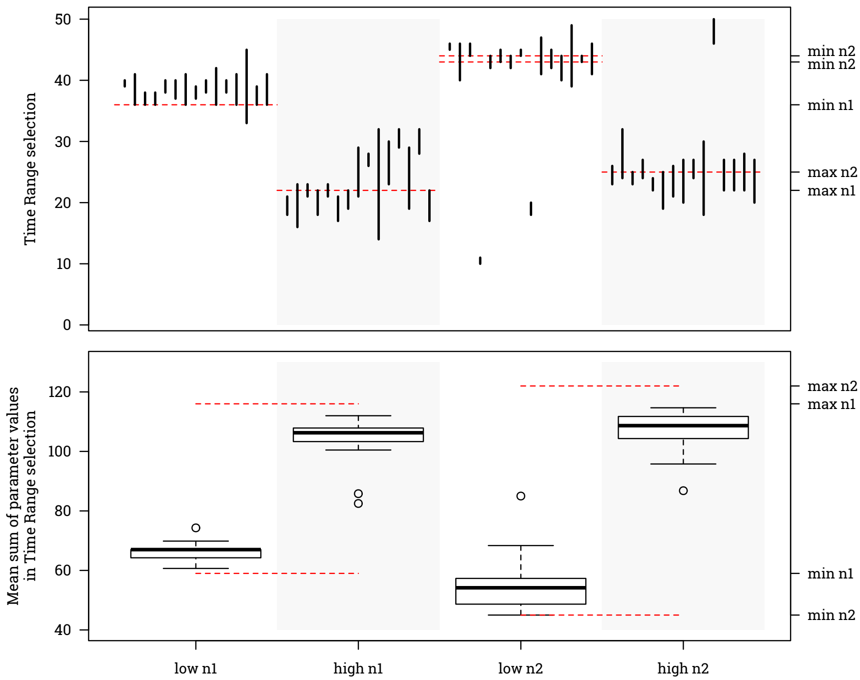


Figure 5.33: Results of task T5, determining a time range selection that includes the most low and high values in each of the two 3D radar charts (n1 and n2) to make a comparative assessment. **Top**: The time ranges that the participants selected. The participant order is the same to allow comparisons. The red lines indicate the 3D radar chart's min and max time indexes (n2 features two very close minimums). **Bottom**: The mean sum of parameter values for the time period selections shown in *Top*. The red lines indicate the theoretical possible min and max values for each 3D radar chart.

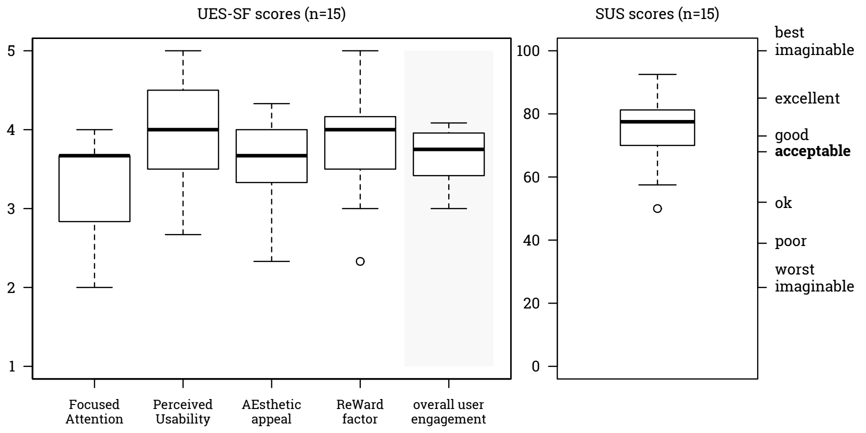


Figure 5.34: Left: Results of the UES-SF, presented according to the different engagement dimensions and the overall user engagement. The median for all individual factor scores (incl. overall engagement) is above average. **Right**: Results of the SUS, presented including the original numerical scale and the supplemental adjective ratings (see Section 2.5.1). The mean value ($M = 75, SD = 11.38$) is well above the *acceptable* threshold.

Three participants noted that moving the time slice sometimes felt "*tricky.*" One participant stated that grabbing and moving the time slice felt "*a bit uncomfortable over time.*" In regard to the two interaction technique alternatives implemented for the time range selection (see Section 5.5.3.2), six participants were observed mainly using the symmetric bimanual gestural command, five mainly used the adapted 2D graphical menu attached to the user's hand, and the remaining four used a mixture of both. Participants who preferred the gestural command technique elaborated on their choice afterwards with comments such as, "*Oh, this [3D gestural input] invites you to do it without buttons.*", and "*Pinching worked fine and felt more natural than using the hand GUI.*" A participant who mainly used the graphical menu argued, "*I preferred the GUI for a more precise selection of the time range.*"

User Session Data Note Taking and Report Ten participants were generally enthusiastic and positive towards the demonstration of the implemented virtual note taking feature (see Section 5.5.3.2) and reviewing the notes of their own study session in the respective report interface via web browser afterwards, finding it "*very useful and meaningful.*" One participant further highlighted that "*Providing an annotation feature is a must for an analysis workflow.*" Five participants were observed to capture very structured and elaborate notes (approximately 1 to 3 minutes in duration per note) of their findings within the tasks.

Analytics One participant explicitly pointed out their ability to visually detect patterns. Another participant found that only limited statistics and textual information in terms of numbers were displayed in the VE. Furthermore, one participant thought that the scaling along one data variable axis (the 2D frequency polygon visualization of the time-series data) looked sometimes very close to each other, while more detailed examination of the actual data variable values revealed that they were more apart than expected.

Technical During five study sessions, minor glitches were observed in the sensory tracking of the used VR hardware (HTC Vive), occasionally causing the participants to experience brief moments of *vection*, i.e., the illusion of self-motion (see Section 2.1.2). The participants stated that it did not impact their experience in the VE in a major capacity. However, it is noteworthy that one participant asked to take a short (approximately 5 minutes) break between tasks T4a and T4b. One participant was observed to unintentionally perform a gestural command to apply a time range selection. Furthermore, one participant was observed wanting to naturally interact with objects outside the Leap Motion Controllers' sensory interaction zone.

5.5.5 Evaluation 2: Uniform 3D Gestural Interface Design

Based on the results and insights obtained throughout the initial evaluation to validate the 3D radar chart visualization design as presented throughout Section 5.5.3, a second empirical evaluation was carried out with the focus on the interaction design. In particular, compared to the initial version, the 3D UI design for the interaction with 3D radar charts in the immersive VE was rigorously iterated in several aspects as follows:

- Beyond the basic set of interactions as described in Section 5.5.3.2, all features as described throughout Section 5.5.2 were made available to the participant. In particular, *Mode Toggle (Reconfigure/Filter)*, *Data Variable Sort*, *Data Variable Filter*, *Zoom (in/out)*, *Reset*, and *Pause/Resume* features were additionally available.
- The design of the 3D UI through 3D gestural input focused on hand-based grasping and gestural command techniques with the objective to provide a uniformed interaction approach, i.e., without the use of any alternative graphical menu-based system control techniques.
- Some overall quality-of-life changes were implemented to further improve the presented 3D radar chart design in general. For instance, the time range selection features now a semitransparent neutrally colored (gray) preview for the time-series data outside the selected time range instead of simply hiding the unselected data, facilitating the user's temporal data analysis.

Based on these changes and extensions, it is possible to contribute with additional insights towards the applied 3D UI design for 3D gestural input within the context of immersive spatio-temporal data interaction. Therefore, this second evaluation was centered around a series of predefined representative interaction tasks in a more walkthrough-like manner compared to the earlier evaluation. Measurements in terms of usability, engagement, and observations enable the assessment of the 3D UI design through the identification of usability issues and remarks on the interaction in the VE. Overall, this second empirical evaluation followed an approach similar to the first one (see Section 5.5.3). The VE setup was identical for all participants. Each study session was conducted in an one-on-one scenario between one participant and one researcher at a time. The conduction of a study session was aimed to take approximately 50 to 60 minutes, whereof the participant would spend approximately 25 to 30 minutes immersed in the VE and wearing the HMD. All study sessions were conducted at the VRxAR Labs research group lab at Linnæus University.

5.5.5.1 Physical Study Space

The setup of the physical study space was for the most parts identical to the prior ones, as described in Section 5.5.3.1. In particular, the researcher had their workstation and the participant their dedicated desk as well as the safe interaction area during their time immersed in the VE. A HTC Vive HMD with a Leap Motion Controller attached in front of it were utilized with respect to the display and interaction technologies. It is noteworthy that this evaluation was conducted during the COVID-19 pandemic, requiring the implementation of some additional practical measures as documented in Section 1.4.

5.5.5.2 VE Setup

For the evaluation of the interaction design, the VE was set up with a representative IA scenario in mind that allows for spatio-temporal data analysis. The artificially generated PWt dataset was utilized to illustrate a scenario in regard to fruit production over time, as introduced in the beginning of Section 5.5. The VE was populated with 39 3D radar charts, each respectively placed at the center of a European country.[23] These countries were displayed as extruded polygons on the floor. Each 3D radar chart featured five data variables (Apples, Oranges, Bananas, Berries, and Grapes), each with a time-series of 150 consecutive time events on a per day basis, thus encoding a total of 750 data variable values.[24] The scenario allows for spatial (European countries) and temporal (time-series at

[23] Although all 39 data entities were displayed and available for interaction in the VE, the task series required the participants to closely interact just with two (see Section 5.5.5.3). The intent of displaying all data entities was to provide an impression of a representative real-world analytical scenario that included all data of a dataset and not just a subset.

[24] $1\ location \times 5\ data\ variables \times 150\ time\ events = 750\ data\ variable\ values$

each country) data analysis featuring an easily understandable data context. All implemented features as presented throughout Section 5.5.2 and Table 5.10 were available to the immersed user in the VE.

5.5.5.3 Task

A series of 31 tasks was created, documented in Table 5.11, comprising a mixture of all implemented features, and structured to be representative of a typical analytical activity, interacting in the VE in a walkthrough-like manner. Furthermore, the tasks featured a mixture of *definite* tasks (e.g., navigate to time event X) and *indefinite* tasks (e.g., select the event X you deem appropriate), enabling the participants to partially make their own data observations and interpretations. All participants started at the same location, i.e., the eastern border of all 3D radar charts. The researcher was responsible for monitoring the task progress of the participants, posing the next task immediately upon completion of the current one. Some tasks required the participants to state a solution, which were communicated spoken aloud to the researcher, who in turn documented these. The same task series order was applied across all participants.

5.5.5.4 Measures

In line with the prior empirical evaluations, a custom pre-task questionnaire was applied with the objective to collect some demographic information (educational/professional background) about the participants and their self-assessed prior experiences with VR technologies. The SUS was utilized to assess the usability of the interactive features provided in the VE. Similarly, the UES-SF was utilized obtain a better understanding about the user engagement with respect to the interactive features. Furthermore, during the task completion, besides taking care of the moderation in general, the researcher also made observations and took notes about the users' interactions. Finally, a brief semi-structured interview was prepared, posing a set of pre-defined questions to each participant, while maintaining the freedom to also pose custom follow-up questions based on their answers and the potentially prior made observations. The prepared interview consisted of an introductory preface and two questions as follows:

> Introductory preface: *3D gestural input, or maybe more commonly referred to as "hand interaction", allows you to interact in a virtual environment, for instance by directly grabbing and manipulating virtual objects, or by making hand postures and gestures that are associated with certain features.*

- Question 1: *How do you feel about hand interaction that allows such an interaction in virtual reality?*
- Question 2: *In regard to the experienced prototype, what is your impression of how the hand interaction was implemented there?*

No.	Task	Feature
T01	Move to *Italy*.	Travel
T02	Move to *Sweden*.	Travel
T03	Activate the 3D Radar Chart at your current location.	Mode Toggle
T04	Navigate to day *120*.	Time Event Selection
T05	Rotate the 3D Radar Chart entirely around its own axis.	Rotation
T06	Name the data variable with the second highest value.	*
T07	Name the data variable with the second lowest value.	*
T08	Select a time range you deem appropriate that contains three peaks in the *Berries* variable.	Time Range Selection
T09	Zoom in into the selected time range.	Zoom (in)
T10	Select a time range you deem appropriate that contains one valley in the *Oranges* variable and one valley in the *Grapes* variable.	Time Range Selection
T11	Zoom in into the selected time range.	Zoom in
T12	Zoom out once.	Zoom out
T13	Switch to the reconfigure and filter mode.	Mode Toggle
T14	Navigate to a time event of your choice that you deem interesting, and briefly describe why it is interesting to you.	Time Event Selection
T15	For the currently selected time event, sort all data variables in ascending order based on their value.	Data Variable Sort
T16	Zoom out once.	Zoom out
T17	Reset the state of the 3D Radar Chart.	Reset
T18	Deactivate the 3D Radar Chart at your current location.	Mode Toggle
T19	Move to *Italy*.	Travel
T20	Activate the 3D Radar Chart at your current location.	Mode Toggle
T21	Switch to the reconfigure and filter mode.	Mode Toggle
T22	Navigate to day *56*.	Time Event Selection
T23	For the currently selected time event, remove all the data variables with a value lower than *20*.	Data Variable Filter
T24	Reset the state of the 3D Radar Chart.	Reset
T25	Pause the 3D hand interaction.	Pause
T26	Attempt to navigate to a different time event.	**
T27	Resume the 3D hand interaction.	Resume
T28	Navigate to a time event of your choice that you deem interesting, and briefly describe why it is interesting to you.	Time Event Selection
T29	Navigate to day *98*.	Time Event Selection
T30	For the currently selected time event, sort all data variables in descending order based on their value.	Data Variable Sort
T31	Deactivate the 3D Radar Chart at your current location.	Mode Toggle

Table 5.11: The series of 31 tasks and their associated interaction features (see Table 5.10) as applied in the second evaluation of the third VE iteration. **Note**: * *ensure understanding of visualization concept* (T06, T07); ** *interaction paused demonstration* (T26). The terms *peak* and *valley* (T08, T10) refer to time ranges that contain high, respectively low, data variable values.

5.5.5.5 Study Procedure

The study procedure followed the same five stages as applied in the first evaluation (see Section 5.5.3.5), anticipating an overall duration of approximately 45 to 60 minutes per study session:

1. Introduction (10 min);
2. Warm-up (5 min in the VE);
3. Task (20 min in the VE);
4. Questionnaires (5 min);
5. Interview (10 min).

The *introduction* began with the participant filling out an informed user consent and a pre-task questionnaire to provide demographic information. Since the chosen data scenario was designed to be easily understandable, there were no specific prior knowledge requirements. The researcher introduced the overall context, scenario, and VE including all its interactive features, utilizing a pre-recorded video (see Appendix A). Each participant was then given some *warm-up* time, allowing them to get comfortable wearing the HMD, and familiarize themselves with the composition of the VE and the 3D gestural input. Once they felt ready, the researcher initiated the *task* stage as described in Section 5.5.5.3. As in the prior evaluations, to prevent a potential insights transfer from warm-up to task stage, different datasets were used. Each participant completed the tasks one by one until all were completed. The researcher observed the participant in the physical real-world space and in the VE from their HMD point of view as mirrored to a screen on the researcher's workstation, and took notes. The researcher read aloud the individual tasks, and noted the participant's answers. Once all tasks were completed, the participant was asked to complete, in order, the SUS and UES-SF *questionnaires*. Finally, the semi-structured *interview* was conducted, after which the participant was thanked and sent off.

5.5.6 Results of Evaluation 2

5.5.6.1 Participants

A total of $n = 12$ participants were recruited, who reported a variety of backgrounds, i.e., 5 *Computer and Information Science*, 5 *Linguistics and Language Studies*, and 2 *Forestry and Wood Technology*. Eight participants stated *a little*, three *average*, and one *a lot* prior experiences with VR technologies. None of them reported any visual perception issues when asked during the warm-up phase, for instance in regard to their ability to differentiate the five data variable axes.[25]

[25] The applied color coding throughout the third VE iteration adopted recommendations by: Cynthia Brewer, Mark Harrower and The Pennsylvania State University. ColorBrewer: Color Advice for Maps. Retrieved June 1, 2022, from `https://colorbrewer2.org/`

5.5.6.2 Task Assessment

All participants were able to successfully complete the task series, as presented in Section 5.5.5.3, and provide correct answers as pre-determined, or otherwise contextually appropriate based on their own selection choices. Based on the analysis of the collected log files, the task completion times averaged with $M = 13.95\ min$ ($SD = 3.15\ min$; tasks were presented in a swift manner without noticeable breaks; participants were instructed to complete them at their own pace). When the participants were asked to select a time event that they deemed as "interesting" and to briefly describe why (T14 and T28), they made their own observations, generally ending up selecting time events that featured either comparatively high or low data variable values. These time events were visually noticeable, allowing them to make comparisons and to begin speculating for potential reasons. Descriptions by the participants (P) included:

- "*Berries are very low, while Bananas and Oranges are high. This could indicate a different season of the year, thus the values across the different dimensions representing a change of season.*" (T14, P1, day 75)
- "*Oranges and Bananas appear to be very high, while Grapes and Apples are very low. It seems like there is a relationship between those, maybe a seasonal event.*" (T14, P7, day 132)
- "*Berries are very low, and then increasing afterwards. This is interesting, what is happening here?*" (T14, P9, day 72)
- "*Oranges and Bananas are very high, while the others are very low. This looks like opposite trends.*" (T14, P10, day 126)
- "*Oranges appear to be very high compared to the time series before and after the selected time event, maybe this could be because of a seasonal effect.*" (T28, P2, day 58)
- "*The values ... seem to be at their dimension's average at the same time. It's a perfect overlap.*" (T28, P4, day 86)
- "*Peak in the Grapes dimension, and it seems that Grapes are generally rather low overall compared to all other dimensions, therefore this is interesting.*" (T28, P6, day 133)
- "*Grapes are high and we are in Italy, so this should be great for the wine season.*" (T28, P12, day 145)

5.5.6.3 Questionnaires

Figure 5.35 presents the results of the reported self-assessments by the participants in regard to user engagement (left) and system usability (right).

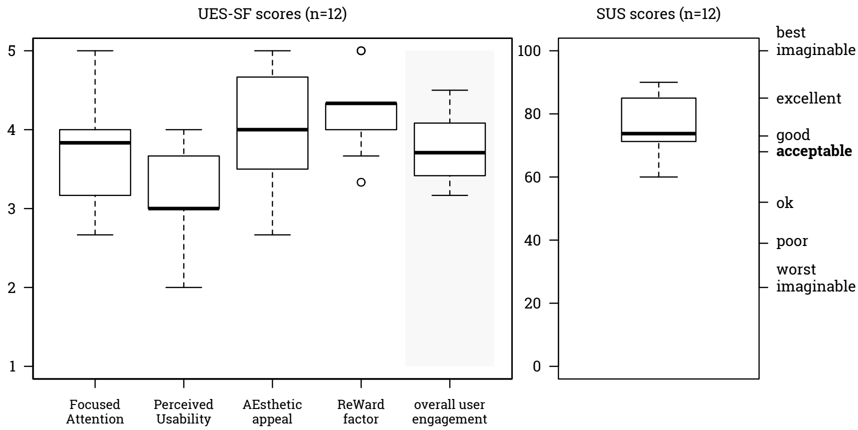


Figure 5.35: Left: Results of the UES-SF, presented according to the different engagement dimensions and the overall user engagement. The median for all individual factor scores (incl. overall engagement) is above average. **Right**: Results of the SUS, presented including the original numerical scale and the supplemental adjective ratings (see Section 2.5.1). The mean value ($M = 76.25, SD = 9.62$) is well above the *acceptable* threshold.

5.5.6.4 Observations

Generally, all the participants appeared to understand the concept and learn the operation of the implemented features rather quickly, enabling them to interact in the VE seemingly natural and in an enjoyable manner. Nevertheless, some interesting observations were made throughout the various study sessions, thematically structured and presented in the following paragraphs.

Usability Issues Most noticeably, the time event selection by grabbing, dragging, and releasing a 3D radar chart's time slice appeared comparatively sensitive during the interaction's conclusion. The participants had seemingly no problems initiating and continuing the grabbing mechanic, navigating back and forth in time while simultaneously interpreting the data and reading the updated labels in the juxtaposed information panel. However, when asked to select a specific time event (T04, T22, and T29), at times the time slice would snap to an adjacent time event during the release of the hand-based grasping. By opening up one's hand, the hand tracking would first interpret a time slice movement before concluding the grasping gesture and discontinuing the time event selection. In these cases, participants had to attempt this interaction more than once until the time slice remained in the desired position. Such reoccurring observations were made during nine study sessions.

The zoom in/out gestural command seemed to require the comparatively longest learning phase. Depending on a participant's hand placement, the tracking sensor would sometimes discontinue detecting the lower hand, as it appeared to be (partially) occluded by the hand above. Once the participants appeared to have gotten a more cautious understanding and feeling of the hand tracking, they were able to perform these gestural commands seemingly fluent. One participant was observed repeatedly attempting the gestural commands in their reverse concept, i.e., moving their hands *together* to zoom in, and moving their hands *apart* to zoom out.

Some instances of unintentional commands were observed, i.e., triggering a feature through the 3D gestural input without the explicit intent. Most noticeably, this occurred during intended mode toggle interactions, resulting in unintended travel. In these cases, rather than touching the 3D radar chart's mode toggle widget with a hand and all fingers extended, the participant would attempt to touch it with only the index finger extended, similar to a "poking" hand posture. This however was in conflict with the same hand posture configuration designed and implemented as part of the travel feature's gestural command (*gaze suggests, point confirms*), thus resulting in an unintentional position transition to another data entity instead.

General Operation and Interaction To make data observations, the participants appeared to use a balanced mixture of actively moving around a 3D radar chart and in-place rotation using its rotation handle widget. Even though not explicitly asked, some participants made on own accord noticeable use of various implemented features to assist them with their task solving process, for instance sorting the data variables before selecting a time range (T08 and T10) or filtering out proclaimed "*uninteresting*" data variables (T14 and T28). The participants were asked to sort the data variables in ascending (T15) and descending (T30) order. However, at no point were they told what these orders mean within the presented context – this was intended by design to observe how the participants interpreted these tasks. The majority associated *ascending* with a *clockwise* and *descending* with a *counter-clockwise* radial arrangement of the data variables with respect to their order in the 3D radar chart's information panel. A few participants appeared rather self-critical with their perceived performance operating the 3D UI, but became seemingly more confident over time as they got "*a better feeling*" for the sensory hand tracking. Sometimes, participants attempted to perform gestural commands rather quickly, while their hands were not yet in the tracking sensor's interaction zone. Although their gestural interaction was correct in concept, the tracking sensor appeared too slow in its initial hand detection, thus preventing them from the practical execution of the respective interaction. This was frequently observed for those features classified as gestural commands, but not so much for the hand-based grasping ones.

5.5.6.5 Interview

At the end of the study session, the researcher conducted a brief semi-structured interview with each participant, inquiring their feedback in regard to hand interaction in a VE using VR in general (Question 1), and how the hand interaction was implemented in the presented VE (Question 2), as described in Section 5.5.5.4. Their feedback can be summarized as follows.

General Hand Interaction Overall, the participants expressed a rather positive attitude towards the concept of hand interaction in VEs. They thought that it has the potential to allow for very natural and intuitive interaction mechanisms. Some of them mentioned their appreciation that no additional sensors needed to be physically attached to one's hands. One participant expressed minor concerns about imprecise command recognition, i.e., when an interaction is not triggered, even though correct in concept, it might make the user feel insecure, as it is difficult to determine whether the detection problem was due to them or the system. Four participants explicitly expressed their appreciation to simply use their hands instead of physical controllers that can "*sometimes feel weird for the interaction, as one is grabbing a controller and the controller is grabbing a virtual object*," therefore having some kind of middle layer impression – which, according to them, is not the case with hand interaction.

3D UI in the presented VE In regard to their impression about the hand interaction implemented in the presented third VE iteration, the participants were generally positive about the provided features. The majority stated that the 3D UI felt very natural and easy to operate once one had learned all possibilities. They acknowledged their impression of learning the various features quickly, with one participant elaborating that it felt like "*riding a bike*" at that stage. Some of them noted themselves that the 3D UI featured logical and coherent analogies for the different hand postures and gestures. A few were genuinely surprised that seemingly many features relied on the utilization of both hands simultaneously, expecting more one-handed gestures. Participants also addressed some of the encountered usability issues, most dominantly mentioning that the precise time slice placement appeared to be "*fairly tricky*" at times, as described in Section 5.5.6.4, making it feel as if the hand tracking was too sensitive in these instances. Some also reflected on experienced unintentional gestural commands.

5.5.7 Use Case: Forestry Data

Similarly in intent as the previously presented use case for the second VE iteration (see Section 5.4.4), a proof-of-concept prototype was also developed to demonstrate the concept of the 3D radar chart data entity visualization and interaction design in a different context compared to the main PWt one. Thus, a use case was conceptualized around the Swedish National Forest Inventory.[26] The inventory contains data that allows the analysis of the standing volume, based on stems for various tree species, in the forests of Sweden from a spatio-temporal perspective, i.e., based on counties in Sweden for the period from 1955 to 2017. Such an analysis is relevant, among others, within the context of forest management to make relevant decisions in an informed and strategic manner.

Figure 5.36 provides some impressions of the immersive data analysis environment that was set up as follows. In line with the overall approach of displaying relevant geographical areas on the virtual floor, extruded surfaces representing the 21 counties in Sweden are visualized utilizing data provided by Statistics Sweden – following a similar approach as described in Section 5.4.4. The inventory contains data differentiating between three categories of tree species, i.e., *scots pine*, *norway spruce*, and all *deciduous tree species*. For each of these groups, data variables exist to indicate the amount of counted tree stems (in millions) across different stem diameter classes. Thus, three individual 3D radar charts were set up for each county, one to represent each tree category, each featuring five data variables axes where each axis represents the temporal data from 1955 to 2017 for the respective stem diameter class. One additional 3D radar chart was created for each county, combining all the data to visualize the *total* amount of tree stems accordingly. The 3D radar chart representing the total data is displayed by default, placed at the centroid of its respective county. The user can switch on demand between the display of either that 3D radar chart or the three others for the individual tree species. Interaction in the immersive VE is generally possible in accordance to the features as described throughout Section 5.5.2.

In addition to the immersive interaction mode, a non-immersive display and interaction variant was developed within the scope of this proof-of-concept prototype.[27] Building on the capabilities of the Unity cross-platform game engine (see Section 4.2.5), an interactive version of the 3D visualization was created that utilizes display through a normal monitor and interaction through keyboard and pointer (mouse) input. All the interactive features as described in Section 5.5.2 were mapped to various keyboard and pointer commands. For instance, the user can use the pointer to click on the extruded surface of a county to make a respective selection, and then use keyboard input to interact with the associated 3D radar chart, for instance utilizing the up and down keys to move the time slice up and down to select a new time event. In addition to the other keyboard commands, the user could also hold the spacebar button to automatically orbit around the selected 3D radar chart. Furthermore, various pointer commands

[26] Sveriges lantbruksuniversitet (SLU). The Swedish National Forest Inventory (NFI). Retrieved June 1, 2022, from `https://www.slu.se/en/Collaborative-Centres-and-Projects/the-swedish-national-forest-inventory/`

[27] The non-immersive display and interaction variant was developed to test the visualization environment outside the usual VR setup and laboratory environment due to the at the time (fall 2020) ongoing work-from-home recommendations in response to the global COVID-19 pandemic.

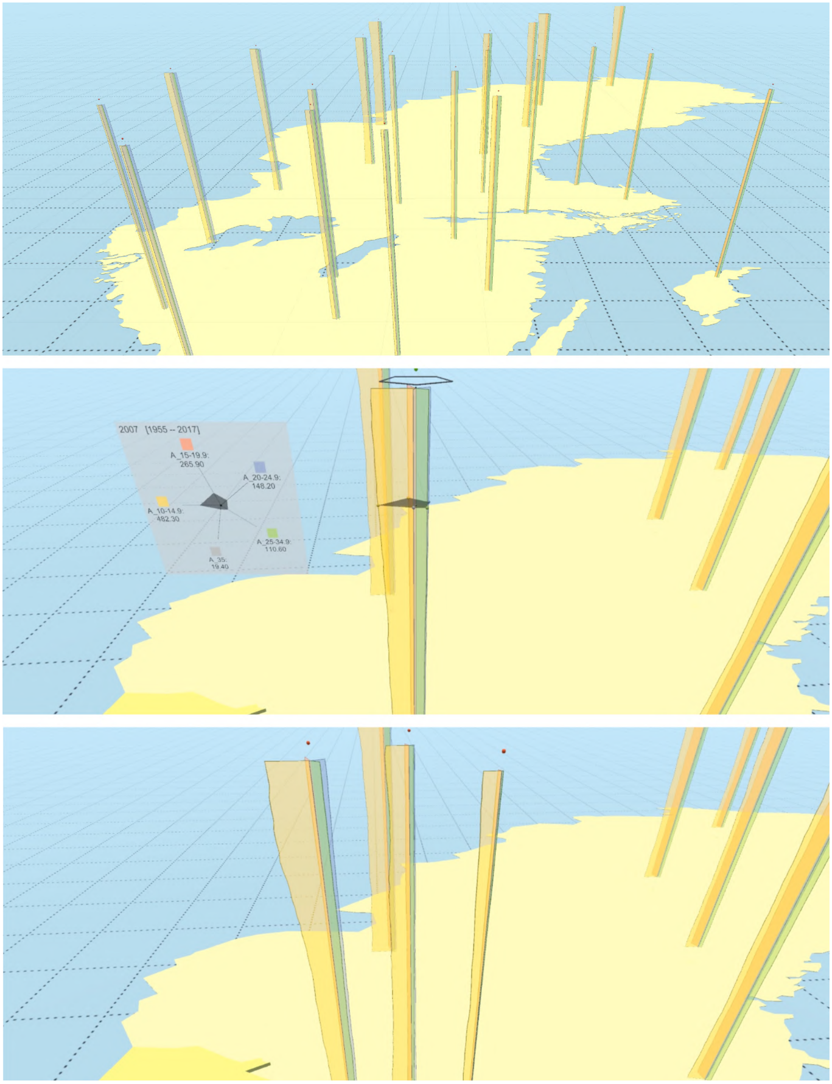


Figure 5.36: Impressions of the third VE iteration, demonstrating the immersive data analysis use case for the Forestry Data. **Top**: The VE composition, from an angled top down view to provide an overall impression. **Middle**: The immersed user's field of view during a details-on-demand temporal exploration with the data entity selected that represents the *total* amount of stems in Dalarna County. **Bottom**: The immersed user's field of view during an overview-like spatial exploration, examining the individual tree species data in Dalarna County (from left to right: *scots pine*, *norway spruce*, *deciduous tree species*).

were made available to support free camera movement and rotation, enabling the user to change the displayed field of view. While inherently different compared to the perception and interaction in the immersive VE, the non-immersive variant can be used to get an overall conceptual impression and understanding of the data entity visualization design.[28]

5.5.8 Use Case: Urban Climate Data

In addition to the use case of utilizing forestry data to demonstrate the third VE iteration, as described in Section 5.5.7, a second proof-of-concept prototype was developed to further illustrate the practical application of 3D radar charts in yet another data context. In particular, the intention with this prototype was to explore the visualization of urban climate data in an immersive data analysis environment. Urban climate data commonly feature spatio-temporal characteristics, such as various climate related data variables over time as well as geolocation information about where these climate measurements were collected – typically, as the name suggests, in urban areas. Among others, the analysis of urban climate data is relevant to identify heat-exposed areas and other aspects of local conditions, which are important with respect to the strategical decision making within the context of urban planning (Vuckovic et al., 2022; Rød and Maarse, 2021). In contrast to climate data provided by general weather stations that commonly cover broader areas, urban climate data typically rely on sensor measurements as voluntary contributions through citizens to build up a desired higher resolution dataset that comprises data from more sensors at various locations (Navarra et al., 2021; Neset et al., 2021; Rød and Maarse, 2021). Adopting the overall IA approach in the format of 3D radar charts as presented in Sections 5.5.1 and 5.5.2, it is possible to create a respective situated visualization, i.e., visualizing and displaying the urban climate data contextually where it was collected (Bressa et al., 2022; Thomas et al., 2018).

Figure 5.37 provides some impressions of the proof-of-concept prototype that was developed in a collaborative exploratory effort involving researchers from Linköping University and Linnæus University, Sweden. The overall third VE iteration approach was utilized to visualize urban climate data within the context of the city of Norrköping, Sweden. Individual 3D radar charts were created, each representing a unique climate sensor, placed at their respective geolocation in the immersive VE. Each 3D radar chart features three data variable axes, visualizing temporal data for *temperature*, *humidity*, and *pressure* for the time period of January 2019 to March 2021 on a per day basis. Data provided by

[28] The described proof-of-concept prototype was presented at *The 6th Big Data Conference at Linnæus University* (Dec 2020) as part of the talk *Interdisciplinary Exploration of Forestry Data Using Machine Learning and Immersive Visualization*.

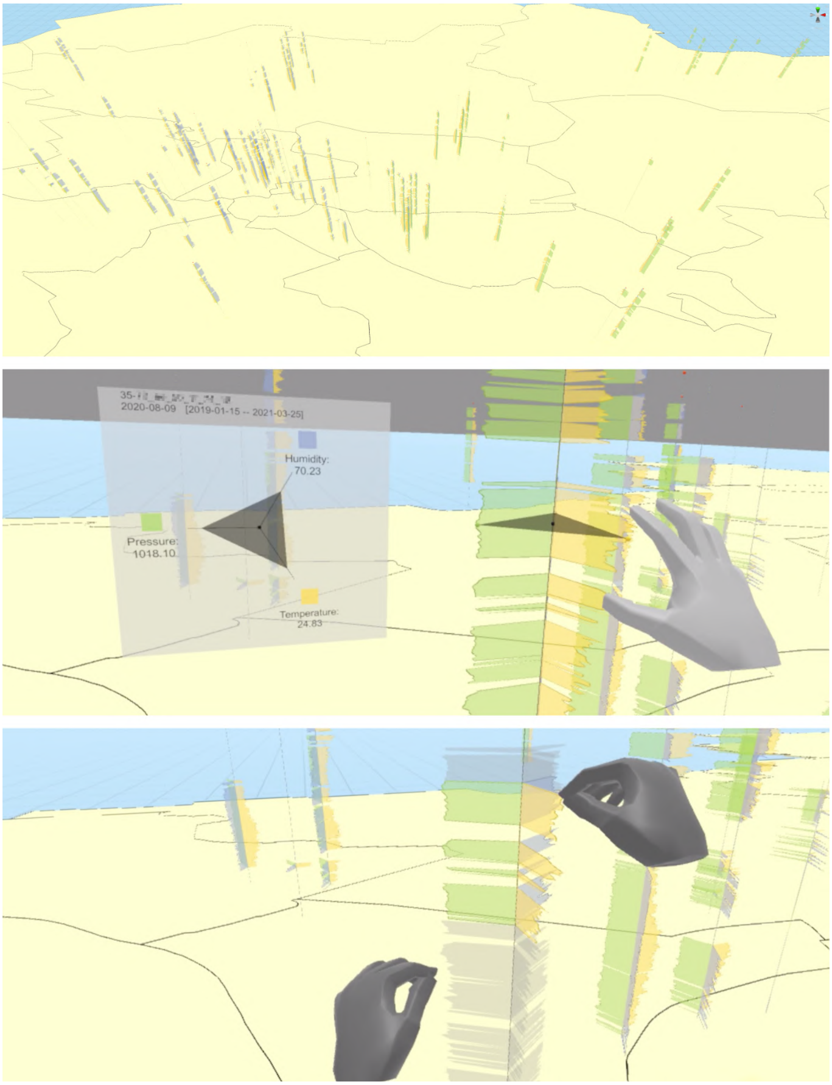


Figure 5.37: Impressions of the third VE iteration, demonstrating the immersive data analysis use case for the Urban Climate Data. **Top**: The VE composition, from an angled top down view to provide an overall impression. **Middle**: The immersed user's field of view during a details-on-demand temporal exploration with the data entity selected that represents *Sensor 35*, displaying the information panel and attempting to grab the 3D radar chart's time slice. **Bottom**: The immersed user's field of view while selecting a time range at the data entity that represents *Sensor 35*.

The Swedish Mapping, Cadastral and Land Registration Authority[29] was used to visualize and display the various districts of the the city of Norrköping as extruded surfaces on the virtual floor, following a similar approach as applied in the prior proof-of-concept prototypes (see Section 5.4.4 and 5.5.7). The user is able to explore the urban climate data as sensor measurements in the immersive VE with respect to their spatial and temporal contexts, extracting insights accordingly. In the future, it is conceivable to visualize such urban climate data juxtaposed with additional information and other visualization artifacts in an immersive VE, for instance by providing a respective visualization about the physical real-world buildings to represent the urban composition of the city of Norrköping. For instance, following a similar approach as presented by Alatalo et al. (2016), who built an interactive immersive 3D representation of the city of Oulu, Finland, urban climate data visualizations could be placed in-situ at a sensor's respective geolocation, enabling further situated visualization and data analysis accordingly (Bressa et al., 2022; Thomas et al., 2018).

5.5.9 Discussion

The third and, within the scope of this thesis, final VE iteration is centered around the 3D radar chart visualization design to represent spatio-temporal data entities that can be interacted with through 3D gestural input. For this purpose, two empirical evaluations were conducted. The first to validate the visualization design in general, including a first set of basic interactive features. Based on the insights from that evaluation, additional features to support analytical tasks in the immersive VE were implemented, following a uniform 3D gestural interface design that focused on hand-based grasping and gestural command techniques without alternative graphical menu-based system control techniques. The extended 3D gestural interface was likewise empirically evaluated. The results of both evaluations, as presented in Sections 5.5.4 and 5.5.6, allow for discussion and various reflections with respect to the presented utilization of 3D radar charts for IA purposes.

5.5.9.1 Visualization Design Validation

The 3D radar chart data entity visualization design, as described in detail throughout Section 5.5.1, enabled the visual encoding of temporal data across multiple data variables not just for one time event, but multiple ones, thus representing a time series. The conducted empirical evaluation allowed participants themselves to analyze data in a scenario of investigating language variability on social media in the Nordic region – a scenario informed by the overall data context of the

[29]Lantmäteriet. Distriktsindelning Nedladdning, vektor. Retrieved June 1, 2022, from `https://www.lantmateriet.se/sv/Kartor-och-geografisk-information/geodataprodukter/produktlista/distriktsindelning-nedladdning-vektor/`

second VE iteration (see Section 5.4), and based on an artificially generated time-series dataset. Based on a series of explorative data analysis tasks that the participants were asked the complete, it was possible to collect various measurements, for instance in regard to usability, user engagement, and task completion. Using the collected quantitative data as well as the additional qualitative data from observations and informal interviews, the overall 3D radar chart data entity visualization design could be validated.

Usability The analysis of the reported SUS scores indicates an overall *good* usability, suggesting that the participants generally accepted the concept and interaction design in the VE, and in turn that the presented approach is usable within an IA context. Some of the lower SUS scores may be attributed to the observed sensory tracking issues in regard to both the HMD and the 3D gestural input, and to some users expressing that the gestural command detection and subsequent interaction not always worked on the first try. Similar technology-related observations inherent from the utilization of such input device types were also made in prior evaluations, for instance as described in Section 5.3.5.1. Nevertheless, all participants were able to learn the various interaction features in a comparatively short amount of time, i.e., during an approximately 5 to 10 minutes warm-up phase, with the majority stating that they found the interaction very natural. This can be regarded particularly positive, especially considering that most of the participants had either no prior or just a little experience with VR technologies prior to the evaluation.

User Engagement The analysis of the received UES-SF assessments indicates an overall *above average* user engagement, suggesting that the participants felt engaged in the VE, interacting and making assessments to complete the given analytical tasks. This result can be supported by the general enthusiasm the majority of the participants expressed. Most of them engaged in the task completion process quite motivated, approaching the search for a solution in a very strategic manner, making use of the various features available in the VE, supported through the capabilities of the immersive VR technologies, i.e., the HMD enabling stereoscopic viewing and the 3D gestural input for hand interaction in the VE. Investigating the results of the individual dimensions of the UES-SF questionnaire, it becomes apparent that the reported perceived usability is in line with the results of the SUS. Even though the researcher provided no indications about the participants' task performance, the majority of participants reported a rewarding experience. Even though the result for the reported aesthetic appeal can be considered generally okay, it also indicates potential for further improvements. For instance, as part of the second evaluation (see Section 5.5.5), various minor changes were applied to improve the 3D radar chart design, such as for instance adjusting the size of various interactive widgets and other elements in the UI, resulting in a subjectively "cleaner" visual design. In comparison, the aesthetic appeal was rated more positively in the second evaluation compared to the first one (see Figures 5.34

and 5.35). The participants reported mixed results with respect to the focused attention, that in retrospect may be attributed to the fact that they had to carefully listen to the researcher phrasing the tasks as well as subsequently reporting back their results – all spoken aloud in a verbal dialog between researcher and participant. Arguably, this may have taken some focus away from their experience immersed in the VE.

Task Assessment Considering the overall motivation of providing a tool that would support the user with their explorative analysis of spatio-temporal data in general, and in the case of the 3D radar chart with particular focus on the temporal data aspects, the reported task completion results can be interpreted as satisfactory. They point towards the participants' ability to make use of the presented data entity visualization design and its provided interaction features, allowing them to complete representative data analysis tasks in a satisfying manner, both in regard to the actual task completion as well as a subsequent rewarding feeling, as indicated by the comparatively high rated UES-SF reward factor (see Figure 5.34). The participants were able to investigate different time events in the data to find appropriate solutions, both with respect to individual dates as well as time ranges. It is noteworthy that the tasks requesting a time range as answer did not specify any duration details, leaving it up to the participant to answer what they thought was most appropriate. Naturally, some reported shorter time ranges as answers, and some reported longer ones. Nevertheless, certain trends in the answers among all participants could be identified, indicating that the majority provided appropriate solutions for the given tasks. Overall, their successful task completions indicate that the participants were able to make sense of the 3D radar chart visualization concept, and thus make adequate assessments across the different data variables and time events.

Interaction The provided set of basic interactive features, implemented using 3D gestural input and a mixture of hand-based grasping, gestural command, and graphical menu interaction techniques as described in Section 5.5.3.2, were generally perceived positively. Except for a few minor disruptions caused through technical issues, that are arguably normal and to be expected given the underlying nature and concepts of the applied sensory hardware, the participants were able to naturally interact in the immersive VE. The majority of them appreciated being able to directly manipulate the time slice by grasping and moving it up and down within the 3D radar chart in order to inspect different time events. At the same time, it was interesting to observe how some participants made targeted use of the adapted 2D graphical menu attached to their hand as an alternative, i.e., they moved the time slice to an area of interest using the hand-based grasping, and then utilized the two-button hand menu to iteratively move forward and backward in time to examine consecutive time events. It seems that the participants utilized the hand-based grasping to move quickly in time, and then utilize the graphical menu for a more detailed step-by-step analysis of the time-series data. All participants were able to make use of the implemented interaction alternatives for the time range selection to solve the respective tasks, arguably approving the usability for both, independent of their overall mixed preferences. Some participants enthusiastically preferred the symmetric bimanual gestural command using a pinching hand posture with each of their hands, stating that the 3D gestural input basically offers itself for such interactions. Others however preferred to use the adapted 2D graphical menu, arguably because it enabled them to be more precise with the respective time slice placement, and thus to be more precise in regard to what events to include in and exclude from their time range selection. Finally, some participants were observed using a mixture of both alternatives, sometimes trying one, and other times the other. For the task completion, the researcher did not provide any indications of whether the participant should use one interaction alternative over the other, but let it open for the participant to decide. This was entirely exploratory to informally investigate, on the side, a potential preference for one interaction technique over the other. At that stage, the results did not indicate a clear preference. However, a slight trend towards the ability to directly manipulate a 3D radar chart in the VE using hand-based grasping and gestural commands was noticed based on some of participants' enthusiasm and feedback. Consequently, the decision was made to further explore this matter by practically removing the graphical menu-based UI elements and focusing on an interaction design that is centered exclusively around hand-based grasping and gestural command techniques in the second 3D radar chart evaluation (see Section 5.5.5). The discussion and reflections on this interaction design decision are subsequently presented in Section 5.5.9.2.

User Session Data Note Taking and Report The implemented proof-of-concept virtual note taking and report feature received positive feedback. The feature was implemented to both conceptually and practically demonstrate the usefulness of user session data transfer as introduced in Section 4.2.3. Each participant was asked to record three virtual notes during their task completion, that were later briefly examined as part of the informal interview using the respective report interface within a normal web browser outside the VE (see Figure 5.26). The participants were in agreement, assessing the feature as a meaningful addition to such a type of VE, and especially useful within the context of IA. Some of the participants even went a step further and categorized the feature as "*necessary*" for serious future IA applications. At this stage, it was particularly interesting to observe the different note taking strategies of the participants, with some just briefly stating a to the point observation, while others recorded very elaborate explanations of their observations including hypotheses of why certain phenomena in the data might be as they are. Consequently, the ability to take such elaborate observations, recorded as virtual notes in the conceptual format of user session data from the immersive analysis activity, and use them as input and guidance for further analysis activities using different tools appears to be logical and desired. For instance, Fonnet and Prié (2021) discovered that so far only few investigations exist that are dedicated to similar immersive annotation features, thus requiring further research.

5.5.9.2 Uniform 3D Gestural Interface Design

The first evaluation of the 3D radar chart design confirmed the overall usability of such a data entity visualization approach for IA purposes, confirming that users can successfully make assessments and data interpretations. Based on these insights, a second evaluation of an extended 3D radar chart design was conducted, as described in Section 5.5.5. This evaluation focused primarily on the evaluation of an extended set of features, implemented through a uniform 3D gestural interface design using hand-based grasping and gestural command techniques, without interactive graphical menus as were available during the initial evaluation (see Section 5.5.3). The decision to abandon the utilization of graphical menu-based system control techniques after the initial evaluation was twofold. First, the participants were generally positive and enthusiastic of the more direct interaction and manipulation possibilities using the hand-based grasping and gestural command techniques. Thus, the operation of the extended features in the VE, i.e., *Mode Toggle (Reconfigure/Filter)*, *Data Variable Sort*, *Data Variable Filter*, *Zoom (in/out)*, *Reset*, and *Pause/Resume*, was designed around these interaction techniques. And second, by not providing alternative operation possibilities for the existing features, a by comparison more uniform 3D gestural interface design could be embraced, rather focusing on a more coherent interaction style overall than a mixed one. Naturally, similar to the other empirical evaluations, to a certain extent this was an exploratory investigation to further examine the utilization of 3D gestural input within the IA context.

The participants interacted with illustrative spatio-temporal data of the PWt dataset (see Section 5.5) to complete a series of tasks in a walkthrough-like manner, aiming to represent a typical data analysis activity in the VE. Generally, all of them were able to do so organically and intuitively using the implemented 3D gestural interface, having a smooth and responsive experience in the developed immersive VE. In contrast to the results of the evaluated 3D gestural interfaces presented by Streppel et al. (2018), the majority of the participants in this second 3D radar chart evaluation managed to learn the various features of the 3D UI comparatively quickly, both conceptually and operationally, completing the different tasks accordingly. Huang et al. (2017) reported similar subjective impressions towards learnability and intuitiveness based on the evaluation of their prototype. When asked to do a certain action within the task series, the 3D radar chart users were able to quickly associate the correct interaction in the VE, i.e., the visual object they had to manipulate or the hand posture/gesture they had to perform. The median score of the measured usability (SUS) was above the *good* threshold. Given the presented focus on hand-based grasping and gestural command techniques, the overall results are satisfactory considering the participants were asked to conduct a multitude of predefined tasks rather than just freely exploring the data at their own leisure. The overall user engagement scores (UES-SF; between 3 and 5, median slightly below 4) are also encouraging, indicating positive engagement with the prototype by the participants. This aligns with researcher's observations, as the participants would often use features such as rotation, data variable sort, and data variable filter, even when not explicitly asked for, seemingly naturally engaging with the artifacts in the VE. Closer examination of the individual engagement factor scores (see Figure 5.35) reveals indications that the participants paid close attention during the task completion, assessing the developed VE as aesthetically appealing, and their experience as rewarding (all three with medians around 4) – all in anticipation with the general design objective for this 3D UI. In comparison to the individual engagement score results received in the initial evaluation (see Figure 5.34), all three of these score slightly better in the second evaluation, arguably reflecting on some of the applied improvements.[30] Somewhat in contrast is the result of the perceived usability (median around 3, tendencies towards 4) that also compares worse to the much better SUS score for usability in general. While the participants were overall excited about the 3D gestural interface and able to intuitively interact with data in the presented context, this may be attributed to some identified usability aspects that can be improved upon.

Hand-based Grasping A major aspect on the 3D gestural interface design was concerned with the utilization of hand-based grasping for the interaction with visible virtual objects in the VE, which was appreciated by the participants. They were able to interact with the data variable axis spheres of the reconfigure and filter handle as an indirect widget to adjust the configuration of the 3D radar chart, similar to the node movement interaction as demonstrated in the prototypes by Osawa et al. (2000) and Huang et al. (2017). The participants could intuitively grab and drag the time slice in order to make respective time event selections, reconfirming the results of the first evaluation that utilized the same approach. While this interaction was valued, some shortcomings were identified when the participants had to place the time slice at a specific time event. The tracking and implementation felt "*too sensitive*" as the time slice would sometimes "snap" into one of the adjacent time events when attempting to release the grab, occasionally resulting in slight frustration and requiring some additional interaction to recover from this error – a cost that should not be ignored at a larger scale (Büschel et al., 2018). The time slice movement is dependent on the detected grab-position of the hand, i.e., the position where fingers and thumb meet. In the process of releasing the grab, this position is likely to be updated slightly before the grab is detected as discontinued, thus no longer updating the time event selection. Based on the

[30] It is noteworthy that the first and second 3D radar chart evaluations featured different tasks and evaluation objectives, and are as such not strictly comparable.

current implementation, this issue is proportionally dependent to the length of the 3D radar chart and the amount of included time events, i.e., the resolution of time events. As reference, a 3D radar chart was scaled to corresponded to a total length of 100 cm in the VE, with a total of 150 time events encoded, resulting in an effective gap distance between two time events of 0.67 cm. A lower amount of included time events over the same length would result in a larger gap between individual time events (for instance in the case when the user zoomed in), which would prevent the time slice from snapping to an adjacent time event accordingly. Vice versa, including even more events in the time series would further increase the perceived sensitivity. While one can expect 3D gestural input technologies to become more precise, it is also possible to envision some solutions that are based on the overall 3D UI design and implementation. For instance, rather than exclusively relying on the finger and thumb positions for the grab detection, one could implement an additional dependency based on the hand's back or palm position. In the presented case of grabbing and vertically dragging the time slice, the hand's back and palm positions are likely to remain relatively static in space during the release of the hand-based grasping compared to finger and thumb movements. A threshold could be implemented to prevent time slice movement in such instances, enabling the system to "interpret" the user's intention to discontinue their interaction. Alternatively, another approach of solving this challenge could be based on an asymmetric bimanual interaction, similar as presented in the prototype by Betella et al. (2014). While grasping the time slice with one hand, a gestural command made with the other could "lock" the current time slice position in place, allowing to safely disengage from the interaction without unintentionally moving forward or backward in time. Furthermore, it is noteworthy that this overall matter was not prominent during the initial 3D radar chart evaluation (see Section 5.5.9.1). The reason for this is arguably twofold. First, even though a 3D radar chart was also scaled to correspond to a total length of 100 cm in the VE, it only featured a total of 50 time events, as opposed to 150 time events in the second evaluation. Consequently, the distance between time events corresponded to 2 cm instead of the much closer gap of 0.67 cm, likely preventing the subsequent adjacent "snap" phenomena when discontinuing the time slice interaction via hand-based grasping. And second, the initial version also featured the alternative graphical hand menu that enabled a more step-wise precision control of the time slice, that one may also have to take into consideration. While focusing on an overall uniform interaction design, one could consider the implementation of interaction mode alternatives that can be seamlessly switched in-situ, for instance as described by Wagner et al. (2021), allowing the user themself to chose the preferred interaction technique for their current task.

Gestural Commands In addition to the interaction with visible virtual objects, various invisible gestural commands were available for feature interaction in the VE. Gestural commands such as for travel and time range selection were positively received. The participants appreciated the responsiveness of the time range selection, allowing them to directly highlight the time ranges they were interested in. The continuous semitransparent neutrally colored visualization of the time-series data outside of these ranges provided them with further preview of the data, which was particularly important for them when making the cutoff and deciding whether or not to include additional time events in their selection. The two-handed gestural commands worked generally well. However, based on the researcher's observations and the received feedback from the participants, some improvements can be made in regard to the zoom (in/out) feature. In the initial hand posture of holding both hands vertically slightly apart with their palms facing each other, the tracking sensor sometimes did not recognize the lower hand as it was occluded through the one above. Thus, even though the participants were holding their hands in the correct configuration, they needed to move them around slightly before the sensor detected, interpreted, and translated them appropriately in the VE. Similar feedback was stated by the participants in the evaluation as reported by Huang et al. (2017), expressing a desire for a more robust gesture recognition in such instances. Moving both hands together and then apart, or vice versa, for zoom operations was also reportedly preferred as a gestural command by the participants in the study of Austin et al. (2020), and are thus in alignment with the implemented 3D radar chart interaction design. Thus, the presented findings as part of the second 3D radar chart evaluation, in alignment with the findings described by Huang et al. (2017), highlight the importance for a reliable implementation of bimanual interactions in future applications to satisfy anticipated user preferences.

Unintentional Commands No unintentional reset interactions were observed, even though the participants were able to perform the command swiftly. Similar to the considerations by Fittkau et al. (2015), the gestural command for the reset feature was intentionally designed to prevent unintentional performance, as resetting a 3D radar chart's configuration is a comparatively drastic operation. However, cases of unintentional gestural commands (Pavlovic et al., 1997) occurred most noticeably when participants wanted to display details-on-demand by touching the mode toggle widget, but instead triggered a travel interaction, as their hand posture was detected as pointing forward. While participants were able to travel back and recover from such an error comparatively quickly, it also caused them a mixture of slight surprise, frustration, and uncertainty towards the mode toggle interaction. This is an excellent example for such an unintentional command, demonstrating that different users may attempt the same interaction differently in regard to their hand posture. One can envision that such an issue can be fixed based on the current implementation in various ways, for instance through the implementation of a distance threshold between the virtual hand model and the mode toggle widget, i.e., preventing travel if a user's hand is detected in close proximity to the widget. Thus, the 3D UI could infer in-situ that the user intends to engage with a 3D radar chart rather than attempting to travel. Of particular relevance in regard to this matter is the discussion by Nehaniv et al. (2005) about the importance of a computer system's ability to infer the user's intent with their interactions (see Section 2.2.4 as part of Chapter 2). Similarly, Lee et al. (2021) highlight and reflect on the potential of a visualization tool's ability to be *context aware*, especially under consideration of the capabilities of modern sensor and input interpretation technologies. An immersive data analysis system could be extended through functionalities that enable the system itself to have a better understanding of the user's in-situ context, and in turn their intent, allowing for more robust user experiences.

Chapter 6

Data Analysis Using Hybrid Asymmetric Collaboration

Contents

Collaboration is major aspect of data exploration and analysis, allowing multiple users to combine their expertise and make meaning of data together. Naturally, the developed immersive data analysis environments presented throughout Chapter 5 hold potential to be extended through respective collaborative features, allowing joint data exploration by more than one user. Rather than enabling multiple users to be immersed in the same Virtual Environment (VE), this thesis aims to investigate possibilities for synchronous cross-platform data analysis, applying a combination of immersive and non-immersive interfaces to bridge Immersive Analytics (IA) with Information Visualization (InfoVis) and Visual

Analytics (VA). To explore the application of heterogeneous display and interaction technologies within the overall context of Collaborative Immersive Analytics (CIA), two collaborative systems have been developed under conceptual and technological guidance of the presented system architecture (see Chapter 4). The first collaborative setup was aligned with the second VE iteration (*Stacked Cuboids*) and centered around an explorative analysis task with physically co-located users, focusing on the joint analysis of the spatial data context. Based on the insights and experiences obtained from this first exploratory investigation, a second collaborative setup was evaluated, aligned with the third VE iteration (*3D Radar Charts*) and centered around a confirmative analysis task with remote users, who explored data in the spatio-temporal context.

This chapter begins by introducing the concept of *Hybrid Asymmetric Collaboration* in Section 6.1, aimed to clearly distinguish between technological (hybrid) and user role (asymmetric) aspects within the context of collaborative data analysis. To aid the empirical evaluation of such a collaborative setup, and within the overall context of spatio-temporal data analysis, the *Spatio-Temporal Collaboration Questionnaire* is constructed and described in Section 6.2.

The collaborative setup aligned with the second VE iteration is described in Section 6.3. Its overall scenario focuses on the exploration of multilingualism in the Nordic region. In addition to the immersive VE that encouraged spatial data analysis, a non-immersive collaborator interface was implemented that focused on the exploration of hashtags in *tweets* with respect to language variability. A set of collaborative features was developed to allow the transfer of respective signals from one interface to the other, providing the ability to make spatial references in the data. Pairs of linguistics students explored the data to extract respective insights from a sociolinguistic perspective, providing valuable feedback and impulses for improvement of the system's various collaborative aspects.

Finally, the collaborative setup aligned with the third VE iteration is described in Section 6.4. Its scenario focused not just on collaboration within the spatial data context, but also the temporal. To explore the design of visual references as collaborative information cues, three design options were conceptualized that informed the subsequent development of several spatio-temporal reference variants. Based on an empirical evaluation, these reference designs were subjectively assessed by various participants with respect to aesthetics, legibility, and general preference. Furthermore, the overall hybrid asymmetric collaboration concept was also evaluated in a dedicated empirical evaluation through participant pairs, who completed a confirmative analysis task by identifying various correlations in a representative spatio-temporal dataset. The positive results of this, within the scope of this thesis, final evaluation are promising for the future application of hybrid asymmetric collaboration, allowing for respective design reflections and considerations.

6.1 Hybrid Asymmetric Collaboration

Nowadays, with large multivariate data collected in various contexts, data analysis is seldom conducted in isolation but in collaboration with multiple analysts, either with the same background or as interdisciplinary efforts where each analyst contributes with their specific domain knowledge. Furthermore, due to the increasing complexity of these multivariate datasets, it is often insufficient to use just one analysis tool or application. Instead, multiple ones are required, each for their own data analysis purpose, to examine dedicated aspects of the data (Vuckovic et al., 2022). These form comprehensive analysis workflows where insights gathered using one tool may serve as input in another, leading to further data discoveries.

Naturally, the interplay and collaboration with other users is also of particular relevance within the context of IA, utilizing immersive display and interaction technologies for data analysis purposes. Important foundational aspects in regard to Collaborative Virtual Environments (CVEs) and CIA have been presented in Sections 2.3 and 2.4.1 as part of Chapter 2. These provide exciting opportunities and impulses for further research in regard to collaborative data analysis experiences that incorporate immersive technologies. For instance, a recently published literature survey of IA research, covering the years from 1991 to 2018, revealed that out of the identified 127 system papers, i.e., papers that describe and potentially evaluate an IA system, only 15 focused on collaboration (Fonnet and Prié, 2021). The authors put this lack of research further into perspective, arguing that collaboration is widely considered one of the major aspects for the future success of IA (Fonnet and Prié, 2021). Their argument is in line with the reports and statements of other IA research (Fröhler et al., 2022; Skarbez et al., 2019; Wang et al., 2019; Billinghurst et al., 2018). As presented in Section 2.4, *collaborative analytics* is considered a major topic of the current IA research agenda, among others including challenges with respect to *supporting behavior with collaborators*, *supporting cross-platform collaboration*, and *integrating current collaboration practices* (Ens et al., 2021). All these further highlight the relevance of integrating immersive data analysis tools with non-immersive ones to bridge collaboration across various interface types. After all, IA aims to provide novel, intuitive, and purposeful three dimensional (3D) data analysis tools that complement and synergize with InfoVis and VA workflows rather than replacing them (Wang et al., 2019; Cavallo et al., 2019; Isenberg, 2014).

To further specify and guide the empirical research efforts within the scope of this thesis, particularly with respect to the defined third research objective (see Section 1.2), the following stance in regard to CVEs and CIA is adopted:

- There exists a synergy between immersive and non-immersive analytics interfaces.

- Different visualization and interaction approaches can satisfy different data analysis needs.
- Collaboration between multiple users is anticipated and encouraged, facilitating their joint analytical reasoning and data understanding, independent of their role and background.

With a focus on utilizing heterogeneous display and interaction technologies for cross-platform collaboration within a data analysis context, Sections 2.3 and 2.4.1 have presented several existing concepts and definitions, for instance Hybrid Virtual Environments (Wang et al., 2019), Collaborative Hybrid Analytics (Cavallo et al., 2019), Hybrid Collaboration (Neumayr et al., 2018), and Cross-Virtuality Analytics (Fröhler et al., 2022). However, arguably all of these put an emphasis on the involved technological aspects as part of their definition, i.e., the use of different types of interfaces for data analysis purposes. It is furthermore noteworthy that these mixtures of interface types are not exclusive to the combination of immersive and non-immersive interfaces, but can also evolve around different types of non-immersive tools, respectively different immersive ones, for instance combining Virtual Reality (VR) and Augmented Reality approaches. The role of the user, specifically the analyst within a data analysis context, is often rather implicitly addressed, as sometimes different interface types are utilized for the overall same purpose, while serving their own distinct purposes at other times. Aligned with the previously presented IA research agenda (Ens et al., 2021), the overall third research objective within the scope of this thesis (see Section 1.2) is to investigate collaborative aspects in a scenario where two analysts explore a multivariate dataset at the same time from different perspectives, immersed and non-immersed, each assuming a distinct role in order to contribute to the joint data analysis activity. While various related works assume the non-immersed user commonly in a more "guiding" or "assisting" role (Ens et al., 2021; Welsford-Ackroyd et al., 2020; Thomsen et al., 2019; Peter et al., 2018), this chapter aims to explore scenarios where the involved analysts contribute more equally, each based on their interface and perspective. Consequently, under consideration of the relevant related literature, the concept of *Hybrid Asymmetric Collaboration* is defined and adopted as follows within the scope of this thesis:

> "*Hybrid Asymmetric Collaboration is the use of immersive 3D and non-immersive 2D display and interaction technologies in a collaborative data analysis activity with two or more analysts where each individual analyst assumes a distinct role, based on their knowledge and facilitated by their respective technological interface, with the objective to equally contribute to the joint data interpretation and analytical reasoning.*"
>
> – Reski

The overall objective with the concept of hybrid asymmetric collaboration is to satisfy a desired analytical workflow that incorporates different display types and interaction modalities (Wang et al., 2019; Cavallo et al., 2019; Isenberg, 2014), where collaborators are potentially coming from different domains, providing each their own perspectives and data insights, anticipating a rather equal interaction and contribution to the joint data analysis activity, instead of a *remote expert* scenario (Ens et al., 2019). A fundamental aspect within this context is concerned with providing features that support and facilitate the interplay between the collaborators. While both the immersive and the non-immersive interface have to serve their own purpose and modality, it is important to consider anticipated means of communication and coordination between the collaborators in order to provide meaningful interface extensions that assist them with these endeavors. The design of collaborative information cues is particularly important within the context of immersive technologies, as these are often user-centric in nature, i.e., they are by default rather tailored to be experienced by a single user (Skarbez et al., 2019). Thus, they introduce more remote-like characteristics in regard to potential collaboration, even in physically co-located scenarios, and important visual information cues (gestures, mimic) are not as easily accessible, if at all. Consequently, nonverbal communication features become particularly important in such setups (Cruz et al., 2015).

Some overall considerations for the practical implementation and connection of immersive 3D and non-immersive two dimensional (2D) data analysis interfaces have been presented in Section 4.2.4. An integral aspect as part of the proposed system architecture is the *Real-Time Networking Interface* between the *Immersive VE* and the *Collaborator Interface*, such as a *Non-Immersive Desktop Terminal*. It allows for the transfer of state updates from one interface to the other, and vice versa, and is thus responsible for providing various synchronous collaborative features. Such state updates can include features that allow each collaborator to send and retrieve signals in their interface, aiming to facilitate their overall collaboration through the ability to share their in-situ data analysis context and to make visual references in the data. Furthermore, it is envisioned that the collaborators are able to communicate verbally, i.e., talk to each other, either locally in close physical proximity or remotely via an established audio link.

6.2 Spatio-Temporal Collaboration Questionnaire

Under consideration of the hybrid asymmetric collaboration concept as described in Section 6.1, parts of this thesis are concerned with the empirical evaluation of the analysts' collaboration. Keeping in mind the mixture of immersive and non-immersive interfaces as well as the focus on the analysis of multivariate data, in particular spatio-temporal data (see Section 5.1), there is a need to assess the collaborators' ability to work together and to solve analytical tasks supported by the collaborative features that are integrated in their respective interfaces. Naturally, evaluation methods such as observations and interviews (see Section 2.5.1) can assist with a qualitative collaboration assessment. However, for the empirical evaluation of collaborative aspects in the presented context, it would arguably also be beneficial to assess the collaborators' own perception of their collaboration after they completed a joint data analysis activity. For that purpose, means of self-reporting through the participants are required, commonly implemented through Likert scale statements as quantitative data collection methods. Within the scope this thesis, and thus the related literature, no dedicated standardized questionnaire for the purpose of investigating collaboration in VEs could be identified. Some potential alternatives exist, for instance the *Social Presence Module* as part of the Game Experience Questionnaire (IJsselsteijn et al., 2013; Poels et al., 2007), but are arguably not specific enough within the presented context where two collaborators analyze and interact with spatio-temporal data – in itself a comparatively common data analysis use case (Fonnet and Prié, 2021). Thus, efforts were conducted with the objective to design a self-constructed questionnaire that satisfies these needs.

Based on relevant literature, among others in regard to CVEs and CIA as presented in Sections 2.3 and 2.4.1, important aspects and dimensions of collaboration could be identified. For instance, Dix (1994) presents a general framework for Computer-Supported Cooperative Work (CSCW), as illustrated in Figure 2.11, dissecting its components in *cooperative work* and various aspects of *computer support*, i.e., *communication*, *computerized artifacts of work*, and *non-computerized artifacts*. The importance of communicative aspects as part of the cooperative work is emphasized throughout the framework, in particular as *computer mediated communication*, arguing for its appropriate integration respectively (Dix, 1994). Five key features that CVEs should strive to support within the context of CSCW are defined by Churchill and Snowdon (1998) as well as Snowdon et al. (2001), and have been described in Section 2.3, namely *transitions between shared and individual activities*, *flexible and multiple viewpoints*, *sharing context*, *awareness of others*, and *negotiation and communication*. Aligned with these five key features, Sarmiento et al. (2014) explored an approach to numerically describe the degree of collaboration in a CVE. For that purpose, Sarmiento et al. (2014) propose a set of ten collaboration aspects across two groups, i.e., five related to *interaction* (*predictability*, *peripheral awareness*, *implicit communication*, *double level language*, *overview*) and five related to *immersion* (*management of coupling*, *simplification of communication*, *coordination of action*, *anticipation*, *assistance*). The conceptual framework and taxonomy by Gutwin and Greenberg (2002) are dedicated to *awareness* within the context of group work. Awareness, seen as a state of being attentive and informed about the events in a situation and environment, can be maintained rather easily and naturally in face-to-face workspaces as opposed to groupware ones that do not feature face-to-face communication (Gutwin and

Greenberg, 2002). Gutwin and Greenberg (2002) differentiate between *situation awareness*, *workspace awareness*, and *awareness maintenance*, and in turn propose a Workspace Awareness Framework to describe aspects related to *environment*, *knowledge*, *exploration*, and *action*. Pinelle et al. (2003) propose a task model to support *collaboration usability analysis*. They categorize the mechanics of collaboration into different aspects of communication and coordination, and go on to describe their task model that consists of *scenario*, *task* (*individual* and *collaborative*), and *action* components (Pinelle et al., 2003). Andriessen (2003) proposes a heuristic classification of the major activities involved in cooperative scenarios according to *interpersonal exchange* processes (*communication*), *task-oriented* processes (*cooperation*, *coordination*, *information sharing and learning*), and *group-oriented* processes (*social interaction*). Within the more specific context of Collaborative Visual Analytics, Heer and Agrawala (2008) discuss important design considerations to facilitate collaborative data exploration, among others relevant to *common ground and awareness*, *reference and deixis*, and *incentives and engagement*.

Based on the insights and impressions gained from the various classifications according to the described literature, all discussing collaboration in regard to similar themes from slightly different perspectives, the subjective decision was made to follow and adopt the descriptions by Churchill and Snowdon (1998) as well as Snowdon et al. (2001) respectively, emphasizing various key aspects that CVEs should aim to support. Arguably, the investigation of transitions between shared and individual activities, negotiation and communication, sharing context, and awareness of others (Churchill and Snowdon, 1998) should allow for the retrieval of insights in regard to different important collaborative aspects, thus providing a "bigger picture" of the collaboration during the completion of a data analysis task.[1] Consequently, to assess these aspects in a setting of synchronous collaboration and within the context of spatio-temporal data analysis, a self-constructed questionnaire, titled *Spatio-Temporal Collaboration Questionnaire* (STCQ),[2] was designed as follows. It features a total of 17 5-point Likert scale statements that are thematically relevant to the four dimensions as adopted from Churchill and Snowdon (1998), and described in the following way:

- *Transitions between Shared and Individual Activities* (TSIA): The interplay between individual and group efforts, including the ability to switch between these, within the scope of collaborative work.
- *Negotiation and Communication* (NC): Verbal conversation (i.e., talk) facilitated through the ability of utilizing nonverbal information cues in order to discuss and interpret any task-related aspects of the activity (e.g., findings in the data, roles and structure of task approach, and so on).

[1] It is noteworthy that the described fifth key feature for CVEs according to Churchill and Snowdon (1998), namely *flexible and multiple viewpoints*, was excluded at this point in time, as it was deemed negligible within the scope of this thesis.

[2] The STCQ was designed in collaboration with Aris Alissandrakis.

- *Sharing Context* (SC): Characteristics and features of the shared space that facilitate and support focused and unfocused collaborative work, leading to shared understandings.
- *Awareness of Others* (AO): The ability to understand your partner's activity during times of (1) focused collaboration and active communication (i.e., *group* efforts), as well as (2) more independent and *individual* work.

Table 6.1 presents an overview of all item statements and their Likert scales across these four dimensions. The design of the individual item statements is held purposefully generic, anticipating re-usability, remix, and further adoption for evaluations in similar contexts in the future.[3] While the STCQ is anticipated to aid the self-assessment of the collaborators' joint spatio-temporal data analysis within the context of a hybrid asymmetric collaboration setup, it is generally considered to be independent of the involved interface technologies. Thus, the questionnaire should also be easily applicable in other collaborative data analysis scenarios that focus on spatio-temporal data. In fact, only the items *AO.2*, *AO.3*, *AO.5*, and *AO.6* as part of the AO dimension (see Table 6.1) are rather context specific in regard to the collaborator's ability to send and retrieve spatio-temporal references using their respective interfaces, inquiring ratings about the collaborator's *location in space* and *time reference* during *group* and *individual efforts*.

In practice, the questionnaire is to be filled out by each collaborator individually, in isolation, and directly after the respective task completion. The evaluation of the answers should allow for a quantitative analysis of a system's collaborative features, and provide insights in regard to the collaboration as perceived by the collaborators themselves. Furthermore, the results should be interpreted within the context of the tested system and against its anticipated design, for instance to assess if an anticipated role distribution between the collaborators was fulfilled as intended, to name just one example.

6.3 Collaboration in VE Iteration 2: Stacked Cuboids

The first empirical efforts in regard to the presented concept of hybrid asymmetric collaboration were conducted within the scope of the second VE iteration, i.e., utilizing the *Stacked Cuboid* data entity visualization design and the Nordic Tweet Stream (NTS) dataset as described in detail throughout Section 5.4. Utilizing the developed VE, the immersed user is able to explore spatial and temporal aspects of the NTS corpus with a focus on multilingualism. However, besides language, location, and time related data variables, each *tweet* contains a variety of additional metadata variables, for instance all the hashtags[4] used in a tweet. Due

[3] A link to an online repository with the STCQ is included in Appendix D.

[4] A hashtag on Twitter is utilized to index keywords and topics, enabling users on the social networking platform to easily follow these.

Item	**Statement**	**Scale**
TSIA.1	How many of your efforts during this task would you consider to have been *individual* efforts?	L1
TSIA.2	How many of your efforts during this task would you consider to have been *group* efforts?	L1
TSIA.3	According to your impression, who was more in a leading/directing role during the *group* efforts?	L2
NC.1	According to your impression, how often did you communicate *verbally* to your partner?	L3
NC.2	According to your impression, how often did you communicate *nonverbally* to your partner?	L3
NC.3	How often would you consider did *dialog* take place?	L3
NC.4	How often would you consider did *negotiation* take place?	L3
NC.5	Who would you say mostly initiated the *negotiations*?	L2
SC.1	The collaborative features of the system allowed me to focus on the same subject as my partner.	L4
SC.2	The collaborative features of the system allowed me to establish a dialog with my partner.	L4
SC.3	The collaborative features of the system distracted me from my *individual* efforts.	L4
AO.1	During your *group* efforts, how much were you aware of your partner's activities?	L5
AO.2	During your *group* efforts, how much were you aware of your partner's location in space?	L5
AO.3	During your *group* efforts, how much were you aware of your partner's time reference (time point/interval)?	L5
AO.4	During your *individual* efforts, how much were you aware of your partner's activities?	L5
AO.5	During your *individual* efforts, how much were you aware of your partner's location in space?	L5
AO.6	During your *individual* efforts, how much were you aware of your partner's time reference (time point/interval)?	L5

Scale	**Format (5-point Likert)**				
L1	none	a few	some	a lot	every
L2	mostly other	more other, some me	both equally	more me, some other	mostly me
L3	never	rarely	sometimes	often	constantly
L4	strongly disagree	disagree	neutral	agree	strongly agree
L5	not at all	a bit	some	a lot	always

Table 6.1: Overview of the items and scales of the designed *Spatio-Temporal Collaboration Questionnaire* (STCQ).

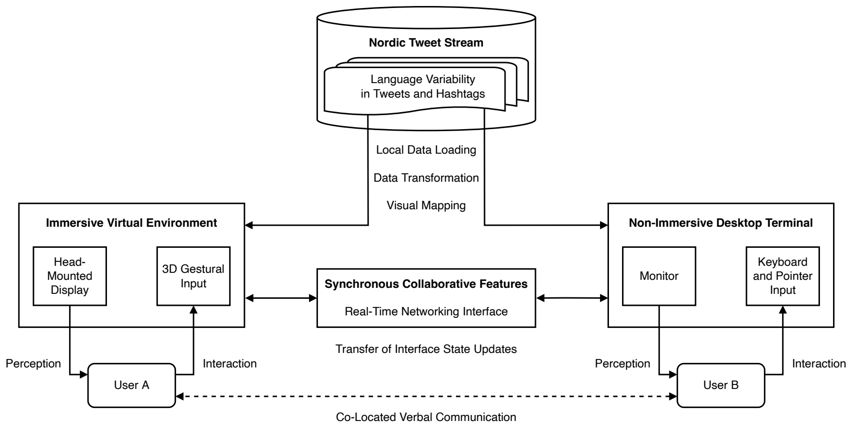


Figure 6.1: Conceptual system overview of the hybrid asymmetric collaboration setup under utilization of the second VE iteration (see Section 5.4). Detailed descriptions about the various system components are provided in Section 4.2.

to the diversity of these metadata variables, various possibilities for collaborative data analysis arise. Even though head-mounted display (HMD) technologies advance in various aspects, among others display resolution and tracking, it is arguably still more convenient to read large amounts of text outside a VE, using a normal computer monitor. Thus, an overall explorative data analysis scenario (see Section 5.2) was adopted that is relevant for linguistics researchers, centered around the collaborative analysis of hashtags in tweets in regard to language variability. While the immersed user analyzes the NTS data with a focus on the contextual spatial aspects of the dataset for a given point in time, identifying potentially interesting locations during their exploration in the VE, the non-immersed user is provided with an interface that enables a closer examination of the used hashtags in the different locations. Consequently, the two users collaborate by utilizing different display and interaction technologies (*hybrid*) and take on different analysis roles (*asymmetric*) respectively, i.e., following the concept of hybrid asymmetric collaboration as introduced in Section 6.1.

As a starting point for the practical investigation of this matter, i.e., hybrid asymmetric collaboration within the context of spatio-temporal data analysis, the decision was made to focus on supporting the collaborators' mutual understanding and their ability to make spatial references using their respective interfaces, but not temporal references. Thus, for the data analysis at this stage, the collaborators were presented with the NTS dataset for a predefined point in time without the ability to manipulate the temporal context using their interfaces.[5]

Figure 6.1 illustrates the overall system architecture, following the proposed *Collaboration Infrastructure* as described in Section 4.2.4. The relevant data from the NTS dataset are served to the respective interfaces, i.e., the immersive VE and the non-immersive desktop terminal. Both interfaces are connected through a real-time networking interface in order to allow support for synchronous collaborative features. Furthermore, it is envisioned that the two collaborators are co-located in physical proximity, such as being in the same room, allowing them to communicate verbally without the need for a respective remote audio link.

6.3.1 Non-Immersive Desktop Terminal

The developed non-immersive desktop terminal, as illustrated in Figure 6.2, is designed as an interactive InfoVis with the objective to support the analysis of hashtags in tweets in regard to language variability. In terms of display and interaction technologies, as indicated by its description, it is intended to be operated using a normal desktop monitor and pointer (mouse) input. The desktop terminal is composed of four views, supporting the overall main tasks of (1) browsing the hashtags as detected in the various spatial locations, and (2) applying subsequent sorting based on frequency and language. The four views are partially interconnected, i.e., interactions in one view will cause others to update accordingly.

View Composition and Interaction The top right part of the interface features a *Map View*, displaying a geographic map of the Nordic region with all spatial data entities represented as circles. Each individual data entity portrays a compiled cluster of real-world places in accordance to the NTS dataset, as described in Section 5.4. Using the pointer, each data entity can be *selected* by clicking on it. More specifically, the user can select two data entities at a time, using the pointer's click and context-click respectively, i.e., left and right click. Data entities selected through a normal click are referred to as *selected* and indicated by a solid green outline, while those selected through a context-click are referred to as *bookmarked* and indicated by a dashed green outline. The *Desktop User Table View* in the top left part of the interface features two tables, each displaying hashtag related information for the respective selected and bookmarked data entities in accordance to the interactions in the map view. The table outlines are equally color-coded to indicate these connections, i.e., solid green and dashed green. Each table is composed of three columns, i.e., from left to right, the *frequency* stating the total count of the hashtag, the *hashtag* itself, and the *language* of the tweet the hashtag was detected in. Each column header can be clicked in order

[5] The support for both spatial and temporal references is explored as part of the subsequent research efforts presented in Section 6.4.

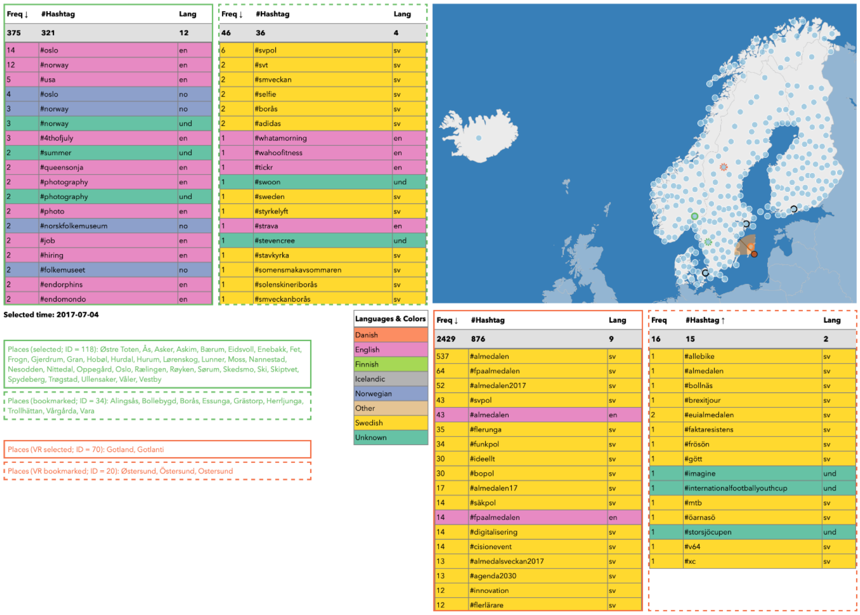


Figure 6.2: An impression of the implemented Non-Immersive Desktop Terminal to explore hashtags in tweets in regard to language variability, and within the context of the hybrid asymmetric collaboration setup with the second VE iteration. **Top Left**: Desktop User Table View. **Top Right**: Map View. **Bottom Left**: Information View. **Bottom Right**: VE User Table View.

to toggle between ascending and descending *sorting* of its respective column. Similarly, as part of the implemented collaborative features (see Section 6.3.2), the bottom right of the interface displays the *VE User Table View*, color-coded with respective orange outlines, and presenting hashtag information based on the immersed user's selected and bookmarked data entities. Finally, the *Information View* in the bottom left displays additional information, such as the temporal context (date or week) of the data analysis, the unique real-world places in the selected and bookmarked data entities, and a legend that illustrates the various languages and their assigned colors.

6.3.2 Collaborative Features Design

To enable a pair of users to collaboratively explore and analyze the NTS dataset, at the same time and using heterogeneous interface types, i.e., one utilizing the immersive VE and the other the non-immersive desktop terminal as described in Sections 5.4 and 6.3.1, additional features to support and facilitate their collaboration are needed. At this stage within the investigation, these collaborative features are designed to address the following important CSCW concepts (see Section 2.3):

- *Common Ground and Awareness*: Facilitate the users' mutual understanding of their in-situ spatial contexts during their joint data analysis activity.
- *Reference and Deixis*: Support for the transmission and reception of spatial references as nonverbal communication cues in each of the collaborators' respective interfaces during their joint data analysis activity.

The collaborative features should support the users' verbal communication through the display of the respective nonverbal information cues, particularly keeping in mind the immersed user in the VE, wearing a HMD and thus unable to establish visual contact with the non-immersed user. Verbal communication and the use of deictic terminology along the lines of "*look at that data entity behind you*" or "*have a look at that one*" under assistance of the provided nonverbal information cues may facilitate their joint data analysis activity. Thus, the collaborative features design aims to satisfy multiple intentions. First, based on additional visual (nonverbal) information cues, the immersed user in the VE should be made aware of what the non-immersed user is exploring, and vice versa. Second, due to the increased context awareness, their verbal communication is eased. And finally, due to the assisting collaborative information cues and the facilitated verbal communication, the foundations for an engaging collaborative data analysis activity are provided, enabling the users to make sense of the data together as well as fostering their individual understanding of the data, while both of them operate tools with different purposes to explore the same data.

Collaborative Information Cues: VE to Desktop Terminal The following collaborative information cues *from the immersive VE* are displayed *in the non-immersive desktop terminal*. The position, orientation, and field of view of the immersed user in the VE are displayed within the map view of the non-immersive desktop terminal (see Figure 6.2 top right). This should allow the non-immersive interface user to have at all times a spatial understanding of where their immersed partner is located, and what direction they are facing. Furthermore, both the selected and the bookmarked data entities of the immersed user are visually indicated in the map view through respective color coding that is in line with the design of the non-immersive desktop terminal, i.e., a solid orange outline for the selected data entity, and a dashed orange one for the bookmarked one. Additionally, hashtag information about these data entities are displayed in the VE user table view and updated in accordance to the immersed user's interactions (see Figure 6.2 bottom right). This allows the non-immersed user to quickly switch ad hoc to the spatial context of their collaborator without the need to make first the respective selections themself (REQ 16 and REQ 17 in Table 4.1).

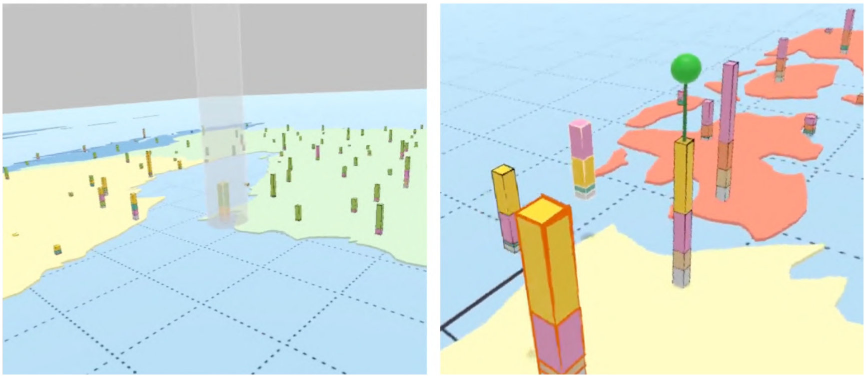


Figure 6.3: Impressions of the additional spatial reference features available in the second VE iteration, and within the context of the hybrid asymmetric collaboration setup. **Left**: Spatial reference *pillar*, set by the user of the non-immersive desktop terminal. **Right**: Bookmark as a *virtual pin*, set by the immersed user in the VE and signaled to the non-immersive desktop terminal.

Collaborative Information Cues: Desktop Terminal to VE The following collaborative information cues *from the non-immersive desktop terminal* are displayed *in the immersive VE*. To enable the non-immersed user to make a spatial reference in order to point the immersed VE user to a specific data entity, a *pillar* design was implemented (see Figure 6.3 left). In particular, a semitransparent white cylinder object is placed at the center of a stacked cuboid in the VE, representing the respective data entity that is bookmarked in the non-immersive desktop terminal. The pillar's height is scaled to be comparatively high, creating a spotlight impression, similar to the bookmark feature initially implemented in the first VE iteration (see Section 5.3.2). This design is intended to be easily identifiable by the immersed user, to catch their attention and guide them to that location, and thus to allow them to align their in-situ data analysis context with the one of their non-immersed partner (REQ 16 and REQ 17 in Table 4.1).

6.3.3 Evaluation: Collaborative Explorative Analysis

An empirical evaluation was set up with the overall objective to explore and validate the practical interplay between immersive and non-immersive interfaces for collaborative data analysis. Naturally, the scenario of the evaluation was centered around multilingualism and the NTS dataset, utilizing the second VE iteration and its stacked cuboid data entity visualization design (see Section 5.4) as well as the non-immersive desktop terminal and the implemented collaborative features (see Sections 6.3.1 and 6.3.2). To evaluate this real-world scenario, pairs of first-year linguistics students were recruited that alternated the roles of one participant being immersed in the VE and the other one using the non-immersive desktop terminal (*within-subject* design).

6.3.3.1 Physical Study Space

All study sessions were conducted at the VRxAR Labs research group lab at Linnæus University, featuring an overall setup as initially described in Section 5.3.3.1. To briefly recap, it features (1) a workstation for the researcher to moderate the study, collect data, and monitor the involved system components, (2) a designated square two-by-two meter area with a visual three-by-three grid on the floor for the immersed user, wearing the HMD, to move freely without obstacles, and (3) an additional desk and several chairs for further use by the participants. The VE user utilized a HTC Vive HMD with a Leap Motion Controller attached in front of it to enable 3D gestural input. Within the scope of the collaborative system evaluation, the additional participant desk was set up as the workstation for the non-immersed collaborator, as shown in Figure 6.4. It featured a desktop computer with a 27-inch monitor, keyboard, and mouse to run the developed non-immersive desktop terminal in full-screen mode. The co-located placement of both collaborators in the same physical location allowed them to communicate verbally (REQ 18 in Table 4.1) without the need of additional technologies. Furthermore, one of the complementary chairs beside the VR system's calibrated safe interaction area was taken over by a second researcher, i.e., a professor of English linguistics in the role of the respective domain expert, assisting with the data collection and the collaboration assessment. Overall, the physical study space provided enough space for the two participants and the two researchers to conduct such an experiment comfortable and uninterrupted.

6.3.3.2 Interfaces Setup

The non-immersive desktop terminal and all the collaborative features across both interfaces were available to the participants as described throughout Sections 6.3.1 and 6.3.2. The VE for the immersed user was generally set up as introduced and described in Sections 5.4.1 and 5.4.2, following the stacked cuboid data entity visualization design. However, a few changes to the overall interface were applied to accommodate the collaborative data analysis scenario of exploring hashtags in tweets in regard to language variability. First, the three *information panels* were modified as follows. Overall additional information and all unique location names are displayed in the *left panel*. A pre-compiled tag cloud, displaying the most prominent hashtags in the data entity, and color-coded in respect to the tweet's language are presented in the *center panel*. A detailed listing about the language distribution of the selected data entity is displayed in the *right panel*. Figure 6.5 provides an impression of the modified information panel setup from the immersed user's field of view. Second, with the current focus on the spatial

Figure 6.4: Two photos for an extended impression of the VRxAR Labs research group lab at Linnæus University. **Left**: A desk and chair where the non-immersed collaborator can be seated to operate the desktop terminal, and the researcher's workstation as a standing desk. A corner of the set up safe interaction area can be seen in the foreground. **Right**: The VR system's calibrated safe interaction area (two-by-two meter), outlined as a three-by-three grid on the floor, for the immersed user to move freely without obstacles.

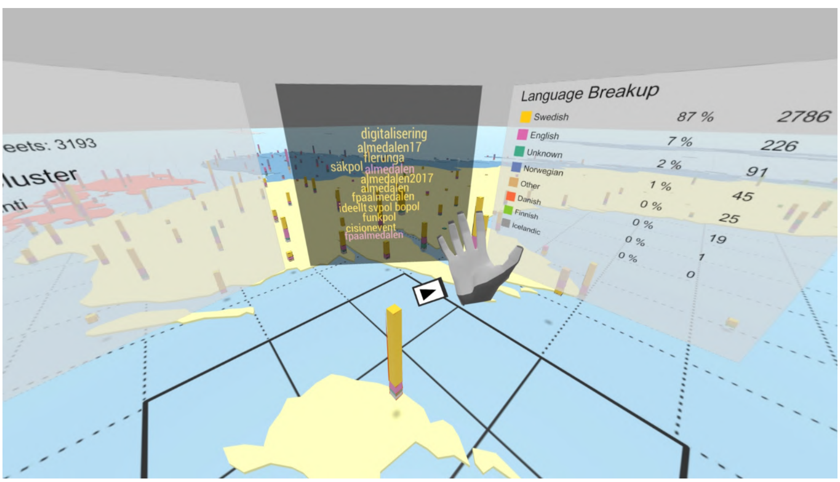


Figure 6.5: An impression of the modified VE interface, presenting the information panel setup in the second VE iteration, and within the context of the hybrid asymmetric collaboration setup. The original interface is described in Section 5.4. **Note**: Additionally, the user's right hand features a juxtaposed adapted 2D graphical menu to temporarily *pause/resume* any kind of interaction – an exploratory feature, further examined in the third VE iteration (see Section 5.5.2).

data analysis and as such the collaborators' ability to make spatial references, the *Time Event Selection* feature in the VE was temporally disabled, i.e., the respective graphical menu attached to the user's hand was hidden. Consequently, the VE user had no means to manipulate the selected time event after application start-up, as intended, to ensure both collaborators remain in the same temporal context during their joint data analysis. Third, similar to the non-immersed user's ability to make a dedicated nonverbal signal to the VE user (see Section 6.3.2), the VE had to be extended for feature parity. To keep track and signal a data entity that the immersed user determines as interesting, a visible marker can be attached by *grabbing* a stacked cuboid and *pulling out a virtual pin* (see Figure 6.3 right). A signal about this "bookmarked" data entity is sent to the non-immersive desktop terminal to update its *VE User Table View* accordingly (see Section 6.3.2).

6.3.3.3 Task

Each pair of participants was presented with an explorative data analysis task (undirected search with no hypotheses given; see Section 5.2), requiring them to collaborate within a sociolinguistic context. In particular, two tasks were defined, requesting the collaborators to make respective assessments for each of the five Nordic regions regarding (1) the language distribution of tweets, i.e., comparing regional data entities, and (2) how closely the distribution of the hashtags' language matches the distribution of the tweets' language (for each region). The collaborators were asked to combine their observations with respect to both of these tasks into a final reasoning for each of the regions. For the purpose of this evaluation, two temporal contexts were selected from the year 2017 within the NTS dataset, i.e., May 23 as dataset A, and October 21 as dataset B.[6] For each study, the order of the datasets was randomized and the participant roles (immersed and non-immersed) would switch accordingly. Consequently, the described tasks were completed twice, once with dataset A, and once with dataset B, with each participant operating each interface type once (*within-subject* design). Overall, the task was designed to encourage an open exploration of the dataset by the participants using their own strategy and pace.

6.3.3.4 Measures

A mixture of task performance and subjective methods were applied (see Section 2.5.1), allowing for quantitative and qualitative data collection as follows. To obtain insights about the usability of the collaborator interfaces, i.e., the immersive VE and the non-immersive desktop terminal, the System Usability Scale (SUS) was administered. A logging system, implemented as part of the collaborative system, enabled the comprehensive collection of all user input across the two interfaces. Furthermore, observations were conducted by two researchers, taking

[6] The birthdays of Carl Linnæus (May 23) and Alfred Nobel (October 21).

notes about the collaborators' verbal communication as well as their interactions with the interfaces and as such with each other. Finally, a brief semi-structured group interview was prepared, conducted after the collaborative tasks on both datasets were completed, inquiring additional insights from both participants with the freedom to pose follow-up questions based on their answers and prior made observations. In particular, important collaboration aspects were defined for the participants, followed by three general interview questions as follows:

> Definitions: *Communication refers to your your verbal communication and the referencing actions in the interfaces. Coordination refers to your actions as a result of your prior communication. Collaboration refers to the outcomes of prior communication and coordination actions.*

- Question 1: *Do you have any comments about your communication/coordination/collaboration during your work on the two tasks?*
- Question 2: *Would you consider the two interfaces (immersed and non-immersed) evenly balanced?*
- Question 3: *In a hypothetical scenario where you are only allowed to use one of the two interfaces (immersed or non-immersed), which one would you prefer?*

6.3.3.5 Study Procedure

In preparation of the empirical evaluation, a class of first-year linguistics students was provided with an introduction to the developed system one week prior to the practical conduction of the evaluation. A video demonstration was used to present all the main functionalities of the immersive and non-immersive interfaces (see Appendix A). They were also briefly introduced to the overall data context, but no details about the tasks were revealed. Overall, this format was chosen to ensure that all participants received the same introduction. With respect to the actual study, each session followed the same procedure of four stages, anticipating an overall session duration of approximately 60 minutes:

1. Introduction (10 min);
2. Collaboration 1 (20 min):
   (a) Warm-up 1 (5 min in the VE),
   (b) Task on dataset A or B (10 min in the VE),
   (c) Questionnaire 1 (5 min);
3. Collaboration 2 (20 min):
   (a) Warm-up (5 min in the VE),
   (b) Task on dataset B or A (10 min in the VE),
   (c) Questionnaire (5 min);
4. Group Interview (10 min).

The *introduction* began with welcoming the participant pair and asking them to fill out an informed user consent. Afterwards, each of the two *collaboration* stages featured the same procedure of first providing a brief *warm-up* phase to allow the participants to familiarize themselves with their respective interfaces, then completing their exploratory data analysis task as described in Section 6.3.3.3, and afterwards completing in seclusion the SUS *questionnaire* for the interface they operated. Once the *first collaboration* stage was completed, the used dataset and participant roles were iterated, after which the *second collaboration* stage was conducted. Two researchers, the thesis author and a professor of English linguistics in the role of the respective domain expert, observed the participants during their *task* completion stages and took notes. Furthermore, it is noteworthy that for each of the respective *warm-up* phases, data of a different temporal context (July 4, 2017) was used to prevent the participants from taking over any data insights from *warm-up* to *task*. Once the two *collaboration* stages were completed, the semi-structured *group interview* was conducted with both participants and both researchers. Finally, the participants were thanked and sent off.

6.3.4 Results

6.3.4.1 Participants

For this empirical evaluation, a total of $n = 15$ participants were recruited, who were teamed up as pairs that allowed the conduction of overall eight collaborative study sessions.[7] The participants were first-year linguistics students, enrolled in the Sociolinguistics module of the English language B.A. program at Linnæus University (see Table 5.9 D). As part of the study preparation, all of them were briefed about the developed collaborative system one week prior to their participation (see Section 6.3.3.5). None of them had any prior hands-on experiences with the developed interfaces.

6.3.4.2 Task Assessment

Based on the explorative analysis task (see Section 6.3.3.3), the collaborator pairs were free to choose their own data exploration and analysis strategy. All of them were able to utilize the provided interfaces to make assessments in regard to the given task. Table 6.2 provides an overview of their assessments and what Nordic regions they chose to analyze during their two tasks. The initially estimated ten minutes limit per task proved insufficient for making assessments for all five regions. Instead, the participants focused on different regions across the two tasks, even though the temporal context during each of the tasks differed.

[7] Due to a last minute cancellation by the sixteenth participant, a doctoral student in linguistics at Linnæus University was recruited as a substitute (their data are excluded from the results analysis), allowing them to pair up with the fifteenth participant.

Session	Assessed Regions	
	First Task	Second Task
1	DK, FI	SE, NO, IS
2	DK, IS, NO	SE, FI
3	SE	DK, IS
4	IS, NO, DK	FI
5	SE, FI	SE, NO, IS, DK
6	SE	DK, NO
7	IS, DK	FI, SE
8	DK, SE	NO, FI

Table 6.2: For the first and second task of each study session, the Nordic regions that the collaborator pairs analyzed. The presented order reflects the order of regions that the pairs chose to analyze during their task completion. **Country Codes**: DK (Denmark); FI (Finland); IS (Iceland); NO (Norway); SE (Sweden).

Exploring Denmark, the students discovered that the use of English and Danish language in tweets (and hashtags) was rather equally distributed, with some increased frequency of Danish towards the countryside. Students were surprised by how dominant and "*omnipresent*" the use of Finnish language was across the entirety of Finland, with English language appearing only in more populated places. Additionally, they noted the low frequency of Swedish in tweets that originated in Finland. These observation were declared as quite different from those of other countries, initiating thoughts such as, "*The Finns really want to use it [the Finnish language] to prevent it from going away*." The students were intrigued with their observations in Iceland, stating that both English and Icelandic language were quite equally used in tweets, but more hashtags were attached to English tweets. Investigating those hashtags, the students concluded that tourism may be a reason. Exploring Norway, the students found that the majority of hashtags attached to tweets were in Norwegian rather than English, even though there seemed to be quite some English traffic along the coasts, arguably attributed to "*tourists visiting the fjords*." Sweden's exploration was more versatile, as students highlighted some differences between the south (fairly uneven distribution of hashtags and languages) and the more center and northern parts (more Finnish language towards the border to Finland; lots of tweets discovered in *other* languages, while hashtags were rather attached to Swedish and English tweets). The students also discovered that the use of Swedish was dominant in the metropolitan areas both in terms of language and hashtag usage. Similar to Norway, some students assessed that a fair amount of the hashtags themselves were in Swedish.

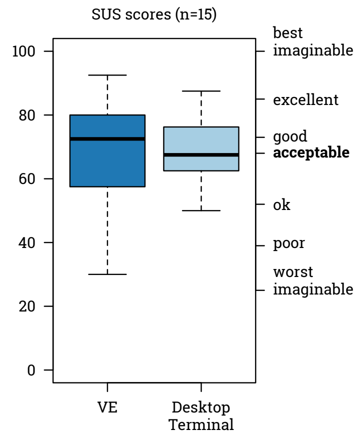


Figure 6.6: Results of the SUS for both interface types, i.e., the immersive VE and the non-immersive desktop terminal, presented including both the original numerical scale as well as the supplemental adjective ratings (see Section 2.5.1). The mean values for both are slightly above the *acceptable* threshold (i.e., score > 68). An independent samples t-test was conducted to compare the two conditions' mean SUS score. There was no statistically significant difference for the VE ($M = 68.5$, $SD = 16.17$) and desktop terminal ($M = 69.67$, $SD = 11.09$) conditions; $t(24.79) = -0.23$, $p > .05$. Samples were tested for normality using the Shapiro-Wilk test.

6.3.4.3 Questionnaires

Figure 6.6 presents the collected self-assessments in regard to system usability for both interfaces of the collaborative system.

6.3.4.4 Logging

Using the collected data from the system logs, it was possible to reconstruct aspects of the collaborative behavior of the pairs. For instance, following an overall similar pathway visualization approach as utilized in the evaluation of the first VE iteration (see Section 5.3.4.4), a pair's travel interactions over time can be visualized, providing visual indications about their spatial data analysis contexts during the task completion. Examining these pathway visualizations, for instance as illustrated in Figure 6.7, it becomes apparent that the collaborator pairs explored only parts of the dataset (due to the set time limitations), and that there were instances where they collaborated closely (each examining the same spatial context) as well as sometimes more individually (each examining different spatial contexts). Furthermore, based on the implemented collaborative features that facilitated the pairs' spatial referencing abilities across the two interfaces

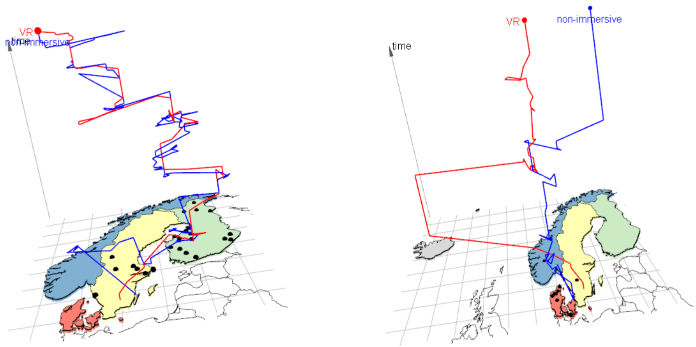


Figure 6.7: Examples of different data exploration strategies based on a participant pairs' travel and selection interactions using their respective interfaces, compiled as pathway visualizations over time. **Left**: The participant pair collaborated closely, examining the same spatial contexts for the most parts. **Right**: The participant pair collaborated more individually, rather examining different spatial contexts. **Note**: An overview of all pathway visualizations is presented in Appendix C.

Pair	**Task**	**Immersed User** *referred (explored)*	**Non-Immersed User** *referred (explored)*
p1	1	8 (8)	10 (10)
p1	2	4 (4)	15 (14)
p2	1	0	12 (10)
p2	2	1 (1)	5 (3)
p3	1	1 (0)	3 (3)
p3	2	1 (1)	1 (1)
p4	1	2 (2)	*substitute participant*
p4	2	*substitute participant*	4 (2)
p5	1	0	1 (1)
p5	2	2 (2)	3 (1)
p6	1	1 (1)	7 (5)
p6	2	4 (3)	4 (3)
p7	1	5 (5)	13 (5)
p7	2	3 (3)	20 (13)
p8	1	6 (4)	5 (5)
p8	2	7 (6)	5 (5)

Table 6.3: An overview of the dedicated spatial referencing interactions for the immersed and the non-immersed users per pair, presenting the number of *referred* data entities by each user and how many of these were subsequently *explored* by their partner through respective travel and selection interactions.

(see Section 6.3.2), the results presented in Table 6.3 indicate that the immersed and non-immersed users referred to specific locations (data entities) that were then explored by their partner. In some cases, these features were used more than in others, but these deictic gestures were to a great extent acknowledged, and benefited their collaboration, enabling them to successfully synchronize their in-situ spatial data analysis contexts.

6.3.4.5 Observations

The notes from the observations conducted by the two researchers were combined, coded, and examined in order to identify reoccurring patterns. Within the context of this evaluation and the results presentation, references are made to *sessions* and *tasks*. Session indicates that the phenomena were observed in both tasks, while task indicates an occurrence in only one of the two tasks of an individual collaborator pair.

Three pairs approached their task solving process noticeable systematical and structured throughout their session, engaging in frequent verbal communication and discussion to make informed assessments. Three pairs appeared to analyze the data closely together, unconnected to whether they approached the task completion systematically or not. One pair was observed exploring the datasets rather individually and in silence during one task. In terms of the decision making with respect to what spatial context to explore next, a rather equal distribution between guidance through the immersed VE user and guidance through the non-immersed desktop terminal user was noticed during three tasks. The immersed user guiding the non-immersed one was observed in one task, while the non-immersed user clearly guided the VE user in four tasks. The use of deictic references and related terminology was explicitly observed throughout four sessions and two tasks, and included phrases, such as:

- *"I am here!"*
- *"Let's look there!"*
- *"Do you want to to go there?"*
- *"Go here!"*
- *"Let's finish here first!"*
- *"Where do you want me to go now?"*
- *"Do you see this/that?"*
- *"Have you seen this one?"*
- *"Where do you want to go? Select a place, and I will go there!"*
- *"Come here!"*
- *"Turn around!"*

- *"Behind you!"*
- *"Next to you!"*
- *"Should I come to you?"*
- *"Look at this one, I pinned another one!"*
- *"I pinned this one here."*
- *"This is quite interesting! Look at this one!"*

During one session and one task, one member of the pair appeared more dominating (arguably by personality) with respect to the collaborative team work compared to their partner. Within one session and three tasks, the pairs made strong use of their contextual and prior knowledge, discussing and commenting on certain phenomena they discovered within the data, covering a variety of topics, for instance anecdotes about places the participants had visited in real life or about major news events that significantly influenced the Twitter hashtags on the task dates. The use of the collaborative features, particular in terms of the collaborators' mutual ability to create spatial references by "pointing" in the data, was frequently observed during five sessions and three tasks (see also Table 6.3). Participants in three sessions gave active verbal acknowledgments that the display of the VE user's position and orientation in the map view of the non-immersive desktop terminal supported their understanding and awareness of their immersed partner's current spatial context. During one session and three tasks, the VE user was observed actively commenting on the underlying map on the floor in the VE and, in conjunction, their ability to navigate and orient themself based on their geographical knowledge. Visible and audible indications of a generally pleasant experience, for instance through noticeable fun and laughter, were observed during four sessions and two tasks. Six participants appeared to be immediately very fluent with the 3D gestural input as interaction modality in the immersive VE. Some additional initial efforts, for instance to *toggle the information panels* by performing a "thumbs up" hand posture (see Section 5.4.2), were observed with six participants.

6.3.4.6 Group Interview

In a joint group interview with both participants of each collaborator pair, they were asked whether they considered their roles based on the two interfaces evenly balanced. Additionally, they were asked to consider a hypothetical scenario where they would only be allowed to use one of the interfaces to analyze the data, and whether they would have a preference regarding the use of the immersive or the non-immersive interface, given that each interface satisfies different data analysis purposes. Four participants expressed the opinion that the balance between the user roles across the immersive and non-immersive interfaces were rather equal. Six participants favored the immersive VE over the non-immersive desktop terminal, while one participant preferred the non-immersive desktop terminal. Four participants stated that their answer regarding preference would depend on the task or the objective of the activity, thus not being able to decide whether they would favor the immersive or the non-immersive interface. However, five participants actively argued that they would prefer a collaborative scenario as experienced within the study, regardless of their interface preference. Six participants could not give a clear preference of one interface over the other. The participants also provided some further general comments. In particular, six participants emphasized that the immersive VE allowed them to get a good overview of the data. Six participants stated that the non-immersive desktop terminal featured a lot of details. Eight participants actively acknowledged experiencing a learning effect in their collaboration and interaction between the first and second task. One pair genuinely appreciated their experience, stating that it felt like "*a two-person job*." Another pair acknowledged the feeling of time passing by rather quickly due to their engagement in the joint data analysis activity. Another pair highlighted that, "*It was so much fun to have both applications, especially for the non-immersed user, otherwise it would be rather dull.*" Two pairs emphasized that they did not feel any barrier within their verbal communication, enabling them to naturally speak and interact with each other using the provided features, self-reporting their perceived coordination and collaboration as "*good and easy*".

6.3.5 Discussion

The integration and bridging of immersive and non-immersive interfaces for collaborative data analysis, interpretation, and meaning making is a subject of high interest in the research community, among others highlighted by Fröhler et al. (2022), Ens et al. (2021), and Wang et al. (2019). In addition to the related literature, some reoccurring observations were made based on prior experiences of demonstrating IA interfaces, for instance within the scope of the developed second VE iteration and the linguistics case study as presented in Section 5.4. At public demonstrations in particular, a typical setup would involve a single user to immerse themself in the VE, mirroring their field of view on a large display for bystanders to follow. However, those had no straightforward means of engaging, communicating, and collaborating with the immersed user. To their own frustration, any such attempt turned out rather difficult to achieve, as the immersed user was not able to successfully identify where the bystander's referred point of interest was. There was either a lack of visual reference due to the immersed user wearing a HMD and in turn not being able to establish visual contact with the respective bystander, a lack of features in the bystander's verbal description, or both, thus preventing collaboration between the immersed *insider* and the non-immersed *outsiders*. Motivated by both the literature and these own experiences, a design space was created to investigate aspects of this subject within the overall context of CIA more closely, namely the concept of hybrid asymmetric collaboration as introduced in Section 6.1.

Under consideration of the developed second VE iteration, the existing IA interface was extended into a collaborative system that connects in real-time to a non-immersive interface, enabling bi-directional referencing to spatial data entities that facilitate the collaborators' awareness and their ability to verbally discuss and interpret the data together. Pairs of linguistics students used the interfaces to collaboratively investigate a large Twitter corpus with a focus on multilingualism in an explorative analysis task to obtain insights about the usage of language and hashtags within the Nordic region. Based on the results of this, within the scope of this thesis, first empirical evaluation with respect to the hybrid asymmetric collaboration concept, the overall presented approach and its designed collaborative features can be considered validated, allowing for reflections with respect to usability and collaboration.

Usability The reported usability scores for the two interfaces, presented in Section 6.3.4.3, point towards an *acceptable* usability for both, generally indicating that the participants were able to operate the interfaces. Thus, one may infer that the whole collaborative system, consisting of both interfaces, is indeed usable within the presented CIA context. Some of the reported lower usability scores for the VE may be attributed to the observed initial difficulty of learning the 3D gestural input (see Section 6.3.4.5). However, given that it was each user's first interaction with the IA interface, and the comparatively short exposure time in the immersive VE (approximately 10 to 15 minutes including warm-up), it is noteworthy that all of them learned to utilize the implemented features for their data analysis. Usability aspects in regard to utilizing 3D gestural input for various analysis tasks have been investigated and discussed more closely within the scope of the third VE iteration's interaction design (see Sections 5.5.2 and 5.5.5).

Collaboration The design of the developed collaborative features (see Section 6.3.2) hold characteristics of remote collaboration, even though the interface users were co-located in the same physical space and able to communicate verbally with each other without additional technological assistance. This is arguably another example illustrating that a simple classification of collaborative scenarios with respect to space and time, i.e., where and when collaboration takes place, becomes increasingly insufficient, as discussed in Section 2.3. Even though the student participants were physically co-located, they were not able to rely on important collaborative information cues due to the nature of the involved interface technologies. A main aspect of the developed collaborative system was concerned with the exploratory investigation of how they would approach their joint data analysis using the provided features. The results, based on log file analysis, researcher observations, and conducted group interviews, can be considered promising in various regards.

First, the students appeared very engaged in the activity itself as well as in the collaboration with their partners. In most cases, their verbal communication was frequent and natural, using a fair amount of close collaboration in order to analyze the data together using the two interfaces, each for their designated purpose. It was particularly interesting to listen to their choice of words, commonly making use of deictic terms and expressions, in combination with the referencing abilities as provided through both interfaces, allowing them to synchronize their data analysis contexts to establish a mutual understanding of where they are focusing their spatial data analysis efforts. Generally, the collaboration between all the students felt subjectively organic, and the system allowed them to approach the task in a rather flexible manner choosing their own data analysis strategies, for instance as illustrated in Figure 6.7. While some pairs approached the tasks rather enthusiastically and actively (some of them even continued with minor data exploration endeavors after the tasks were completed), a few appeared to be rather shy and conservative at first. Arguably, this can be attributed to the different personality traits of the students, individually as well as collectively, particularly with respect to Churchill and Snowdon's (1998) acknowledgment, stating "*(...) that cooperative and collaborative activities involve considerable negotiation, and teams vary tremendously in their negotiation strategies as well as in their task accomplishment process*."

Second, it was particularly interesting to investigate how the students would make use of the provided collaborative features. The detailed display of the VE user's position, orientation, and selection in real-time in the non-immersive desktop terminal enabled the *outside* student to be aware of the VE user's context, subjectively at all times, thus providing somewhat of a monitoring tool. Often the non-immersed student verbally acknowledged their awareness of the VE user's data analysis context. This allowed the VE user, visually immersed, to focus on the exploration of the virtual 3D space. However, the non-immersed student often used the provided spatial referencing feature, guiding the attention of the immersed student ad hoc to highlighted points of interest in the VE. Arguably, the non-immersed student had more (visual) awareness of their immersed partner than vice versa. In most of the cases, the VE user found something of interest and simply stated (along the lines of) "*look what I found*" or "*look where I am*" and their non-immersed partner was able to do so. These observations are in line with the log file analysis, confirming the relatively less frequent use of the dedicated spatial referencing feature by the immersed user (see Table 6.3). It appears that the collaborative information cues integrated in the map view of the non-immersive desktop terminal, which continuously provided updates about the immersed user's spatial data analysis context, were sufficient for the non-immersed interface user to follow along. Contemplating this matter and the collaborative features design, an interesting reflection can be made. While the non-immersed user was constantly aware of their immersed partner's context without the requirement for them to make dedicated spatial references by attaching a virtual pin to the respective data entity (see Section 6.3.3.2), the immersed user was more reliant on trusting the verbal acknowledgments of their non-immersed partner, stating that they followed along. After all, an indication about the non-immersed user's spatial data context, as "visual confirmation" so to say, was not automatically displayed in the immersive VE. Even though it did not prevent them from collaborating during their joint data analysis, based solely on visual information cues, the immersed student was arguably less aware of their non-immersive partner than vice versa.

Third, all student pairs were able to complete the explorative data analysis tasks and to make a variety of observations within the given scenario, as documented in Section 6.3.4.2. One of the two observing researchers, the domain expert within linguistics, was satisfied with the participants' answers, rating them as reasonable and along the lines of what is expected of first-year students in terms of complexity and critical thinking. Furthermore, it was pleasant to observe that the students' dialog with each other encouraged them to base some of the assessments on meta information, such as their own contextual knowledge. Generally, the interactions between the immersed and non-immersed students seemed to facilitate experiences of shared discovery, which has great potential for collaborative data analysis activities within educational contexts, such as the presented linguistics and higher education context where students are more and more frequently introduced to the large datasets that are produced in the humanities scholarship today.

6.4 Collaboration in VE Iteration 3: 3D Radar Charts

The results of the first exploratory investigation in regard to the concept of *hybrid asymmetric collaboration*, presented throughout Section 6.3, motivates to further examine the subject of bridging interactive InfoVis and IA. After all, the results of the conducted empirical evaluation validated the presented interaction and collaboration between a pair of users where one was, so to speak, *inside* a VE, while the other remained *outside*, each for their own dedicated data analysis role. To move further in this direction within the scope of this thesis, a collaborative data analysis scenario was designed that incorporates the latest immersive VE iteration, i.e., applying the *3D Radar Chart* data entity visualization and interaction design as described in Section 5.5. In particular, the data analysis scenario is centered around the Plant-Weather timelines (PWt) dataset, likewise introduced as part of Section 5.5 and illustrated in Figure 5.18. The dataset features correlated time-series data for various locations with respect to *plant* and *weather* data variables, and each plant data variable is either *positively* or *negatively correlated* to each of the two weather data variables. Using the PWt dataset, it is possible to design a collaborative task with the objective to analyze the data and identify

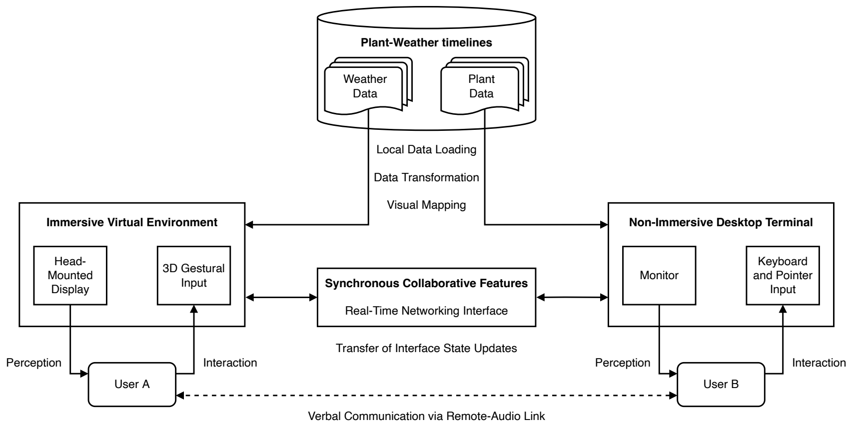


Figure 6.8: Conceptual system overview of the hybrid asymmetric collaboration setup under utilization of the third VE iteration (see Section 5.5). Detailed descriptions about the various system components are provided in Section 4.2.

these correlations by using the immersive and non-immersive interfaces as well as their collaborative features. In contrast to the initial exploratory investigation, the additional empirical efforts towards hybrid asymmetric collaboration, as presented throughout the remainder of this section, differ in several major aspects:

- Both the immersive and the non-immersive interface focus on the analysis of spatio-temporal data under utilization of appropriate visualization approaches, i.e., the 3D radar chart data entity visualization design as part of the immersive VE, and typical 2D data visualization techniques such as described by Ward et al. (2015, Chapters 6 and 7), Munzner (2014, Chapter 12), and Lundblad et al. (2010) as part of the non-immersive desktop terminal.
- Several generalized design options were compiled with the objective to guide the implementation of various spatio-temporal reference designs as visual, nonverbal collaborative information cues for the user in the immersive VE.
- The collaborative features to allow referencing across the interfaces are integrated more seamlessly through continuous signaling without the need to take dedicated actions to send discrete signals.
- A pair of collaborators utilizes the interfaces with the aim to complete a confirmative analysis task, i.e., a directed search to extract insights from the data (see Section 5.2).

- In addition to usability, the empirical evaluation of the collaborators' data analysis activity examines aspects of user engagement and collaboration in VEs (see Section 6.2), including a quantitative audio activity analysis of the collaborator pairs.
- Instead of being physically co-located, the collaborators assume a distributed remote setup where they communicate verbally via respective audio link, as illustrated as part of the overall collaborative system architecture presented in Figure 6.8.

6.4.1 Spatio-Temporal Reference Design

Within the context of a collaborative data analysis setup that utilizes an immersive VE centered around data visualization using the 3D radar chart design as presented throughout Section 5.5, the question arises of how to implement visual references. Visual references as nonverbal communication cues aim to assist the immersed collaborator to focus their attention towards a point of reference as indicated through the non-immersed collaborator. While the visual reference design as part of the collaborative setup around the second VE iteration was exploratory, utilizing a pillar approach to make spatial references in the VE (see Section 6.3.2), it is intriguing to investigate visual reference designs in a more structured manner, among others allowing for reflection on design considerations for CIA experiences.

A promising direction to enable collaboration in VEs, and thus support various collaborative information cues, is the use of avatars, i.e., a virtual representation of the other user(s) in the VE, either co-located or connected remotely (Xia et al., 2018; Steed and Schroeder, 2015). The visual design of such avatars has been investigated in a multitude of studies, for instance to determine differences between realistic and other types of avatar representations (Pakanen et al., 2022; Sun et al., 2019), to explore effects of nonverbal expression using highly expressive avatars (Wu et al., 2021), or to determine how avatar appearances influence aspects of communication and interaction (Heidicker et al., 2017). Collaboration and the use of avatars often imply that all users are utilizing some type of immersive technology to track and share the contexts in the 3D information space. However, such an approach is not always feasible for the collaboration on the same data across different interface types, for instance when using non-immersive display and interaction technologies that do not feature such capabilities. Consequently, within the scope of the presented hybrid asymmetric collaboration concept (see Section 6.1) and the design of collaborative information cues across immersive and non-immersive interface types, rather than utilizing an avatar-based approach, an abstract visual reference design is followed that is centered around the manipulation of the interface through visual artifacts that are similar to annotations, indicators, or alike.

Thus, three generalized *design options* for the creation of visual references in immersive data visualizations were defined (see Section 6.4.1.1). These design options are partially informed by related work around the subject of collaborative information cues in VEs, for instance as presented in Section 3.4.

Naturally, the reference design of collaborative information cues is inherently dependent on the applied visualization, it's purpose, and it's scenario, such as what type of data are displayed and using what technique. In other words, the reference design is dependent on the complexity and composition of the virtual 3D space. This is important in order to determine what to potentially signal to, allowing for the application of a reference design option accordingly. Within the presented scenario of spatio-temporal data analysis, the *task type* definitions as described by Andrienko and Andrienko (2006, Chapter 3) can be followed, differentiating between *elementary* and *synoptic* tasks. In particular, elementary tasks are concerned with the reference to one individual data entity, for instance, one location (spatial) or one point in time (temporal). Synoptic tasks are concerned with the reference to multiple data entities, such as a group of locations (spatial) or a range of multiple points in a time-series (temporal).

Under consideration of the 3D radar chart data entity visualization design and the different task types, the overall VE composition, described in detail in Section 5.5.1, can be summarized as follows. Individual 3D radar charts can be uniquely identified by their geospatial location (country). Thus, in regard to the spatial reference design, signaling to individual and multiple locations should be possible. Each 3D radar chart features multiple time-series data variables. Thus, it should be possible to refer to a single time event as well as to a range of consecutive points in time. Time event and time range references should be possible both across all time-series data variables as well as just for individual data variables. Based on this understanding of the VE composition, it is possible to design spatial and temporal references as collaborative information cues. Their implementation is described in Sections 6.4.1.2 and 6.4.1.3, allowing for subsequent empirical evaluation (see Section 6.4.4). Some of the presented reference designs are also utilized as part of the overall presented collaborative features design (see Section 6.4.3).

6.4.1.1 Design Options

To provide a reference that catches the user's attention and thus guides them towards an artifact, a signal is required that allows to be distinguished from the conventional environment. Within the scope of this thesis, the focus is set on *visual* sensory input, i.e., the manipulation of the immersive data visualization through means that allow its user to visually perceive references by looking around.[8] To guide the creation of visual references as collaborative information cues, three

[8]Naturally, other sensory input, for instance from auditory or somatosensory interfaces, may be used as collaborative information cues in VEs.

design options were defined: *Modify Artifact*, *Add Artifact*, and *Modify Environment*. These design options, as presented in Table 6.4, are held purposefully generic and low-level to allow their utilization across many different contexts and scenarios.

Modify Artifact (MA) The MA design option follows the concept of temporarily *modifying the visual appearance of the referred artifact*, aiming to distinguish it from all others accordingly. It is important that such a modification enables the user to detect the referred artifact, even if the process of the actual modification was not observed, i.e., the transition from normal to referred visual appearance. This is arguably particularly important within the context of VR experiences, as it cannot be guaranteed that the referred artifact is within the immersed user's field of view at all times, thus the change in visual state might not be observed. Depending on the complexity of the implemented modification to the (existing) artifact, this option can be considered to be comparatively friendly in regard to required computational resources, as it is likely that no new artifacts and geometry need to be added to the scene. At the same time, the visual alteration of the referred artifact should be carefully considered, as potential visual mapping and data encoding may be lost through the modification of its original appearance.

Add Artifact (AA) The AA design option follows the concept of temporarily *adding a visual artifact in close proximity to the referred artifact*, serving as a visual annotation. Such an added artifact should strive to (1) enable clear identification of the associated artifact it is intended to signal to, allowing the user in the VE to effectively focus on the referred artifact, (2) be easily detectable and distinguishable from all other artifacts in the scene, but at the same time (3) not obstruct or occlude other important information in the scene. Using this option, the visual appearance and integrity of the referred artifact is maintained in its original state, thus not loosing any potentially applied visual mapping and data encoding, which are likely to be relevant within analytical scenarios. Adding an artifact to the scene, although temporary, requires some practical considerations. For instance, depending on the added artifact's complexity, such as its geometry, its introduction to the scene will likely demand additional computational resources. Thus, it is important to ensure that this is implemented in a way as to avoid a noticeable impact in the immersive application's performance. Furthermore, based on the designer's and developer's assessments, temporarily adding (and removing) an artifact should be possible implementation-wise in a comparatively effortless and reasonable way.

Modify Environment (ME) The ME design option follows the concept of temporarily *modifying existing artifacts of the environment* that are in close proximity or can otherwise be directly associated with the referred artifact. As opposed to the AA option, rather than introducing an additional artifact to the scene, the ME option builds upon the utilization of existing elements or features in the computer-generated environment to establish a visual reference. Naturally,

Design Option	**Summary**	**Categorization of Presented Work**	**Examples from Related Work**
Modify Artifact	Modification of the visual appearance of the referred artifact.	node (see Figure 6.9), highlight (see Figures 6.10 and 6.11)	interaction availability filter (Casarin et al., 2018), ghost avatar (Lacoche et al., 2017)
Add Artifact	Addition of a visual artifact in close proximity to the referred artifact.	pillar (see Figure 6.9), pointer (see Figures 6.10, 6.11 and 6.12) symbol (see Figures 6.10, 6.11 and 6.13)	extended grid (Lacoche et al., 2017), light beam (Peter et al., 2018), virtual drone (Peter et al., 2018), virtual hand (Sugiura et al., 2018), marker (Welsford-Ackroyd et al., 2020)
Modify Environment	Modification of the environment around the referred artifact.	location (see Figure 6.9), outline (see Figures 6.10 and 6.11)	safe navigation floor (Lacoche et al., 2017), outline (Peter et al., 2018)

Table 6.4: Summary of the generalized design options to guide the implementation of visual, nonverbal references, including a categorization of the implemented spatial and temporal reference designs as well as the designs described in various related work.

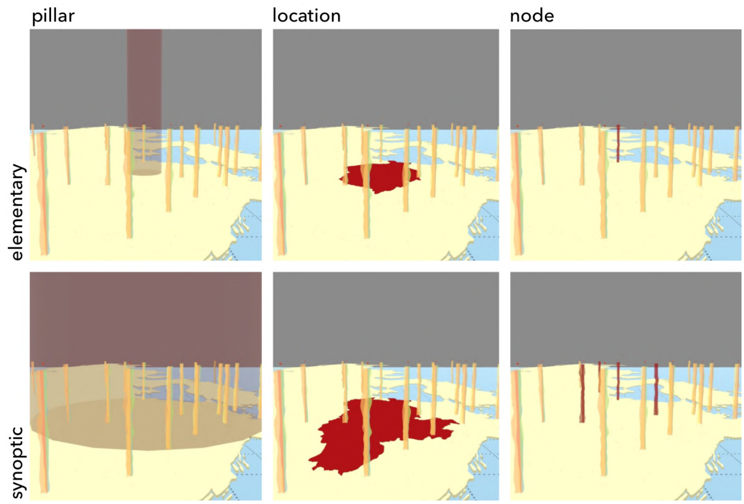


Figure 6.9: Impressions of the three spatial reference designs implemented in the third VE iteration, from the immersed user's field of view, and presented as elementary and synoptic task configurations.

this requires that the overall visualization and virtual scene are complex enough to provide such modification opportunities to begin with. If this requirement is met and the environment indeed provides artifacts or alike that allow for a semantic inference to the referred artifact, their appearance may be modified in order to act as a visual signal accordingly. Generally, this approach is likely to have a comparatively low computational impact as it utilizes existing artifacts and geometry that already exist. At the same time, the original visual integrity of the referred artifact is maintained (as opposed to the MA design option).

6.4.1.2 Spatial Reference Design

For the purpose of referring to specific 3D radar chart instances in the VE, three different spatial reference approaches were designed in accordance to the presented design options (see Figure 6.9). First, the *pillar* design follows the AA option, creating a semitransparent cylinder object with the 3D radar chart at its center, surrounding it accordingly. The pillar's height is scaled to make it appear to "shine from the top down" in the VE, similar to a spotlight. It is noteworthy that this design was also previously used in the exploratory investigation of hybrid asymmetric collaboration within the context of the second VE iteration (see Section 6.3.2) and as a bookmark in the first VE iteration (see Section 5.3.2). Second, the *location* design follows the ME option, modifying the color of the extruded country polygon on the virtual floor that each 3D radar chart is directly associated with. Third, the *node* design follows the MA option, visually separating the referred 3D radar chart from all others by uniquely coloring all its data variable axes.

6.4.1.3 Temporal Reference Design

For the purpose of referring to individual time events and time ranges of multiple consecutive time events within a 3D radar chart, four different temporal reference approaches were designed in accordance to the presented design options (see Figure 6.10). First, the *highlight* design follows the MA option, visualizing a colored mesh for the time event across all data variables as elementary task reference, respectively coloring the time range segments in each data variable axis as synoptic task reference. Second, the *outline* design follows the ME option, creating a closed visual loop along the outside of all included time events. Third, the *pointer* design follows the AA option and adds two artifacts to the visualization as reference, i.e., each included time event is encapsulated by a small visual sphere, further assisted through a juxtaposed 3D pointer model that directly indicates the respective time event or time range. Fourth, a *symbol* design was implemented, also based on the AA option and following a similar approach as the pointer design. However, instead of a pointer, the virtual sphere is juxtaposed with a symbol that can be interpreted by the user to infer further meaning. As a first exploratory illustration, the decision was made to use a magnifying glass symbol in a "*let us investigate this [temporal reference]*" analogy.

The complexity of the 3D radar chart data entity visualization design allows for different temporal reference configurations. Making a reference across all data variables is illustrated in Figure 6.10, whereas references in individual data variables are presented in Figure 6.11. Additionally, different configurations of the pointer and symbol designs are explored, as presented in Figures 6.12 and 6.13. The placement of the pointer indicator could encode further analysis related information, such as through a *neutral*, *positive*, or *negative* pointing direction, for instance to provide a comparison to a prior data value or to indicate an overall trend across the referred time range. Similarly, the application of different symbols could indicate an additional collaborative information cue, for instance to provide a reason why the collaborator is making a reference in the first place, such as to *investigate*, because they found something they deem *exciting*, or because they want to further *talk* about the referred data.

6.4.2 Non-Immersive Desktop Terminal

Within the presented context of the PWt dataset and in alignment with the collaborative data analysis scenario, the non-immersive desktop terminal is designed as an interactive InfoVis. Its main purpose is to enable its user to explore the two weather data variables, i.e., sunlight and humidity, over time

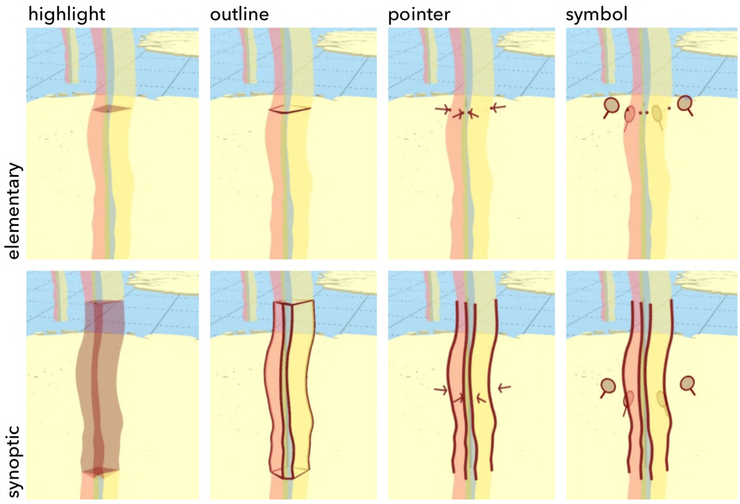


Figure 6.10: Impressions of the four temporal reference designs implemented in the third VE iteration, from the immersed user's field of view, and presented as elementary and synoptic task configurations across all data variables.

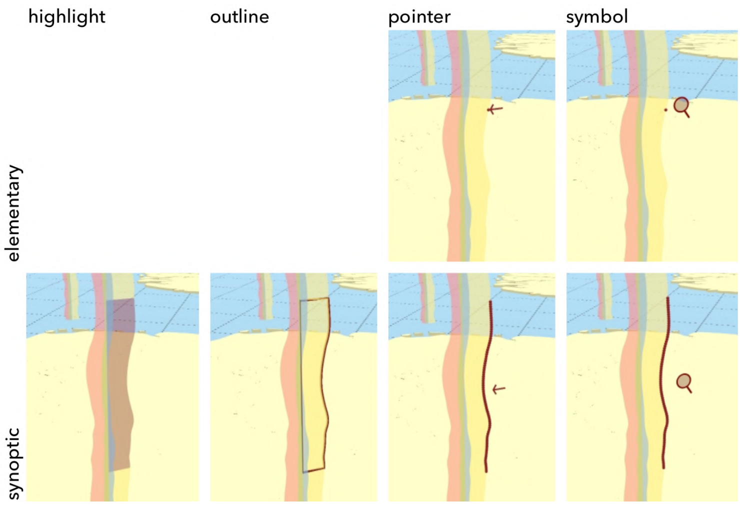


Figure 6.11: Impressions of the four temporal reference designs implemented in the third VE iteration, from the immersed user's field of view, and presented as elementary and synoptic task configurations for one individual data variable. **Note**: Due to the nature of the *highlight* and *outline* designs with respect to the 3D radar chart data entity visualization, implementations for a reference as elementary task configuration for one individual data variable are not available.

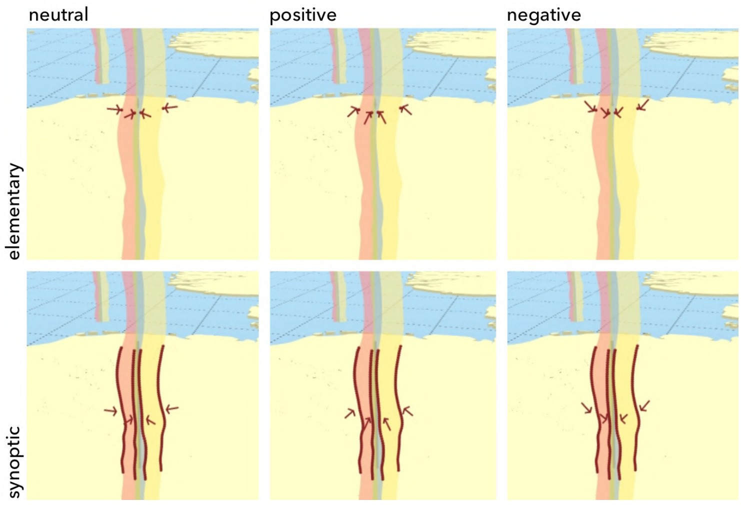


Figure 6.12: Impressions of the three different configurations for the temporal *pointer* reference design implemented in the third VE iteration, from the immersed user's field of view, and presented as elementary and synoptic task configurations across all data variables.

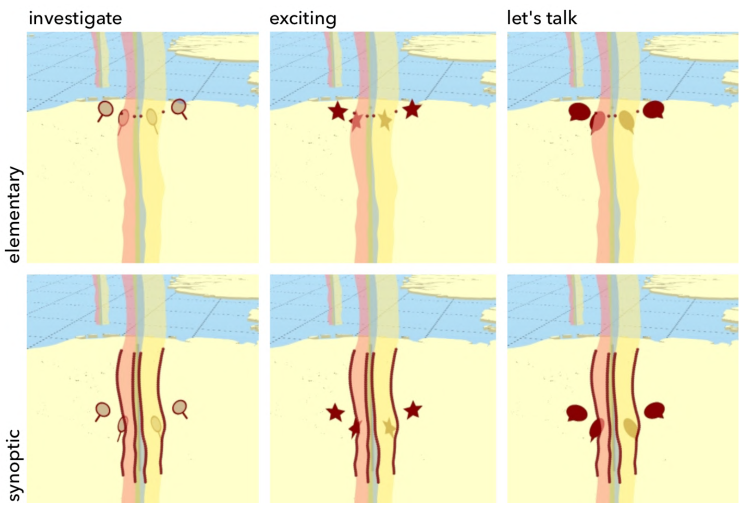


Figure 6.13: Impressions of the three different configurations for the temporal *symbol* reference design implemented in the third VE iteration, from the immersed user's field of view, and presented as elementary and synoptic task configurations across all data variables.

(temporal) and in regard to the various different country locations (spatial). The interface is held purposefully minimalistic but representative given the dataset and analysis task, applying typical views and visualization techniques for geospatial and time-series data, for instance as described by Ward et al. (2015, Chapters 6 and 7). In fact, Lundblad et al. (2010) described an interface for the interactive visualization of Swedish road weather data that features a similar setup compared to the presented non-immersive desktop terminal. The decision to hold it rather minimalistic is inherent from the intention to focus on the integration of interactive views that are relevant within the data context and will contain collaborative features. Figure 6.14 provides an overview of the non-immersive desktop terminal's view composition.

View Composition and Interaction Generally, the non-immersive desktop terminal is operated through a normal desktop monitor using keyboard and pointer (mouse) input, similar to the previous one that was used to explore hashtags (see Section 6.3.1). The interface is composed of two main views. First, a *Map View* is placed on the right part of the interface, displaying outlined all countries across Europe, including an interactive circle at the center of each country that enables the user to *select* it through a left-click interaction. Naturally, as each of these circles represents an individual data entity, it is noteworthy that their placements on the map are in line with placement of the individual 3D radar charts in the VE (see Section 5.5.1). A selected data entity is indicated through a colored outline in the map view. Once such a selection has been made, the *Weather View* on the left part of the interface is updated. The weather view itself is composed of two line graphs, one for the visualization of the sunlight time-series data of the selected country, and one for humidity. Each line graph's horizontal axis encodes time, while the vertical axis encodes the respective data variable value. Once the user *hovers* over a line graph, a vertical dashed *Preview Line* is displayed, providing visual feedback in regard to the hovered time event. The user may select the hovered time event via left-click. Additionally, the user may also *select a time range*, i.e., a series of multiple consecutive time events. With a single time event selected as a starting point and utilizing a combination of holding the COMMAND-key and left-click, the user can select a second time event as the respective end point, effectively establishing a time range. The selected time range is indicated through a visual overlay in the interface. Time event and range selections can be updated by simply applying new selections in the interface, replacing the prior ones accordingly. Furthermore, these selections are synchronized across the two line graphs of the weather view, i.e., a selection made in the sunlight line graph will automatically display the same selection in the humidity one, and vice versa.

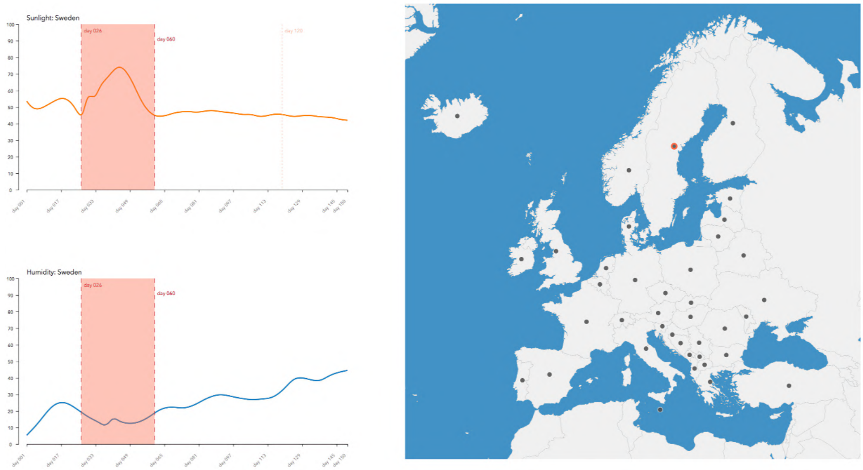


Figure 6.14: An impression of the implemented Non-Immersive Desktop Terminal to explore spatio-temporal weather data, and within the context of the hybrid asymmetric collaboration setup with the third VE iteration. **Right**: Map View. **Top Left**: Weather View - Sunlight, including hovered Preview Line and applied Time Range Selection. **Bottom Left**: Weather View - Humidity, including the synchronized Time Range Selection.

6.4.3 Collaborative Features Design

To facilitate the collaborators' verbal communication to support the way they discuss, interpret, and make meaning of data during their joint analysis activity, independent of their interface type, various collaborative features are needed. Naturally, the design of these features is informed by the previously obtained insights and experiences described in Section 6.3.5. However, in comparison to the collaborative information cues implemented in the prior interfaces (see Section 6.3.2), there are a couple of key aspects that need to be addressed, specifically in regard to the collaboration around the PWt dataset and under utilization of the presented interfaces. Arguably most noticeable, while the prior features focused on the support for spatial referencing, the new collaborative feature set additionally requires means to establish temporal references, i.e., indicating a time event or time range at a specific data entity. Furthermore, the spatial referencing strategy implemented in the prior interface featured rather discrete signaling characteristics, i.e., each of the users had to take a dedicated action (*bookmark*) to indicate a location for the respective collaborator. At times, this subjectively led to some minor waiting times and interruptions when one user was waiting for the referencing action of the other. To address this matter, the new collaborative features are implemented utilizing a more continuous signaling strategy, aiming to make the collaborators aware of all their respective interactions without the need to send dedicated signals. To summarize, the presented collaborative features across the immersive VE and the non-immersive desktop terminal are designed to support the users' hybrid asymmetric collaboration with the following main aspects in mind:

- *Common Ground and Awareness*: Facilitate the users' mutual understanding of their in-situ spatial and temporal contexts during their joint data analysis activity.
- *Reference and Deixis*: Support for the transmission and reception of spatial and temporal references as nonverbal communication cues in each of the collaborators' respective interfaces during their joint data analysis activity.
- *Continuous Signals*: Integration of the collaborative features in a seamless and ubiquitous manner, aiming to add collaborative information cues to the respective interfaces without unnecessarily increasing the complexity of their operability.

Collaborative Information Cues: VE to Desktop Terminal The following collaborative information cues *from the immersive VE* are displayed *in the non-immersive desktop terminal* (see Figure 6.15 right), addressing REQ 16 and REQ 17 in Table 4.1. Following the same design as implemented in the previous collaborative features, the map view displays in real-time the position, orientation, and field of view of the immersed user, enabling the non-immersed collaborator to have an understanding of their immersed partner's location in space. The data entity that the user in the VE is potentially actively interacting with (details-on-demand) is outlined in the map view as well, providing an indication about the immersed user's status accordingly. If both collaborators have selected the same data entity, the weather view features a vertical dashed line that represents the VE user's current time event selection, i.e., the position of the *time slice* in the 3D radar chart. Similarly, if the VE user applied a time range selection, an overlay in both of the line graphs is visualized in the weather view. All interface elements representing information cues of the VE user are color-coded differently to easily discern them from the desktop terminal user.

Collaborative Information Cues: Desktop Terminal to VE The following collaborative information cues *from the non-immersive desktop terminal* are displayed *in the immersive VE* (see Figure 6.15 left), addressing REQ 16 and REQ 17 in Table 4.1. Location selections made in the map view will temporarily highlight the corresponding country, extruded in 3D on the floor in the VE, following the *location* spatial reference design as presented in Section 6.4.1.2. It is noteworthy that if the VE user is already interacting with data entity at the selected location,

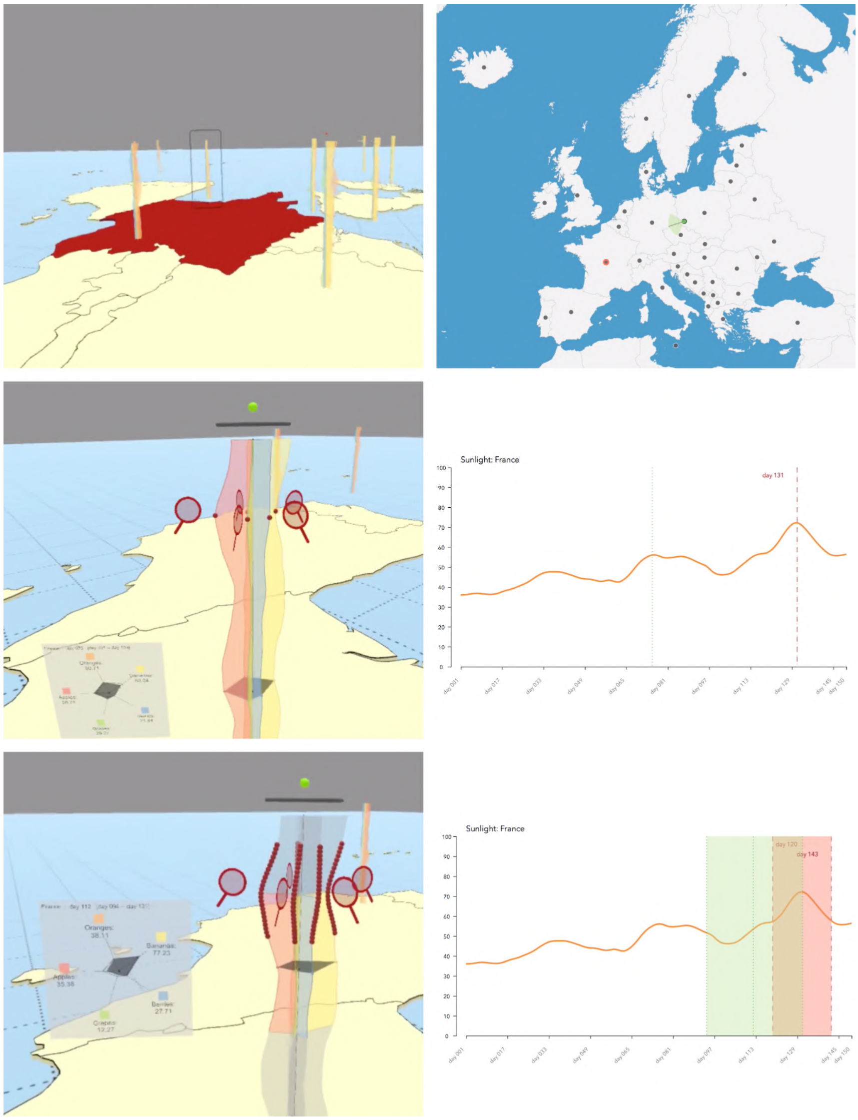


Figure 6.15: Impressions of the implemented collaborative features design across the two interfaces, and within the context of the hybrid asymmetric collaboration setup with the third VE iteration. **Left**: Immersive VE, with visual references as signaled from the non-immersed user. **Right**: Non-Immersive Desktop Terminal, with visual references as signaled from the immersed user. **Top**: Spatial reference (location). **Middle**: Temporal reference (time event). **Bottom**: Temporal reference (time range). **Note**: Left and right interface views are synchronized, illustrating both interfaces at the same point in time during the collaborative analysis activity.

the spatial reference will be omitted. Time event and range selections made in any of the two line graphs in the weather view will be represented as virtual annotations in the corresponding 3D radar chart following the *symbol* temporal reference design as presented in Section 6.4.1.3, aiming to catch the VE user's attention and indicating that the non-immersed user is currently *investigating* this temporal context.

6.4.4 Evaluation 1: User Preferences of Reference Designs

In order to gain insights about the various spatial and temporal reference designs as documented throughout Section 6.4.1, an empirical evaluation was set up with the objective to subjectively assess user preferences in general. The different reference designs are intended to function as visual, nonverbal collaborative information cues that support the immersed user in their ability to focus their attention towards a point of reference as indicated by a non-immersed collaborator. Thus, this evaluation is concerned with the assessment of the different reference designs in regard to the immersed user's subjective response, in particular with respect to *aesthetics*, *legibility*, and *general user preference*. The gathered insights should allow for respective reflections on the reference designs, and thus provide impulses that can guide the design of similar collaborative information cues in the future.

6.4.4.1 Physical Study Space

All study sessions, planned as one-on-one sessions involving the participant and the researcher, were conducted at the VRxAR Labs research group lab at Linnæus University. As the researcher would systematically present the different reference designs to the participant for assessment, no additional "real" collaborator was needed. Therefore, the overall physical study space was organized similarly as initially described in Section 5.3.3.1. For the immersion in the VE, the participant was utilizing a HTC Vive HMD and a Leap Motion Controller attached in front of it to enable 3D gestural input. Furthermore, as this evaluation was conducted during the COVID-19 pandemic, some additional practical measures were implemented as described in Section 1.4.

6.4.4.2 VE Setup

The VE was set up in accordance to the overall presented collaborative data analysis scenario, utilizing the 3D radar chart data entity visualization design and the PWt dataset as described in the introduction to Section 6.4. While the Leap Motion Controller was active and as such tracking and displaying the immersed user's hands in the VE for their general reference, no additional interactive features were provided during this evaluation. Overall, a total of 39 3D radar charts were placed across the different European countries. Each 3D radar chart

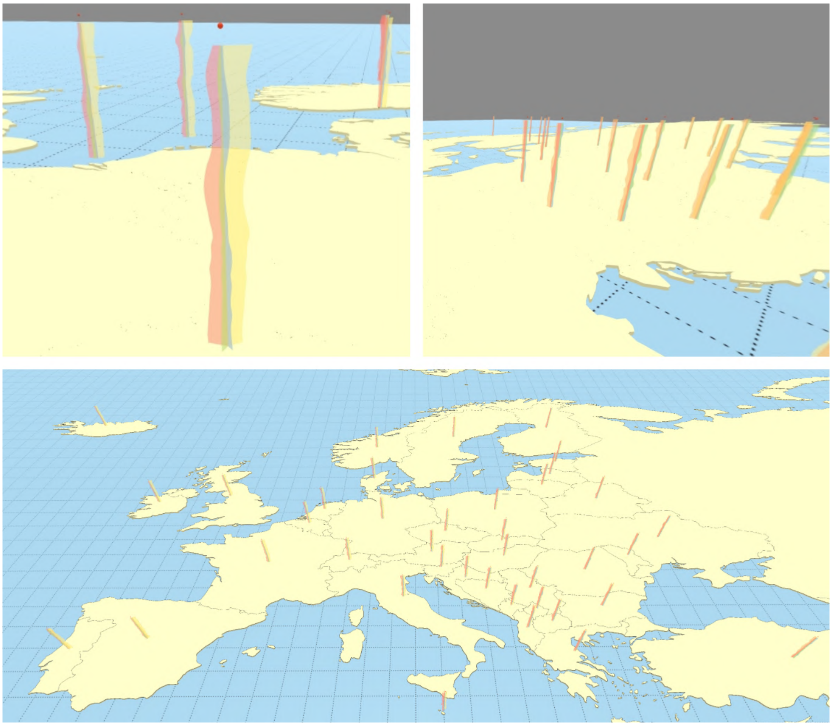


Figure 6.16: Impressions of third VE iteration, set up for the evaluation of the user preferences in regard to the various spatio-temporal reference designs, and within the context of the hybrid asymmetric collaboration setup. **Top Left**: The immersed user's field of view, looking at the 3D radar chart in close proximity (within their safe interaction area). **Top Right**: The immersed user's field of view, looking at faraway 3D radar charts. **Bottom**: The VE composition, from an angled top down view to provide an overall impression.

featured five data variable axes, each comprised of a time-series with 150 events. The immersed user was placed in Central Europe, able to move freely within the calibrated two-by-two meter area, and make observations. There was one 3D radar chart placed within their safe interaction area in the VE, all others were beyond their reach. Figure 6.16 provides an overview of the set up VE.

More specifically, each 3D radar chart in the VE featured a total height, i.e., vertical length, corresponding to 100 cm. All 3D radar charts were placed to float 40 cm above the virtual floor, thus reaching an effective height at 140 cm. The 3D radar chart used to display all temporal reference configurations was placed directly at the center of the immersed user's calibrated two-by-two meter area, enabling them to move freely around to investigate that 3D radar chart from all sides if so desired. The distance between the center of the immersed user's safe interaction area and the spatial reference for the elementary task (see Section 6.4.4.3) was approximately 848 cm in the virtual space. Furthermore, the distances between the center of the immersed user's area and the spatial references for the synoptic tasks (see Section 6.4.4.3) were approximately 573, 645, 848, and 1024 cm. These distances resulted from the properties of the underlying PWt dataset used within the study, as introduced in Section 5.5.

6.4.4.3 Task

To assess aesthetics, legibility, and general user preference for the different reference designs, presets for all the configurations were prepared, asking each participant to rate them by *speaking aloud*, enabling the researcher to document their ratings accordingly via pen and paper. In particular, assessments were inquired for all the elementary and synoptic spatial and temporal references as illustrated in Figures 6.9, 6.10, and 6.11. Additionally, user assessments for the temporal *pointer* and *symbol* reference designs in general were inquired, i.e., one combined rating for each, as illustrated in Figures 6.12 and 6.13. The reference presets were configured to make the same reference, i.e., referring to the same location/locations (spatial) or time event/range (temporal). The inquired assessments from each participant were collected in random order throughout two stages. First, they were presented with the individual reference presets and tasked to *rate* their perceived aesthetics and legibility on a 7-point Likert scale as follows:

- Aesthetics: *On a scale from 1 (aesthetically unpleasing) to 7 (aesthetically pleasing), how would you rate this design of [spatial / temporal] referencing?*
- Legibility: *One a scale from 1 (not at all) to 7 (very well), how clearly can you determine what is [spatially / temporally] referenced?*

Second, once they provided numerical assessments for all the presented reference designs, a general user preference for one over the other in a series of pair-wise comparisons was inquired. In particular, for each logical category, i.e., spatial elementary, spatial synoptic, temporal elementary, and so on, all possible pair permutations were created, and the participant was asked for each:

- General User Preference: *Out of these two [spatial / temporal] reference designs, which one do you prefer?*

6.4.4.4 Measures

A pre-task questionnaire was administered with the objective to obtain some overall demographic information about the participant sample, in particular with respect to their educational/professional background and their self-assessed prior experiences with VR technologies. The empirical evaluation of the different reference designs is centered around the subjective assessment with respect to three measures, i.e., *aesthetics*, *legibility*, and *general user preference*, as previously introduced in Section 6.4.4.3. Within the scope of this evaluation and the presented context, *aesthetics* is defined as how much one appreciates the visual design and appeal of a reference design, and whether or not one finds it is pleasing and beautiful to look at. *Legibility* is defined as how well one can understand, determine, and detect what is highlighted, and how clear it is what to focus on. Collecting quantitative ratings for these two measures for each reference design should allow for respective comparison. Additionally, the third metric is concerned with the *user's general preference* based on pair-wise comparisons, indicating which one of two reference designs the user would rather work with if they were to use such an immersive data analysis environment with collaborative features frequently. The results of these pair-wise comparisons allow for respective tallying,[9] providing an overall preference indication as well as functioning as a potential tie breaker between two designs in the case of equal aesthetics and legibility ratings.

6.4.4.5 Study Procedure

Each study session, aimed to take approximately 45 minutes in duration, followed the same procedure of three stages:

1. Introduction (10 min);
2. Warm-up (5 min in the VE);
3. Task 1: Aesthetics and Legibility Ratings (15 min in the VE);
4. Task 2: Pair-wise Preference Comparison (15 min in the VE).

During the *introduction*, each participant was welcomed and asked to complete an informed user consent and pre-task questionnaire. Afterwards, the researcher presented the overall context and scenario of the immersive VE in regard to its analytical and collaborative aspects, ensuring that each participant understood the 3D radar chart data entity visualization design as well as the composition of the VE. Participants were then provided with a brief *warm-up*, allowing them to familiarize themselves wearing the HMD and with the VE. Once they felt comfortable, they proceeded to the *task* stages. First, based on random order, a participant's *aesthetics and legibility ratings* for the different spatial and temporal reference designs were inquired via Think Aloud technique. Second, based on random order, each participant provided distinct answers for each *pair-wise preference comparison*, selecting one reference design over another. The researcher

[9]The pair-wise preference comparison is inspired by the prior experiences of utilizing the Task Load Index (TLX), in particular its *weighing* process, as described in Section 2.5.1.

noted all ratings and preferences on a predefined task answer sheet. Throughout both task stages, the participants were allowed to provide additional remarks as desired that were also noted by the researcher. Finally, they were thanked for their participation and sent off.

6.4.5 Results of Evaluation 1

6.4.5.1 Participants

For the described user preference evaluation, a total of $n = 12$ participants from a mixture of different academic backgrounds (5 *Computer and Information Science*, 5 *Linguistics and Language Studies*, 2 *Forestry and Wood Technology*) were recruited. Eight participants reported to have just *a little* prior experience with VR technologies, three *average*, and one *a lot*. During the warm-up phase in the VE, none of the twelve participants reported to have any visual perception issues, neither in regard to the 3D radar chart data entity visualization design and the applied color coding of the various data variable axes nor with respect to the VE composition in general.

6.4.5.2 User Assessments and Remarks

Figure 6.17 presents the results of the participants' ratings for the aesthetics and legibility measures as defined in Section 6.4.4.4. The results of the pair-wise user preference comparison, combined with the rating medians, are presented in Figure 6.18.

Some participants provided additional remarks for the different spatial and temporal reference designs. For instance, participants stated that they can envision usefulness and relevance for both temporal pointer and symbol reference designs in a real-world scenario (see Figures 6.12 and 6.13). According to them, the pointer design subjectively represented more precision and urgency, while the symbol one was easier to recognize and featured better clarity. One participant stated that the pointer design makes more sense during synchronous collaboration (as the users are likely talking to each other), while the symbol design may be better suited for asynchronous collaboration (encoding an additional semantic meaning). It was also noted that a visual connection to the 3D radar chart's center, i.e., its time axis presenting the origin, is missing and would be preferred in the synoptic reference designs (pointer and symbol), as included in the outline design for an individual dimension reference (see Figure 6.11). Some participants were unsure whether the pointer for the synoptic referencing task was referring to the entire time range or to one specific time event within it. One participant expected the pillar design to be multiple small ones as opposed to one large one when making a spatial synoptic reference (see Figure 6.9). Another one stated that it might increase the legibility of the spatial location reference design by additionally slightly extruding the respective 3D country model on the virtual floor. Similar as

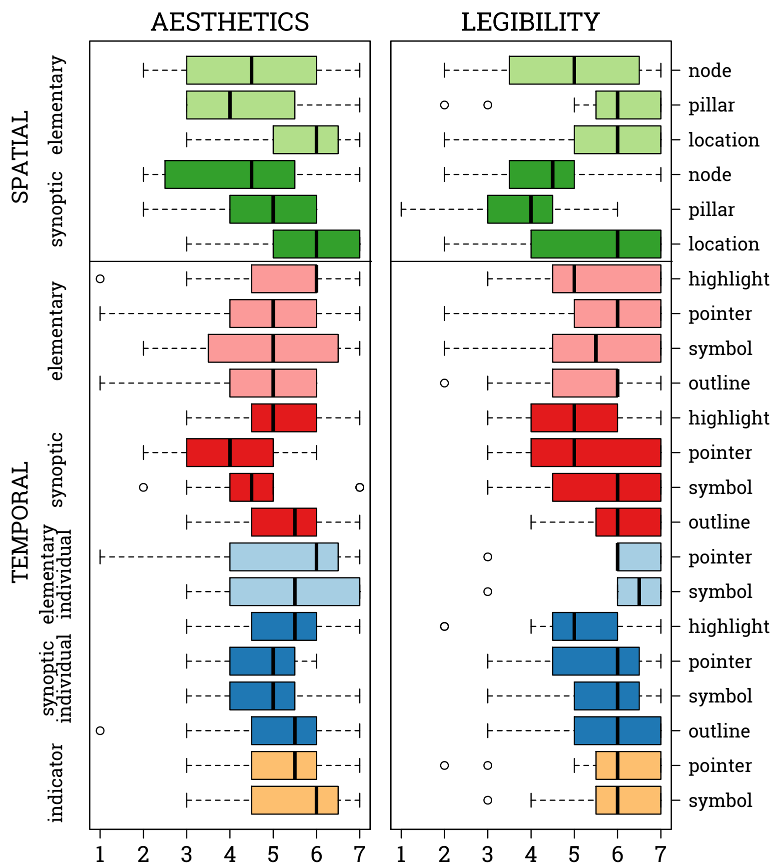


Figure 6.17: Results of the aesthetics and legibility ratings for the implemented spatial and temporal reference designs based on the $n = 12$ participants' subjective assessments.

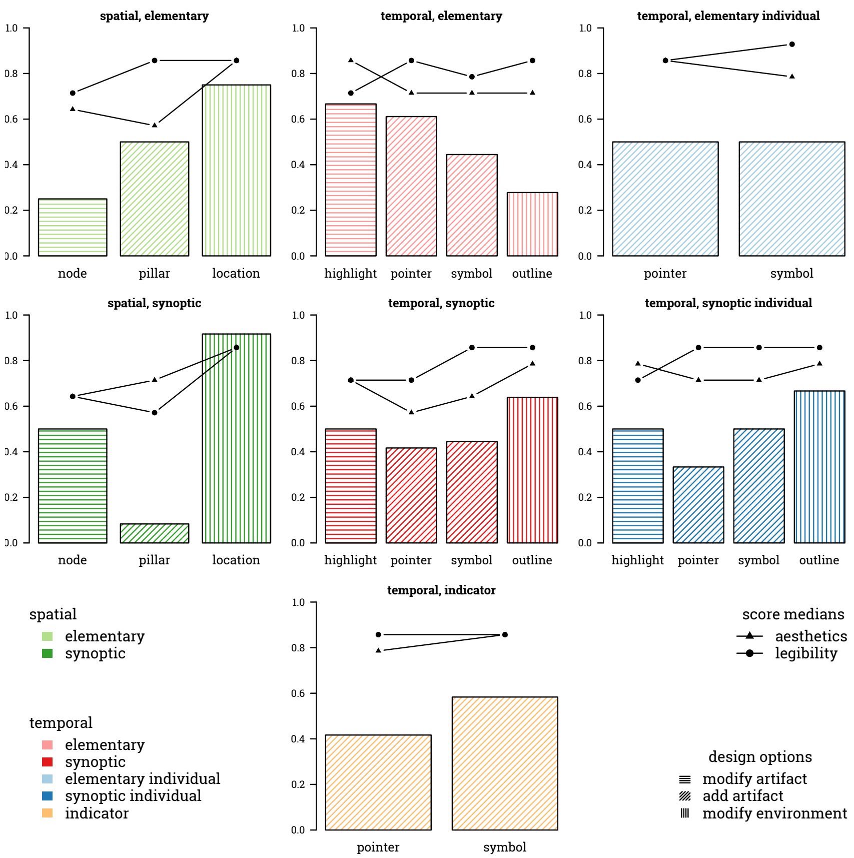


Figure 6.18: Results of the general user preference comparison for the implemented spatial and temporal reference designs based on the $n = 12$ participants' subjective assessments. A total of seven categories were examined based on the applied pair-wise comparison of the different spatial and temporal reference designs. The score medians for the aesthetics and legibility ratings are included (see Figure 6.17). *Plot created by Aris Alissandrakis.*

in the investigation reported by Peter et al. (2018), participants made suggestions for some hybrid designs, combining two of the presented designs into one, such as pillar+node, location+node, highlight+pointer, outline+symbol without the visual spheres, and pointer without the pointer using just the visual spheres – similar to the marker design presented by Welsford-Ackroyd et al. (2020).

6.4.6 Evaluation 2: Collaborative Confirmative Analysis

Within the scope of this thesis, a final empirical evaluation was designed with the objective to further investigate the overall hybrid asymmetric collaboration concept as introduced in Section 6.1. The overall technological setup, aligned with the presented conceptual system overview (see Figure 6.8), is composed of the third VE iteration, a non-immersive desktop terminal, and various collaborative features, as described throughout Sections 5.5, 6.4.2, and 6.4.3. Compared to the initial collaboration evaluation within the context of the second VE iteration (see Section 6.3), the design of this final evaluation differs in various major aspects as outlined in the introduction of Section 6.4. Arguably most noticeable besides utilizing the third VE iteration and the PWt dataset, it is centered around a confirmative analysis task (directed search with some hypotheses given; see Section 5.2) that demands collaborative analytical reasoning and data interpretation in regard to both spatial and temporal contexts. To evaluate the developed collaborative data analysis system, pairs of participants were recruited, alternating the roles of one person being immersed in the VE, and the other using the non-immersive desktop terminal (*within-subject* design). As the PWt dataset and its overall approachable and easy to understand scenario was utilized, there were no specific domain knowledge or other expert requirements, allowing for an inclusive participant recruitment.

6.4.6.1 Physical Study Space

As with all other empirical evaluations, this study too was conducted at the VRxAR Labs research group lab at Linnæus University, featuring an overall similar physical study space compared to the initial collaboration evaluation (see Section 6.3.3.1). The immersive interface was based on a HTC Vive HMD with a Leap Motion Controller, and the non-immersive interface utilized a desktop computer with a 27-inch monitor as well as keyboard and pointer (mouse) input. However, there were a few key differences as follows. Instead of two researchers, one was responsible for the study conduction, including its moderation and data collection. Furthermore, instead of the participant pair being co-located in the same room, a distributed remote setup was applied. In particular, the setup for the immersed VE user remained unchanged in the research group lab, while the non-immersed user's workstation was set up in a dedicated separate office room, where they were seated alone and free of distractions. A remote audio link utilizing the Zoom Cloud Meetings[10] teleconferencing software was installed on both the researcher's and the non-immersed user's workstations, allowing for verbal communication between the two physical locations (REQ 18 in Table 4.1). The researcher remained generally at their workstation during the evaluation, while the participants changed the locations once as part of iterating their assigned user roles across the task completion stages. The conduction of this evaluation during the COVID-19 pandemic required the implementation of some additional practical measures as described in Section 1.4.

6.4.6.2 Interfaces Setup

Within the scope of this evaluation, the immersive VE was generally set up as introduced and described throughout Section 5.5, utilizing the 3D radar chart data entity visualization design to display the PWt dataset. More specifically, the VE setup featured an overall identical setup as described in Section 5.5.5.2, documenting the VE setup for the empirical evaluation of the uniform 3D gestural design, with two exceptions. First, all implemented interactive features were available except the *Zoom (in/out)* feature, which was deemed expendable within the context of the presented collaborative analysis task. And second, the VE was extended to include all the designed collaborative features in accordance to the descriptions in Section 6.4.3. It is noteworthy that the choice of the applied reference designs, i.e., the *location* design as spatial reference and the *symbol* design as temporal reference, was informed by the results of the user preference evaluation (see Section 6.4.5).[11] The non-immersive desktop terminal was available as described in Section 6.4.2, including all its collaborative features.

6.4.6.3 Task

Under consideration of the hybrid asymmetric collaboration concept, the pair of participants used different interfaces (hybrid) that additionally affected which part of the PWt dataset they had access to, thus determining their role (asymmetric). The non-immersive desktop terminal provided access only to the weather data variables, while the immersive VE provided access only to the plant data. Thus, the pair assumed the roles of *weather* and *plant* expert respectively. Their individual roles would iterate as they switched interfaces for the second task.

Generally, for one confirmative analysis task, the pair was asked to collaboratively explore the presented weather and plant data in space and time, and to use the provided tools to make assessments that describe the relationship between each plant and the two weather variables (sunlight and humidity). More

[10]Zoom Video Communications, Inc. Official Website. Retrieved June 1, 2022, from `https://zoom.us`

[11]For the reference design as part of this evaluation, a uniform design strategy was followed (instead of applying a mixture of visually different designs). Both the *location* and the *symbol* reference design received positive feedback in the respective user preference evaluation.

specifically, they were tasked to determine the *type of correlation* between each weather and plant data variable, and indicate how *confident* they were with their assessments. The non-immersed user was additionally tasked to write down their joint answers on a prepared physical task answer sheet (see Appendix D).

The PWt dataset, and more specifically the *correlated-timelines* project as described in Section 5.5, was utilized to generate not one but two datasets, i.e., one featuring *fruits* and one featuring *vegetables* as plant data variables. Naturally, the respective correlations between the fruits/vegetables and the two weather data variables differed in each of the two datasets.[12] Utilizing these datasets, the pair could be asked twice to conduct an overall identical data analysis task, allowing each user to operate each interface type once, without any potential insights transfer in regard to the data between the tasks. Although the overall data scenario was easy to understand and relate to, i.e., climate and flora at different locations across Europe and how the weather conditions affect the plant growth, any previous knowledge of geography and agriculture had to be dismissed. Therefore the study was presented with a "science fictional" description to the participants that had to suspend their disbelief and pretend that they were exploring a parallel universe in the far future instead of working with real observations that followed known phenomena (see Appendix D).

6.4.6.4 Measures

A pre-task questionnaire was applied to inspect some general demographic information about the participant sample, i.e., their educational/professional background as well as their self-assessed prior experiences with VR technologies. A mixture of task performance and subjective methods was applied to collect quantitative and qualitative data that enable the subsequent assessment of various relevant aspects within the presented hybrid asymmetric collaboration context. More specifically, system logs of all the participants' interactions with their respective interfaces during the task stages were collected. This should allow for analysis in regard to various concerns, for instance to identify when the collaborators were actively investigating the data in the same spatial location, or in regard to when each collaborator moved from location to location, to name just two examples. Furthermore, based on the designed confirmative analysis task as described in Section 6.4.6.3, the pair's ability to collaboratively identify the potential correlations between the plant and the weather data variables can be assessed. For each task, this results in a total of ten correlation answers, i.e., five plant data variables × two weather data variables, each indicating either a *negative*, *positive*, or *no* correlation. While the option to answer *no correlation* was provided, a correlation was always defined as part of the generated datasets across both the fruits and the vegetables scenario. Each of the ten correlation answers included an associated confidence (*low*, *medium*, *high*, or *do not know*), describing the pair's

[12] Both datasets are included in the *correlated-timelines* project repository (see Appendix D).

reported confidence for their respective answers. The researcher kept notes based on observations made during the pair's task completion.

To make assessments about the usability of each interface as part of the collaborative system, the SUS questionnaire was administered. Additionally, the User Engagement Scale - Short Form (UES-SF) questionnaire was utilized to obtain insights about the collaborators' general engagement with their respective interface. Both questionnaires are described in Section 2.5.1. Assessing system usability and user engagement within the scope of the presented collaborative task allows to obtain further insights into the implemented data analysis interfaces and their collaborative features. While both interfaces are using fundamentally different display and interaction technologies, it is important to potentially identify factors that might impact the pair's collaboration. The results of both questionnaires may indirectly contain some assessments in regard to the pair's collaboration, i.e., through the implemented collaborative features as part of the interfaces. However, it is arguably beneficial to inquire some more focused and structured feedback with respect to important collaboration aspects. For this purpose, the self-constructed STCQ was administered (see Appendix D), as introduced and described in detail in Section 6.2.

Finally, as the participants were located in two physically separated locations during the task stages, means of verbal communication were required – besides the nonverbal communication features as part of the provided system, i.e., the spatio-temporal referencing. As described in Section 6.4.6.1, the Zoom Cloud Meetings teleconferencing software was utilized for that purpose. In addition to allowing the participants to talk to each other via an audio call, it also provided the opportunity to record their conversation for each task stage. In fact, Zoom conveniently allows the recording of separate audio streams for each call participant. Thus, at the end of each task stage, three audio files (one of the combined audio, and one from each user) were obtained. Using the Audacity[13] audio editor software and its *Sound Finder* tool, it is possible to obtain timestamps that describe when sound was detected[14] in each participant's audio file, and therefore roughly when they were (individually) speaking. Summing up the time intervals provides an estimation of each participant's *speaking* amount, and it was also possible to calculate when and how much participants were *overlapping* (talking at the same time). These timestamps were further synchronized with the system log timestamps, by knowing when the audio recording of each session started (and ended). This can allow the correspondence of verbal communication activity and system events (including nonverbal communication cues).[15]

[13]Audacity. Official Website. Retrieved June 1, 2022, from `https://www.audacityteam.org`

[14]Based on a decibel (dB) threshold and minimum duration of silence between sounds. The default settings of 26 dB and 1 second were applied.

[15]An analysis of the audio activity with respect to when a participant pair shared the same spatial context is published as Supplementary Material of the article by Reski et al. (2022), available at: `https://www.frontiersin.org/articles/10.3389/frvir.2021.743445/full#supplementary-material`

6.4.6.5 Study Procedure

Based on the overall study setup (physical study space, interfaces setup, task, and measures), each study session followed the same procedure of four stages with an overall anticipated duration of approximately 120 minutes:

1. Introduction (10 min);
2. Collaboration 1 (45 min):
   (a) Warm-up 1 (5 min in the VE),
   (b) Task on PWt *fruits* dataset (30 min in the VE),
   (c) Questionnaires (10 min);
3. Preparatory Break (5 min);
4. Collaboration 2 (45 min):
   (a) Warm-up (5 min in the VE),
   (b) Task on PWt *vegetables* dataset (30 min in the VE),
   (c) Questionnaires (5 min).

In the *introduction*, the participants were first welcomed and then asked to complete an informed user consent in regard to their participation, and subsequently complete the pre-task questionnaire. The researcher in their role as moderator then provided an overview about the two interfaces and their collaborative features, as well as about the data context and the task for their upcoming joint data analysis, i.e., the first and second *task* stages.

The initial choice of which participant assumed which role and interface was random. For the two *task* stages, the participants were encouraged to analyze the data and complete their task at their own pace. However, for practical purposes, the pair was given a duration of approximately 30 minutes to aim for and to have a frame of reference in regard to their task completion progress. Whether the pair required more or less time, was up to them. For the *first task*, each participant of the pair assumed their role and respective interface. Using a special warm-up dataset, different from each of the two task scenario datasets, the participants were provided with the opportunity to warm-up and become familiar with their interfaces and the collaborative features. Once the pair felt comfortable, the researcher loaded the task scenario dataset and issued the start for the pair's task completion by initiating the audio recording. During the tasks, the pair could only talk to each other, while the researcher refrained from making any comments to the pair, only writing down noteworthy observations. Once the pair considered themselves to be done with their task by speaking aloud "*We are done with the data exploration*" (or equivalent), the researcher stopped the audio recording. The participants were then asked to complete the three *questionnaires*, i.e., in order, the SUS, the UES-SF, and finally the STCQ.

After a short *preparatory break* in which the researcher made several arrangements, including sanitizing the technical equipment and re-initializing all software components, the participants switched their assumed roles and interfaces, and the *second task* stage started by following the same procedure as the first one (warm-up, task, questionnaires). Finally, the pair was thanked for their participation and sent off. If they inquired about their task performance, they were informed after the study completion.

6.4.7 Results of Evaluation 2

6.4.7.1 Participants

A total of five collaborator pairs were recruited, resulting in a total of $n = 10$ participants. The two participants of each pair knew each other prior to the study.[16] Two pairs reported a background in *Information Visualization and Visual Analytics*. One pair reported a *Computer Science* background, and another pair one in *Applied Linguistics*. The participants of the remaining pair stated a mixed background, i.e., *Linguistics* and *Psychology* respectively. Furthermore, with respect to the verbal communication between the participants during their task completion, it is noteworthy that the study was conducted in English language which all participants were fluent in, although none of them were native English speakers. Only one participant with a background in Computer Science considered having *a lot* of prior experiences with VR technologies, while all others reported *a few*. None of the participants reported any visual perception issues in regard to the applied color coding throughout both interfaces. Figure 6.19 provides some impressions of several participants during their collaborative task completion, immersed and interacting in the VE.

6.4.7.2 Task Assessment

All collaborator pairs were able to complete the two tasks based on the PWt fruits and vegetables datasets, providing an estimation for each of the ten correlations in each task scenario, i.e., five for the sunlight–plants data variable pairs, and five for humidity–plants. Their answers are presented in Table 6.5. Overall, they were on average 84% correct and 10% incorrect by estimating that there was *no* correlation at all (but not being wrong by estimating the *opposite* correlation, which only happened for 6%). Their reported confidence for any wrong or no correlation answers was medium or low, and only the participants of the pair p4 stated a high confidence (for their two mistakes).

[16] Among other considerations and even though the participants were physically distributed in different locations as described in Section 6.4.6.1, this type of recruitment was part of the implemented safety precautions due to the COVID-19 pandemic at the time of the study.

Figure 6.19: Immersed participants during their collaborative task completion in the third VE iteration, and within the context of the hybrid asymmetric collaboration setup, wearing a HMD and interacting in the VE using 3D gestural input to perform various features (see Section 5.5.2 as part of Chapter 5).

6.4.7.3 Questionnaires

System Usability Scale and User Engagement Scale The reported system usability and user engagement scores for both interfaces, i.e., the immersive VE and the non-immersive desktop terminal, are presented in Figure 6.20. The scores for both measures are very positive, ranging between *good* and even *best imaginable* for the SUS, and having median values at or above four (out of five) for all the factors of the UES-SF. There were three instances where the scores for the two interfaces noticeably differed, i.e., (1) *focused attention* (UES-SF), (2) SUS, and (3) *perceived usability* (UES-SF).

First, the immersive interface's focused attention was rated higher than the non-immersive one, which is encouraging given its anticipated IA characteristics. However, a Wilcoxon signed rank test with continuity correction was conducted to compare the focused attention score medians for the immersive and non-immersive interfaces; $V = 27, p = .23$, there was no significant difference of medians.

Second, the non-immersive interface's usability score (SUS) was rated higher than the immersive one. A paired t-test was conducted to compare the SUS score means for the immersive ($M = 79.75, SD = 7.31$) and non-immersive ($M =$

Pair	Task	Correct Answers	Wrong Answers	No Correlation
p1	fruits	8	2	0
p1	vegetables	6	2	2
p2	fruits	9	0	1
p2	vegetables	7	0	3
p3	fruits	10	0	0
p3	vegetables	8	1	1
p4	fruits	9	0	1
p4	vegetables	9	1	0
p5	fruits	9	0	1
p5	vegetables	9	0	1
	mean	84%	6%	10%

Table 6.5: The participant pairs' answers across the two task scenarios (fruits/vegetables). Out of ten required answers for each scenario, the number of correct answers (according to the correlation model of the respective task's dataset), the number of wrong answers (positive when negative correlation was the correct answer, and vice versa), and the number of times the pair assessed that there was no correlation (there was always a correlation according to the model). Two chi-squared tests were performed to determine whether there was a difference between the answer frequencies (correct or wrong/no correlation combined) and either the task scenarios (fruits/vegetables) or the weather variable (sunlight/humidity). In both cases there was none: $X^2(1, N = 100) = 1.86, p = .17$ and $X^2(1, N = 100) = 0, p = 1$, respectively.

$91.5, SD = 8.1$) interfaces; $t(9) = -3.38, p = .008$, the means were significantly different. Upon closer examination of the received answers for the individual SUS items, it appears that the difference was due to an item on whether support from a technical person would be required to use the system.[17] Requiring technical assistance was not an issue during any of the study sessions. However, all but one of the participants declared minimal prior experiences with VR technologies. This can be an expression of lack of confidence from the participants' part.

Third, the non-immersive interface's perceived usability was rated higher than the immersive one. A Wilcoxon signed rank test with continuity correction was conducted to compare the perceived usability (UES-SF) score medians for the immersive and non-immersive interfaces; $V = 5.5, p = .027$, there was no significant difference of medians. Upon closer examination of the received answers for the individual perceived usability (UES-SF) items, it appears that this was mostly due to an item about frustration,[18] which seems understandable given

[17] SUS Item 4: "*I think that I would need the support of a technical person to be able to use this application.*" (Brooke, 1996)

[18] UES-SF Item PU-S.1: "*I felt frustrated while using this Application X.*" (O'Brien et al., 2018)

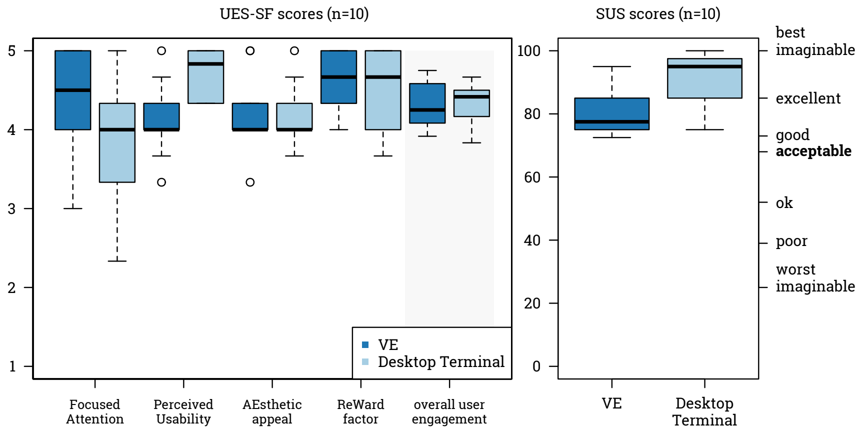


Figure 6.20: Left: Results of the UES-SF for both interfaces, presented according to the different engagement dimensions and the overall user engagement. The median for all individual factor scores (incl. overall engagement) for both interfaces is well above average. **Right**: Results of the SUS for both interfaces, presented including the original numerical scale and the supplemental adjective ratings (see Section 2.5.1). The mean value for both interfaces is well above the *acceptable* threshold.

the relatively higher complexity of the immersive interface and its 3D gestural input modality.

Spatio-Temporal Collaboration Questionnaire The reported collaboration assessments based on the STCQ from the perspectives of the immersive and non-immersive interface users are presented in Figure 6.21. The results presentation and respective discussion are based primarily on the median values and, where significant, the interquartile range.

The participants reported that there were a few individual efforts (TSIA.1), and a lot of group efforts (TSIA.2). Furthermore, the non-immersed user also had the impression that they took more of a leading role compared to their immersed partner (TSIA.3).

The participants reported constant verbal communication (NC.1), and often nonverbal communication (NC.2). The participants also reported that they were almost constantly in dialog (NC.3) and that they sometimes negotiated (NC.4). Similar to the responses as to who took more often a leading role (TSIA.3), the non-immersed users considered that they initiated these negotiations more often than their immersed partners. However, this was done equally on median (NC.5). Noteworthy for the context of the NC items, all paired medians were the same. Wilcoxon signed rank tests were conducted to determine whether the median

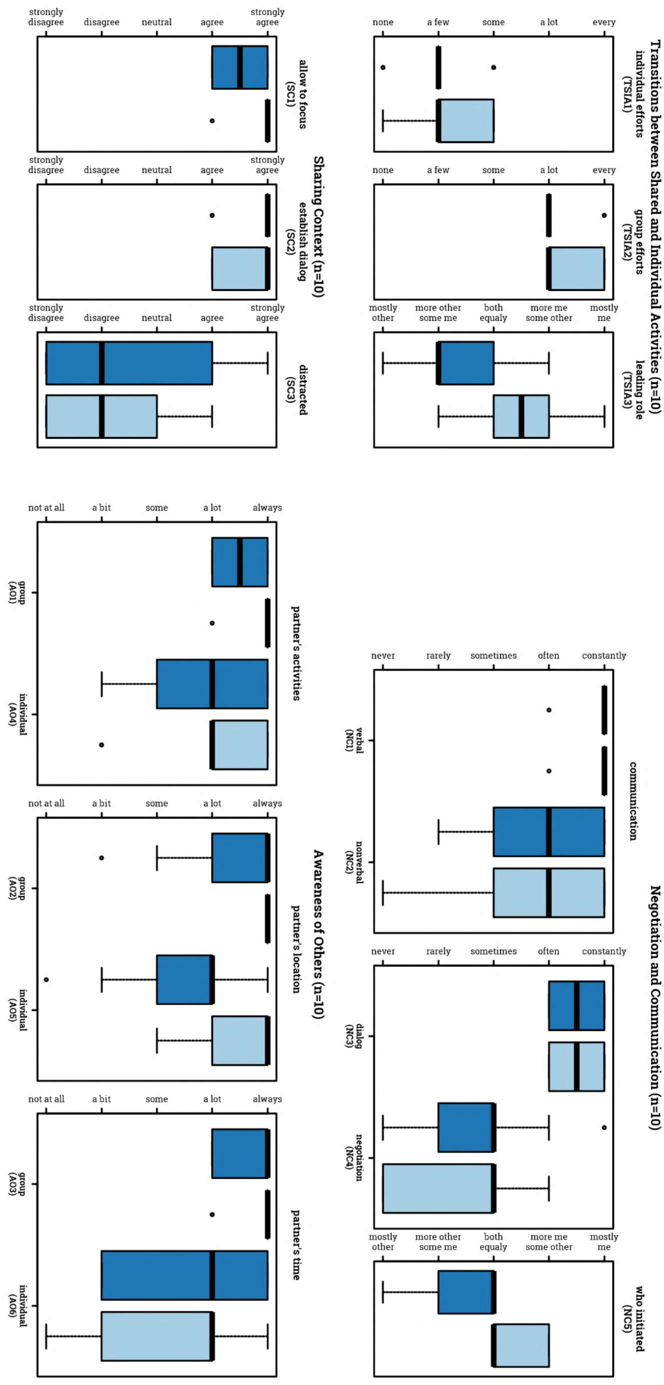


Figure 6.21: Results for the STCQ for both interfaces (VE in dark blue; Desktop Terminal in light blue), presented in groups according to the four different collaboration dimensions. **Top Left**: Transitions between Shared and Individual Activities. **Top Right**: Negotiation and Communication. **Bottom Left**: Sharing Context. **Bottom Right**: Awareness of Others.

values were the same for the immersive and non-immersive interface users, for all NC items; in all cases there was no significant difference ($p = 1$ for NC1, NC2 and NC3; $p = .42$ for NC4 and $p = .018$ for NC5).

The participants strongly agreed that the collaborative system allowed them to focus on the same subject as their partner (SC.1), and also to establish dialog (SC.2). While the participants disagreed that the collaborative features of the system distracted them from their individual efforts, the interquartile range for the immersed users was wider (SC.3).

The participants reported to have been always aware of their partner's activities during group efforts and a lot during individual efforts (AO.1 and AO.4). Their awareness of their partner during individual efforts was a little lower than during group efforts, as expected. The participants were always aware of their partners' location during both group and individual efforts, except for the immersed users during individual efforts (AO.2 and AO.5). The participants were always aware of their partner's temporal context during group efforts (AO3), and a lot during individual efforts with quite wide interquartile range for both users (AO.6).

6.4.7.4 Logging and Audio Analysis

Based on the collected log file data and the recorded audio files (see Section 6.4.6.4), the pairs' collaboration can be assessed in various aspects. Table 6.6 presents the session duration for each task scenario in minutes, the number of unique locations the participants visited (both, together, and independently), how long they were at the same location at the same time, how long the speaking time was for the immersive and non-immersive interface users, and how long their speaking time overlapped.

One noticeable outlier regarding session duration was the second session by the fourth pair. However, the short duration had no impact on their task performance (see Table 6.5) and was not due to the participants being in a hurry to complete their session. Overall, there was no significant difference of session duration means comparing the fruits ($M = 30, SD = 7$) and vegetables ($M = 26, SD = 11$) scenarios; paired t-test, $t(4) = 0.65, p = .55$. A paired t-test was conducted to compare the normalized amounts that the participants were at the same place for the fruits ($M = 88\%, SD = 7\%$) and vegetables ($M = 87\%, SD = 8\%$) scenarios; $t(4) = 0.18, p = .86$, indicating that the means were not significantly different. A paired t-test was conducted to compare the immersed and non-immersed users speaking time means; $t(9) = -1.80, p = .11$, indicating that the means were not significantly different. However, every participant that changed role from using the non-immersive to the immersive interface spoke less, and every participant that changed role from using the immersive to the non-immersive interface spoke more or about the same, as illustrated in Figure 6.22.

Similar to the log file analysis conducted as part of the previous collaborative evaluation (see Section 6.3.4.4), pathway visualizations were created to visualize the pairs' spatial data analysis contexts over time as they moved between the different locations. Figure 6.23 illustrates two representative examples, indicating their overall close collaboration throughout the data analysis activities.

6.4.7.5 Observations

Data Exploration and Task Solving Strategy Throughout all ten task sessions, the collaborators appeared to be very engaged and motivated to solve the given task as best as possible by identifying the appropriate correlations. They would come up with a hypothesis for a plant–weather correlation based on their joint observations in one location (data entity), and then move on to confirm this by investigating the same data variables in one or several other locations before confirming or rejecting their initial hypothesis. This behavior was observed in nine of the ten sessions. Only one pair (p1, fruits scenario) deducted all correlations based on their observations from a single location. Furthermore, most sessions followed a rather systematic approach, guided through the answer sheet the non-immersed user was in charge of, seemingly providing somewhat of a starting point for their investigation. However, as the collaborators were focusing on one plant–weather correlation, they were also often able to make interesting observations relevant for other correlations along the way, effectively diverting from the structure of the answer sheet and collecting their insights rather organically as their investigation proceeded. Particularly towards the end of their task session, they would together refer back to the answer sheet to identify which plant–weather correlations remained unexplored. None of the participant pairs appeared hectic, stressed, or otherwise pressured for time in their task sessions. Three of the five pairs approached the task completion noticeably objective-oriented, considering what would be the best or most effective way to solve the task using the provided interfaces. The other two pairs appeared to be more freely and openly analyzing the data and making observations. At times, the choice of what location to explore next seemed to be influenced by the collaborators' prior knowledge or relation to a specific country, providing another point of reference for their ongoing investigation.

Collaboration In six sessions, the collaborators appeared to be equally guiding and directing the task completion, going back and forth based on their respective observations. The non-immersed user seemed to be somewhat in a more leading role during the remaining four sessions, providing more directions in regard to what and where to explore next. Generally throughout all the sessions, the collaborators were able to communicate in a seemingly organic manner with each other, using various deictic and reference-related terms (see Section 2.3) to support their contextual information exchange. The implemented collaborative features in both interfaces appeared to further facilitate their natural interaction, resulting in comments such as, *"You see, the point [in time] you selected is actually*

Pair	**Task**	**Duration** (*mins*)	**Unique Locations Visited By** both	together	VE	DESKTOP	**Were At The Same Location (%)**	**Speaking Time (%)** VE	DESKTOP	overlap
p1	fruits	31.2	3	1	2	2	97.7	22.1	36.7	0.7
p1	vegetables	40.3	16	9	13	13	89.8	28.0	28.2	0.9
p2	fruits	24.4	16	9	14	11	88.8	8.1	24.8	0.0
p2	vegetables	25.4	23	12	22	13	86.7	14.3	35.9	1.2
p3	fruits	22.4	12	7	11	8	81.2	20.8	69.1	4.0
p3	vegetables	24.8	8	3	7	3	97.1	53.6	33.4	4.8
p4	fruits	40.6	38	10	22	37	87.8	29.3	37.9	1.7
p4	vegetables	9.4	16	3	8	14	85.6	27.1	28.1	7.4
p5	fruits	32.5	36	8	12	33	82.8	45.3	48.4	4.5
p5	vegetables	28.5	34	8	16	33	75.0	38.8	46.2	3.9
	mean	28.0					87.3	28.7	38.9	2.9
	standard deviation	9.1					6.9	13.9	13.0	2.4

Table 6.6: Results of the logging data and the audio analysis for all participant pairs across the two task scenarios (fruits/vegetables). **Note**: The percentage values are normalized according to the pair's task session duration.

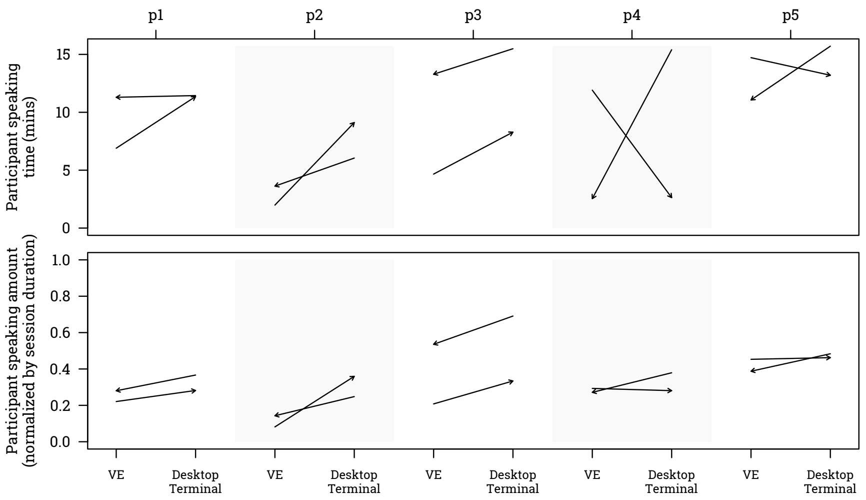


Figure 6.22: A comparison of how much each participant spoke during the collaborative data analysis activity (see Table 6.6), from the interface of their first task (either VE or Desktop Terminal) to their second task, for each pair. **Top**: Speaking time in minutes. **Bottom**: Speaking amount normalized by task duration. A pattern can be observed that the participants spoke less when immersed in the VE, compared to when using the Desktop Terminal.

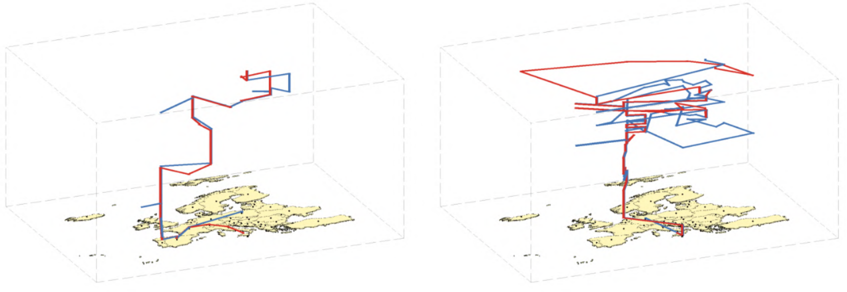


Figure 6.23: Examples of different data exploration strategies based on a participant pairs' travel and selection interactions using their respective interfaces, compiled as pathway visualizations over time (VE in red; Desktop Terminal in blue). **Left**: The participant pair collaborated closely, examining the same spatial contexts for most parts of their joint data analysis activity. **Right**: The participant pair collaborated closely in the beginning of their joint data analysis activity, while making some more individual efforts later on to find "interesting" locations to verify their task assessments. **Note**: An overview of all pathway visualizations is presented in Appendix C.

interesting for me too, because (...)." In at least four sessions, the collaborators were observed laughing at various occasions, appearing to overall enjoy themselves during their joint data analysis activity. At times, the collaborators also made inquires to one another, requesting observations about the data explored by their partner. Among others, such inquires included:

- *"Could you please highlight < x >?"*
- *"Can you do one more [highlight]?"*
- *"How does the < plant dimension > look?"*
- *"How does the period I marked now look like [for you]?"*
- *"Let's try something different: Can you see where a peak for < plant dimension > is in < location >?"*
- *"Can you check < here / this / these days >?"*
- *"Can you describe the trend for the entire time?"*
- *"Can you suggest one more location [from looking around]?"*
- *"It's so great that you can tell me where < location > is, because I am terrible at geography."*

System Features Interaction The majority of interactions of all participants in the immersive VE, wearing the HMD and utilizing the 3D gestural input, appeared natural and fluent. One minor usability issue with the implemented 3D gestural input was noticed, potentially resulting in an unintended travel interaction when instead they anticipated to display details-on-demand by touching a 3D radar chart's *Mode Toggle* widget. A reflection on such unintentional gestural commands is also presented in Section 5.5.9.2. Nevertheless, in the rare cases that this occurred, the immersed user was able to quickly recover from this, traveling back to their desired location in order to continue the investigation, while letting their partner know that they "*accidentally traveled someplace else*." In line with the prior described inquires, the non-immersed user often created temporal references in their interface, upon which the immersed user was able to describe their data variables at that point in time, or that time range respectively. In seven sessions, the immersed user performed various times what can be best described as *live annotation*, i.e., they would grab the time slice in the 3D radar chart, move it slowly in time, and describe how the time-series of a plant data variable is evolving as the position of the time slice updates – naturally, their non-immersed partner was able to follow along based on the collaborative information cues implemented in their interface (see Figure 6.15 middle). These live annotations would result in descriptions such as for example, "*It is fairly low here, now it rises, more, and more, now it is at its peak, and now it goes down again*." Similarly, at times they would also grab the time slice or make a time range selection and move it quickly back and forth to signal a specific period to their partner, while commenting on observations along the way (see Figure 6.15 middle and bottom). Furthermore, the immersed users were also able to detect patterns in the data, allowing them to make deductions accordingly. For instance, one VE user expressed "*If we find out what happens with* $< plant\ dimension\ A >$*, we also know happens with* $< plant\ dimension\ B >$*, because it is exactly the inverse.*"

Reference and Deixis Terminology Even though numerical value information of the different data variables were available in both interfaces, the collaborators largely appeared to ignore these throughout the majority of the task sessions. Instead, they used various descriptors in order to explain to each other their observations of the time-series data as presented in their respective interfaces. A selection of such descriptors, as noticed by the observer, include (in alphabetical order): *bump/bumpy, curvy, down, high, inverse/opposite, low, (local) minimum, (local) maximum, mountain, peak, period, slope, spikes, top, uniform, up, valley*. Furthermore, general deictic terms for both spatial and temporal references included: *here, from here to there, this [point in time/location], these [time range], earlier, later*.

6.4.8 Discussion

To enable collaboration in an immersive VE, special features are required that facilitate the collaborators' mutual understanding during their joint data analysis activity. For instance, the ability to make references to an artifact in the CVE allows the collaborators to synchronize their in-situ contexts and to establish common ground for the subsequent data analysis and interpretation. The use of an avatar as 3D virtual representation of a collaborator's partner in the CVE may provide important expressive information cues that facilitate mutual understanding, similar to how co-located humans interact in the real world. However, an avatar-based approach is not always feasible, for instance when the collaborators are utilizing heterogeneous display and interaction technologies. It may be the case that one of the collaborators is using an interface that is based on non-immersive technologies that simply do not feature any 3D tracking or other technologies that allow the respective implementation of their avatar for their partner. Particularly in a hybrid asymmetric collaboration setup that endorses the concept of utilizing immersive and non-immersive interfaces for collaborative analytical meaning making and data interpretation, as described in Section 6.1, alternative approaches for the implementation of collaborative information cues are needed.

Utilizing as a foundation an immersive data analysis environment that is based on the 3D radar chart data entity visualization design and respective VE composition as described in Section 5.5, various questions arise, for instance how to enable the immersed user to detect referred artifacts in the VE that originated as signals from a non-immersed collaborator, or how to reflect the data analysis context of the immersed user in the respective interface of the non-immersed partner. To investigate such matters more closely, two dedicated empirical evaluations have been conducted within the scope of this thesis. First, an exploratory empirical evaluation was conducted, investigating various visual designs as spatial and temporal data references in the immersive VE that were guided by three generalized design options. The subjective assessments as reported by various immersed participants aid the identification of user preferences and allow for respective design reflections. And second, based on a representative confirmative data analysis task, various collaborative aspects in regard to the interplay between immersed and non-immersed analyst were empirically evaluated, building on the outcome of the initial efforts in that direction as discussed in Section 6.3.5. Respective reflections on system design and collaboration provide further insights towards anticipated multimodal analysis workflows.

6.4.8.1 User Preferences of Reference Designs

Spatial References Out of the three implemented spatial reference designs, the location design was favored the most by the participants, universally across elementary and synoptic task configurations. Their perceived aesthetics and legibility ratings were higher compared to the node and pillar designs. The implementation of the location design was possible due to the composition of the VE, in particular due to the availability of the individual extruded country polygons on the virtual floor. Using the features of the VE seemed to have been appreciated by the participants, allowing an identification of either one or multiple referred data entities. At the same time, this design maintained the original visual composition of the data entity, which was very much favored by the participants under consideration of the presented CIA context. Interestingly, the user preference in favor of the implemented location design is somewhat in contrast to the results reported by Lacoche et al. (2017), where their *safe navigation floor* approach rated worse compared to the others. The actual design was visually similar in both cases, but the signal's intent differed: While the purpose of the presented location design was to actively guide the user to the highlighted area, theirs was instead to avoid it (Lacoche et al., 2017).

Comparing the node and pillar designs, it is interesting to see the disparity in regard to their assessments across the elementary and synoptic task configurations. The pillar design was rated better for referring to one data entity, likely because it allowed the participants to quickly and easily identify the referred data entity (legibility), while it took them subjectively a bit longer to identify the node design. However, the synoptic design of using one large pillar to highlight a group of data entities did not scale appropriately, as it was not clear to them what data entities are referred to exactly. Maybe such a design would be better suited if one was to make a reference to an otherwise not specified large spatial area, instead of a group of identifiable data entities in the VE. These results are quite interesting, as they indicate a favor for different designs based on the task configuration. Arguably from a user interface perspective, one would normally strive to apply a uniform design strategy, implementing the same designs for the same, or similar, tasks. Within the context of the presented scenario, the question comes to mind, whether or not one should implement the same design across both tasks, thus following a rather coherent approach, or instead implement different designs, each in better support for their given task. Clearly, this requires careful consideration from the application's interface designer, weighing pros and cons for either design given a purpose and task at hand. Arguably, in the presented data analysis context, legibility and user preferences independent of the design are more important, allowing analysts to be as precise as possible in their referencing and subsequent collaboration. Consequently, if environmental features were not available for the implementation of the location design, instead a mixture of pillar for elementary referencing and node for synoptic referencing should be applied. The pillar design is conceptually similar to the *light beam* technique as presented by Peter et al. (2018), who reported mixed results from their evaluation, describing that sometimes the participants would consider the light beam being part of the VE instead of being a dedicated signal. The realistic design of the light beam seemed to have conflicted with the realistic setting of their VE, which prevented participants from clearly identifying the signal through the *VR-Guide* (Peter et al., 2018). On the other hand, the pillar design, as initially experimentally applied within the first collaborative data analysis setup, worked reasonably well for the elementary spatial referencing of data entities that were represented as stacked cuboids (see Section 6.3.2). Compared to the realistic VE composition as presented by Peter et al. (2018), the immersive data analysis environment featured a rather abstract visual composition, in that case arguably making it easier for the immersed user to clearly distinguish the pillar design from the other artifacts in the VE.

Temporal References There was no distinct temporal reference design that was favored across elementary and synoptic tasks within the 3D radar chart scenario. For time event (elementary) referencing, the participants generally preferred the highlight design. Interestingly, while it scored better aesthetics ratings, its median for legibility is the worst compared to the three other designs. The pointer design is a close runner-up in regard to user preference, scoring generally better legibility ratings. It seems that the participants liked the analogy of literally "*pointing to a point in time*" – even though a rather abstract design was used instead of a realistic one (Sugiura et al., 2018).

Similar to the results presented by Peter et al. (2018), some trends were identified towards the participants' favor of the outline design compared to all others for the synoptic task configuration, both across all data dimensions and for an individual one. This is particularly interesting and also somewhat odd, as the outline design received the lowest user preference for the elementary task, even though it scored comparatively good in regard to aesthetics and legibility ratings. These results reveal a similar preference disparity as for the spatial node and pillar designs, requiring careful consideration in favor for or against a visually uniform interface design.

Two designs based on the AA option were implemented using different indicator types, namely pointer and symbol. The symbol design was rated slightly better when directly compared to the pointer one. However, the results indicate trends towards a rather equal preference of both designs when examining the bigger picture. A participant's comment in regard to using the pointer design during synchronous collaboration, while encoding additional meaning in the symbol indicator during asynchronous collaboration, in a more annotation-like manner, was particularly interesting, encouraging some further investigation. No major advantages or disadvantages for one over the other were identified, making both potentially valid designs depending on reference task and purpose.

Based on the received assessments and within the presented scenario, one can argue as follows. The outline design appears to be a favored reference design for making time range references (synoptic), especially under consideration of the comparatively good aesthetics and legibility ratings. However, an implementation as temporal reference for an elementary task configuration of an individual data variable was not possible using the outline design (neither for the highlight one). Given the preference for the pointer design over the symbol one for the elementary task, one can see its application respectively, thus again recommending a mixture of different design approaches across elementary and synoptic tasks for making temporal references.

6.4.8.2 Collaborative Confirmative Analysis

Usability All reported usability scores for both interfaces are above the good margin, indicating that the collaborators were generally able to operate their respective interfaces for their intended purpose. Given the overall purposefully minimalistic but representative design of the non-immersive desktop terminal as an interactive InfoVis, the comparatively high usability scores (median above excellent) are not that surprising, as it relied on rather established visualization techniques, i.e., line graphs and a map from a bird's-eye view (Ward et al., 2015, Chapters 6 and 7; Munzner, 2014, Chapter 12; Lundblad et al., 2010).

The received positive usability feedback for the immersive VE is overall in alignment with the results of the prior empirical evaluation, dedicated to the investigation of the implemented 3D gestural interface design (see Section 5.5.6). In fact, comparing the usability scores (see Figures 5.35 and 6.20), the feedback received within the scope of the collaborative scenario was better compared to the conducted single user experiment. It is noteworthy to acknowledge some underlying differences in this comparison though, i.e., (1) the collaborative scenario allowed the users to interact in the VE to their own accord, while the single user experiment featured a more rigid structure and task protocol, and (2) the zoom (in/out) feature was not available in the collaborative scenario, which could be improved based on the feedback received in the single user experiment. Nevertheless, the slightly better usability feedback for the VE in the collaborative scenario is encouraging, as it arguably presented a more realistic scenario where the immersed user utilized the various provided features freely to their own desire to explore and analyze the data.

Generally, all participants were able to quickly pick up and learn the various aspects of the immersive interface during their brief warm-up stage. They understood the concept of the 3D radar chart data entity visualization design and the various collaborative information cues, became comfortable with wearing the HMD, and utilized the 3D gestural input to interact in the VE. This is in line with the general anticipation of utilizing immersive technologies for their natural interaction techniques (Skarbez et al., 2019; Büschel et al., 2018). After all, enabling users to simply pick up the technology and start using it for data analysis purposes in an intuitive manner without extensive training allows them to focus on the subject matter at hand.

The self-reported usability scores coincide also with the observations, confirming the VE user's ability to operate the immersive interface in a natural and fluent manner. In fact, the majority of the participants made reflections at the very end of the study, i.e., after the completion of their second task, positively highlighting the "*smoothness*" of their experience in the VE and that they overall could have easily spent even more time with their joint data analysis activity. Considering these comments in regard to the measured session duration ($M = 28, SD = 9.1$) and within the presented confirmative analysis scenario and task, this can arguably be considered a step towards the direction of moving beyond comparatively brief "just a few minutes"-VR experiences. This is also important keeping in mind the complexity inherent of collaboration in general (see Section 2.3), i.e., interpretation of data, information exchange, as well as discussion and negotiation take time (Heer and Agrawala, 2008; Andriessen, 2003). Following this line of thought and with respect to such multi-user interplay, one can potentially anticipate longer exposure times in CVEs within CIA scenarios compared to single user experiences.

Among others, the above good usability scores for both interfaces are important within the scope of the conducted empirical evaluation for two reasons in particular. First, it validates that the interfaces could be operated as intended and without any major usability flaws. Consequently, a potential negative impact on the pairs' overall collaboration due to an "unusable" interface can be excluded. And second, it also indirectly validates the usability of the designed and integrated collaborative features as part of each respective interface. The visual information cues (see Section 6.4.3) were easy to recognize and provided important contextual spatial and temporal references about their respective partner's activity, as also emphasized by Cruz et al. (2015). Additionally, these references were integrated in a rather seamless manner, allowing for the automatic transmission of information as the users naturally interacted with their interfaces. As opposed to introducing additional dedicated actions in regard to what information to share and when, such as discussed as part of the initial collaborative evaluation in Section 6.3.5 or as used by Welsford-Ackroyd et al. (2020) and Peter et al. (2018), this seemed to have allowed the collaborators to naturally interact with each other, seamlessly picking up and referring to their partner's context without noticeable action delays, i.e., without the need to wait for a specific collaborative signal.

User Engagement The median user engagement scores for both interfaces are at or above 4, indicating an overall high user engagement with the provided collaborative system during the confirmative analysis task. This corroborates the observer's impressions of the collaborators being motivated and eager to use their interfaces to analyze the data and to find the correct answers, noticeably enjoying themselves and their collaboration during their task. These results align well with the often stated argument that immersive technologies have the potential to provide engaging experiences that encourage data interaction and interpretation (Ens et al., 2021; Dwyer et al., 2018; Hackathorn and Margolis, 2016).

Furthermore, the results allow for a discussion in regard to the individual user engagement factors. The immersive interface users reported a higher focused attention compared to the users who operated the non-immersive desktop terminal. There are arguably a couple of potential reasons for this. Primarily, the characteristics of the applied display and interaction technologies have to be taken into account. Perceiving a VE through a HMD and allowing for natural hand interaction, i.e., a comparatively high level of immersion, may have required the immersed user to be generally somewhat more attentive, as there are many visual stimuli to process – both in regard to the data visualization in the VE itself as well as due to the integrated visual information cues triggered through their non-immersed partner. Additionally, it also needs to be considered that while the non-immersed user explored two data variables (weather data) per location, the immersed user was presented with five data variables (plant data) per location (see Section 6.4.6.3). Nevertheless, the general high focused attention across both interfaces can be attributed to the overall close collaboration between the users, for instance the pairs investigating the same respective locations for the majority of the task duration (see *were at the same location* column in Table 6.6), collaboratively interpreting and discussing the time-series data of a specific data entity. In regard to the slightly lower focused attention score reported by the non-immersive interface users, another aspect comes to mind, i.e., the note taking and completion of the pen-and-paper task answer sheet (see Section 6.4.6.3). During the task, they were in charge of keeping track and filling out the provided plant–weather answer matrix, which arguably may have affected their attention to the interface as they were required to temporarily switch their focus to the physical task answer sheet.

The reported perceived usability scores for the two interfaces were in line with the reported SUS scores, as discussed in the prior thematic paragraph.

Aspects in regard to the aesthetic appeal were rated similarly positive across both interfaces as well. At this stage of the collaborative system, the focus in the immersive and the non-immersive interface was on the essential parts that allow the users to explore and analyze data, trying to avoid unnecessary information or distracting elements in general. The received aesthetic appeal scores are satisfactory, overall indicating that the participants enjoyed the chosen graphical elements and visual design for each of the interfaces accordingly.

The positive reward scores, reported with medians above 4.5 for both interfaces, are especially interesting and encouraging. The collaborators were observed being particularly motivated to solve the given task correctly, often verifying their observations of the time-series data across multiple different locations to ensure that their answer was correct. A reoccurring expression across the different task sessions was along the lines of "*I am sure we got it [right], but let's just check one more [location]*." The investigative nature of the confirmative data analysis task provided the pairs with a clear purpose for this activity, which is also important in regard to moving beyond initial novelty reactions (Snowdon et al., 2001). At the same time, it was completely up to them to organize their task solving approach, resulting in interesting data analysis strategies. The freedom of task approach, combined with the fact that they had to work together, using different types of display and interaction technologies, but still being able to have a notion of what their partner was up to, supported through the collaborative features integrated in the interfaces, likely contributed to these positive reward scores. The participants appeared genuinely excited that "*it [the collaborative system] really worked*" and "*we [the collaborators] are able to see each other*," thus successfully enabling them to be mutually aware of each other (Heer and Agrawala, 2008; Benford et al., 1994; Benford and Fahlén, 1993). The positive reward scores for both interfaces within the presented collaborative context indicate arguably also a comparatively equal user contribution. Of course each interface served their own purpose, but they were engaging for both collaborators alike, motivating them to partake in and equally contribute to the data analysis activity – as anticipated (see Section 6.1). It is also important to consider two more factors within the context of the reward scores, i.e., the overall high rated usability, and the absence of an explicit time limitation. The users' ability to use the interfaces as intended in a scenario where they were not pressured for time, informally confirmed by some of the participants' expressions that they could have spent even more time with the developed system, likely also contributed beneficently to these reward scores.

System Design Reflections Overall, both the immersive and the non-immersive interface were assessed positively in regard to usability and user engagement. It is important to highlight again that the reported scores are not meant to be compared in a "X is better than Y" manner, nor can they (due to the asymmetric role setup), as the interfaces serve different purposes. Instead, with the primary objective to investigate collaborative aspects when integrating and combining immersive and non-immersive interfaces into the same data analysis workflow, it is arguably crucial to have an understanding and assessment of the applied tools that are likely to impact the collaboration. After all, the empirical evaluation of collaborative immersive systems is complex, as highlighted throughout Section 2.5 and by Ens et al. (2021), Skarbez et al. (2019), and Billinghurst et al. (2018). Having received a similar assessment by the participants, one can assume that the two interfaces are appropriately balanced in regard to their purpose and operability within the scope of the presented context as well as with respect to their anticipated hybrid asymmetric collaboration. This assumption serves as a good foundation for the assessment of various collaborative aspects, particularly when using heterogeneous device types and an asymmetric user role setup.

System Logs, Audio Analysis, Observations, and Task Assessment The analysis of the system logs as presented in Table 6.6 reveals that the collaborators spent the majority of their time investigating the time-series data at the respective same location, i.e., the same spatial data entity (min 75.0%, max 97.7%). This also becomes apparent when examining the audio analysis and pathway visualizations as illustrated in Section 6.4.7.4. Consequently, one can infer that the collaborators found themselves in a state of rather *close collaboration* for the majority of the session duration, i.e., directly interacting with each other in the same spatial data context, making efforts as a *group* to solve the task by investigating the respective time-series data.

They communicated about their observations and findings both verbally, i.e., by talking to each other to explain, discuss, and negotiate, and nonverbally, i.e., by making spatial and temporal references to point and highlight data for their partner using the provided collaborative features. Collaboration relies inherently on complex personal and social processes (Billinghurst et al., 2018; Heer and Agrawala, 2008), therefore every human user has slightly different ways and approaches of interacting with one another. This is well reflected in the results of the audio analysis in regard to the speaking time during the task sessions (see Table 6.6 and Figure 6.22). For instance, some pairs (pair p2, fruits scenario, 24.4 min: 8.1% and 24.8%) communicated verbally less compared to others with much higher speaking time rates (pair p3, fruits task, 22.4 min: 20.8% and 69.1%). There were also some instances when the collaborators made some more *individual* efforts. These usually occurred when a pair set out to find a new data entity (location) to explore, either in regard to yet completely unexplored plant–weather correlations, or in order to verify and confirm previously made deductions (see Figure 6.23 right).

At times, both collaborators explored parts of the data independently in order to find a location that contained interesting data, or in the words of the participants, time-series data visualizations that are "*curvy or bumpy*" and feature "*peaks, spikes, slopes, or valleys*." An interesting case is pair p4 within the vegetables task, whose verbal activity was much lower during these phases compared to the ones when they analyzed the same location. However, during similar individual spatial exploration phases of other pairs, their overall verbal activity did not seem to change that much compared to the remainder of their joint data analysis, indicating that they generally kept talking to each other independent of whether they were making individual or group efforts.

In general, all pairs were able to collaboratively complete the tasks of identifying the various data variable correlations (ten in total per task and subsequent dataset) in a satisfactory manner (see Section 6.4.7.2). Given the hybrid asymmetric collaboration setup, there was no knowledge carryover anticipated from the fruits to the vegetables task in regard to the respective interface's operation. A carryover of data insights was also not possible as both tasks featured different datasets. There is the possibility for the pair's task solving approach in the vegetables task to be somewhat influenced and informed by their applied strategy in the prior fruits task. However, there was no significant impact on their task performance (see Section 6.4.7.2). Given these circumstances as well as the prior described influences of the personal and social processes on collaboration (Billinghurst et al., 2018; Heer and Agrawala, 2008), it is unlikely that there was a noticeable knowledge carryover between the two tasks.

STCQ: Transitions between Shared and Individual Activities In addition to the discussion of the pairs' collaboration based on the previously described system logs, audio analysis, observations, and task assessment, the participants also provided self-reported assessments as collected through the administered STCQ (see Figure 6.21). The pairs' own reporting in regard to the occurrence of *individual* and *group* efforts is in line with the observations and system log analysis. They considered having made a lot of shared group efforts, while only a few individual efforts during the task solving activity. Furthermore, the pairs had the impression that the non-immersed collaborator was in somewhat of a leading or directing role compared to the immersed one. While the VE users reported a median of *more other, some me*, the median for the non-immersed users lies between *both equally* and *more me, some other*. Based on the researcher observations, it is likely that the task answer sheet responsibility caused the non-immersed user to assume the role of being the task director at times. Even though assuming such a "leading" role just temporarily, it arguably cannot be considered to be the same as dedicated guiding roles, for instance as discussed by Welsford-Ackroyd et al. (2020) or Peter et al. (2018), but instead an overall rather balanced interplay between the collaborators for their own purposes that is conceptually similar to the scenarios described by Lee et al. (2020), Sugiura et al. (2018), and Gugenheimer et al. (2017).

STCQ: Negotiation and Communication In regard to the pairs' verbal communication frequency, both interface users reported that they talked pretty much *constantly*, in line with the observer's subjective impressions. The pairs stated that they *often* utilized the nonverbal communication features, i.e., the provided collaborative synchronous features. Furthermore, the pairs considered dialog taking place for the majority of their verbal communication. Negotiation was reported taking place only *sometimes* if at all, rather similarly initiated by both interface users. All the above is interesting for a couple of reasons. First, the medians from both interface users across these five items are equal, overall indicating that the collaborators had a rather similar impression about their negotiation and communication independent of the interface type. Second, considering the higher share of dialog compared to the lower amount of negotiations, it seems that the collaborators were rather successful in their verbal and nonverbal communication, being overall able to follow their joint data descriptions and interpretations, without much need for additional negotiation. And third, the reported amounts of verbal and nonverbal communication, most of the time categorized as dialog, further indicate a close collaboration between the two interface users. It was also interesting to observe pairs establishing their individual reference terminologies, including common and reoccurring expressions as well as more unique ones.

STCQ: Sharing Context The results indicate that the implemented collaborative features allowed the users to focus on the same subject as their partner, and to establish a dialog accordingly, which is a foundational aspect for successful collaboration (Cruz et al., 2015; Heer and Agrawala, 2008; Snowdon et al., 2001). Overall, the collaborators *disagreed* that these features distracted them from their individual efforts, however with a slightly bigger range of provided impressions. Generally, all the results in this category are favorable within the context of the presented hybrid asymmetric collaboration setup and task. The ability to focus on the same subject and to establish dialog are crucial for any kind of collaboration (Dix, 1994). With both interface users confirming that they were able to do so, the overall design of the collaborative features across both the immersive and the non-immersive interface can be considered validated within the presented context, of course assuming that a verbal communication channel is available (see Figure 6.8). These results are also relevant within the context of physically distributed collaboration environments (Skarbez et al., 2019), as it enables analysts to work together remotely, independent of their distance to each other.

Furthermore, while the collaborators did not assess the implemented collaboration features as distractions during their individual efforts, it is important to consider the amount of reported individual efforts that took place, i.e., only a few. On the one hand, the collaborators assessment is a promising trend in regard to the provided features design, allowing them to focus ad hoc on their partner's context without interfering with their own individual efforts. On the other, some further investigations using tasks that involve more individual efforts throughout the collaborative task are necessary to confirm or reject this trend.

Finally, it was also interesting to observe that some pairs came up independently from each other with similar ways of utilizing the implemented collaborative features, such as the live annotation interaction behavior.

STCQ: Awareness of Others Both interface users reported high awareness about their respective partner's activity in general, their location in space (spatial data context), and their time reference (temporal data context). These assessments allow again some reflections and discussion on the designed collaborative features that aimed to facilitate their awareness of one another.

First, the assessments for joint awareness were reported slightly higher during group efforts, which is desired as this is arguably the situation when it is more important to know about the collaborator's activity. Nevertheless, the awareness was still rated fairly high even during the few individual efforts, and seemingly in a non-distracting manner as discussed before.

And second, the awareness of their partner was slightly higher perceived through the non-immersed user, with everyone agreeing that they were *always* aware of the immersed user's activities in the VE. Reflecting on the characteristics of the implemented visual information cues across both interfaces, one aspect becomes apparent in regard to seemingly different update frequencies. The location and time reference updates from the immersive interface appear much more continuous in the non-immersive interface, i.e., the position and orientation of the immersed user are constantly updating (following the positive insights as obtained in the initial collaboration evaluation; see Section 6.3.5), and even new time event and time range selections appear much more in motion and fluent due to the characteristics of the designed time event and time range selection features as part of the interaction with a 3D radar chart (see Section 5.5.2). These arguably provide smooth visual transitions from one state to another, naturally updating the non-immersive interface accordingly. The other way around, collaborative information cues from the non-immersed user update more discrete and "event"-like in the immersive VE, i.e., new selections appear when they are done (through respective pointer and keyboard events; see Section 6.4.2), providing comparatively fewer visual transition cues. This may be a good starting point for further investigations into this matter.

Still, based on the implemented collaborative information cues across both interfaces, it appears that each user was able to follow and have an understanding of their partner's current investigation, closely coupled with the reported sharing context results. As described in Section 2.3, considering the general importance of mutual awareness for the design of collaborative systems (Cruz et al., 2015; Snowdon et al., 2001; Benford et al., 1994), not least as an important foundation to establish communication that results in the subsequent interpretation and discussion of the data (Andriessen, 2003), the received awareness assessments can be interpreted positively. The designed visual approaches for supporting spatial and temporal references worked well and as intended in both the immersive and the non-immersive interface, allowing the collaborator pairs to point and highlight data to their partner accordingly. The results also align well with the insights reported by Nguyen and Duval (2014), stating that comparatively simple awareness cues can often be sufficient to provide the collaborator with an understanding of the shared workspace.

Collaboration Reflections Throughout all task sessions, the collaborators were able to work closely as a group for the majority of their joint data analysis activity in order to solve the given confirmative analysis task in a satisfying manner (see Table 6.5). Considering that they had to provide ten answers (including ten complementary confidence indications), they were overall quite busy during an average half-hour for data exploration, observing, interpreting, and discussing their findings. They frequently and extensively communicated verbally through dialog that was further facilitated through the various spatial and temporal referencing features of their interfaces. The setup allowed them to closely analyze and interpret the spatial and temporal contexts of the data, making important observations and deductions along the way. One pair in particular highlighted the "*detective work*"-like nature of the task and their collaboration, reflecting on the great interplay between the two interfaces, and rating the experience in a very positive manner. The participants' overall excitement and the natural way of interacting with each other were reoccurring themes throughout the different task sessions, likely positively contributing to their self-reported collaboration assessments. This can be further underlined through a selection of noteworthy participant comments after their task completion:

- "*Oh, this was really fun and worked really well.*"
- "*Oh wow, did we take that long? It was so much fun.*"
- "*It was a lot of fun actually.*"
- "*This was so cool.*"
- "*It worked really well.*"
- "*I was really able to see you!*"

Comments such as particularly the last one are quite interesting given that there were no avatar-based representations of the users in neither of the interfaces, for instance as utilized by Nguyen et al. (2019), Heidicker et al. (2017), or Benford et al. (1994), but just the provided visual references. It appears that the participants made themselves the mental association between the visual references and their partner. Similar observations within the context of remote collaboration around interactive tabletop systems, also including rather abstract and minimal visual representations for the collaborator's input, were made by Kim et al. (2010), reporting that users in their study "*(...) felt as though the remote participants were in the room itself.*" It would be interesting in the future to investigate effects on collaboration and empathy when there is no virtual user avatar, but instead other more abstract means of user representation, identifying requirements and use cases where one approach is potentially preferable over another. For instance, a CIA system presented by Nguyen et al. (2019) used a virtual avatar representation for a collaborator in a VE that utilized a similar VR approach. How could an alternative approach without such an avatar look like, and what would the difference be in the (perceived) collaboration?

The expressed appreciation for the collaborative system and the rewarding experience through the participants is much in line with the visions for such hybrid analysis environments, combining different types of technologies, as motivated throughout Section 6.1. All in all, using different types of display and interaction technologies and facilitated through various collaborative features, all pairs within the presented empirical evaluation were able to successfully collaborate with each other in a rather balanced shared workspace manner, as opposed to more common remote expert scenarios in similar technological setups (Ens et al., 2019).

Chapter 7

Towards Design Guidelines for Collaborative Immersive Analytics

Contents

The goal of this thesis was to thoroughly investigate the application of immersive display and interaction technologies, in particular utilizing head-mounted display (HMD) devices and interaction through 3D gestural input, to design and develop data analysis tools that are based on a Virtual Reality (VR) approach. Within the overall context of Immersive Analytics (IA) and a focus on spatio-temporal data, the two core themes explored throughout this thesis have been *interaction* and *collaboration* from a user-centered perspective. Chapter 4 set the stage by presenting a conceptual and technological system architecture that served as the foundation for the subsequent design and development of various data analysis interfaces and their infrastructure. While Chapter 5 investigated the immersive interaction with spatio-temporal data, Chapter 6 explored possibilities to enable cross-platform collaboration using immersive and non-immersive interfaces. In particular, three major Virtual Environment (VE) iterations were developed, two of which were even extended through support for synchronous collaborative data analysis with a non-immersed partner. Conducted at various stages throughout the thesis, a total of six dedicated evaluations as well as one case study, with several hands-on experiments and demonstrations, were used to obtain empirical insights for all the developed data analysis environments. The results of these empirical efforts have been discussed and reflected upon, also under consideration of important foundational concepts and relevant related work, as presented in Chapters 2 and 3.

As a synthesis of the presented interdisciplinary research efforts and the obtained empirical insights, various *design guidelines* can be derived. These design guidelines aim to contribute to the research community by providing helpful directives for the design and development of data analysis interfaces that utilize

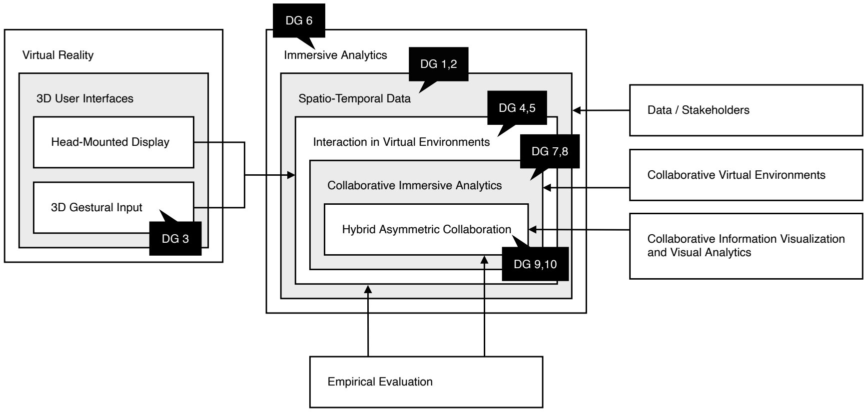


Figure 7.1: The ten derived design guidelines (DG 1-10) are aligned and mapped to the respective thematic modules of the presented interdisciplinary thesis design space (see Section 2.6).

immersive display and interaction technologies – within the contexts of IA and Collaborative Immersive Analytics (CIA) in general, and the presented thesis design space in particular (see Section 2.6).

A total of ten design guidelines have been derived. Figure 7.1 illustrates their alignment and mapping to the thesis design space. Section 7.1 begins with the presentation of the first six guidelines, centered around the immersive interaction with spatio-temporal data. The remaining four guidelines, focused on various aspects of collaboration, are described in Section 7.2.

7.1 Interaction Design Guidelines

Design Guideline 1

Consider providing supporting artifacts that facilitate orientation and interpretation of the spatial data context.

Under consideration of the three-dimensional (3D) VE, the visual representation of an individual data item, i.e., the data entity, has to be placed "somewhere" to allow subsequent interpretation and interaction. For instance, with a focus on spatio-temporal data, data entities in the VE may be placed in accordance to geolocation coordinates. Naturally, other position mapping approaches may be applied based on dataset and scenario. Independent of the applied mapping, it can be beneficial to not just visualize data entities in an *empty* VE space, but additionally provide artifacts that support the user's contextual spatial understanding. These artifacts can provide a frame of reference that facilitates the interpretation of the visualized data, enabling the user to establish an association between data entity and the supporting artifact, and thus assist with the data entity's identification in the spatial context. Additionally, supporting artifacts can provide important information cues that facilitate the immersed user's ability to orient themself in the VE, and in turn their ability to navigate and travel in the 3D space – to identify not yet visited areas as well as to find their way back to previously visited ones. This is arguably of particular importance when the user explores the data in a more overview-like manner, i.e., looking and moving around in the VE with the objective to identify "interesting" data entities that are worth of further investigation. Ideally, the implemented supporting artifacts should ensure unambiguous identification. To provide an example, assuming data entities are positioned based on their geolocation coordinates, the VE could feature additional visualizations of relevant geographical areas, rendered as extruded surfaces on the virtual floor. With individual data entities placed directly on top of these surfaces, one could easily associate their spatial context, even before displaying further details-on-demand. The user's awareness of their own in-situ spatial context is not just important for their own data analysis, but also relevant during collaboration, allowing the establishment of a shared vocabulary for spatial referencing.

Design Guideline 2

Consider the visual mapping for the integration of the temporal data variables into each data entity.

When creating the design for a data entity, i.e., a data item's visual representation in the VE, carefully consider the integration of temporal data variables through appropriate visual mapping in accordance to the dataset and the scenario. A good starting point is the identification of the amount of temporal data variables for each data item as well as their value range. Consider whether to visually map a single time event onto a data entity, or a larger time range that consists of multiple time events. Encoding a time series of consecutive time events in the visual representation of a data item should enable the user to identify trends over time, allowing them to easily spot "interesting" events that are worth of further inspection. When encoding a time series, consider the amount of visualized time events. Maybe there are instances in the data analysis activity when it is sufficient to display just a limited time range as a subset of the entire time series, allowing the user to focus on a specific period.

Design Guideline 3

Design for hand interaction.

3D gestural input, commonly referred to as hand interaction, can provide intuitive mechanisms for the interaction with data in the immersive 3D space. Within the context of immersive data analysis, easy to understand interactions can facilitate data discovery through active user engagement, encouraging the exploration of the VE's content. For the design and implementation of interactive features as data analysis tasks, consider the utilization of established 3D interaction techniques. Hand-based grasping can allow the immersed user to interact with *visible* graphical artifacts in the VE, while gestural commands enable the performance of *invisible* operations. Furthermore, consider the use of real-world analogies for the design of the 3D gestural interface. For instance, looking and pointing at a specific data entity can be easily associated with the mental thought of "*I want to go there*", and used as a respective travel feature in the VE.

Design Guideline 4

Design with hand posture complexity in mind; utilize simple unimanual techniques for frequent tasks, and more complex bimanual techniques for less frequent ones.

Similar to physical comfort considerations that should be kept in mind when designing a 3D gestural interface, consider also the overall complexity of hand postures, for what data analysis task they will be formed, and estimate their frequency accordingly. Consider the utilization of comparatively simple unimanual (one-handed) techniques for tasks and interactions that are commonly expected. For less frequent tasks, consider the implementation of more complex bimanual (two-handed) techniques. Furthermore, consider also other properties inherent of a data analysis task within the overall context of the anticipated interaction in the VE, for instance error recoverability. Take into account the implications of performing comparatively "drastic" tasks, such as resetting the visual configuration of a data entity, and consider guarding such interactions from unintentional performance by assigning them to complex and unique hand posture configurations that are presumably not formed by accident.

Design Guideline 5

Limit available interactions based on the user's in-situ context.

The design of interactive features for immersive 3D VEs is a complex endeavor in general. Based on the VE's purpose, it is likely that different types of interactions need to be provided. This is of particular relevance within the context of immersive data analysis environments that require support for various analysis tasks, and thus demand a rich set of interactions to facilitate the data analysis activities in the VE. To assist the design of a growing interactive feature set that is made available to the user, their in-situ contextual interaction should be taken into account. More specifically, consider whether or not all features need to be at the user's disposal at all times during the data analysis activity. Some features may only be relevant under certain circumstances, for instance during details-on-demand investigation of a data entity or with an applied selection. Based on the user's in-situ interaction context, consider temporarily enabling or disabling some features. Consider the implementation of such a feature logic, aiming to assist the system's ability to interpret the user's input, their intention, and anticipated interaction. Among others, this may assist with the prevention of unintentional commands, particularly in immersive environments that utilize a 3D gestural interface with a rich feature set where some hand postures and gestures are similar – in their design, how users form them, and based on the system's input interpretation.

Design Guideline 6

Consider workflow integration with non-immersive tools.

While the user is immersed in the VE, i.e., wearing the HMD and interacting in their safe interaction area using the 3D gestural interface, their ability to multitask and to switch ad hoc their mode of operation is arguably limited. During more traditional data analysis activities that utilize non-immersive display and interaction technologies, the user is typically able to easily switch between different contexts, for instance taking notes on a physical medium or even iterating between multiple applications on their device. This enables them to be versatile with their overall data analysis workflow, utilizing and quickly switching between different tools. Due to the characteristics of immersive display and interaction technologies, particularly with respect to HMD devices, such an alternation between different tools is rather cumbersome. Consequently, the role of the immersive VE within the context of the overarching data analysis workflow should be taken into account. Consider the mechanisms and tools that usually assist the user *outside* the VE, and how to potentially integrate them as part of the activity *inside* the VE. Consider how to "export" insights and discoveries obtained from the immersive analysis for later revisit and potential use as input parameters to other data analysis tools.

7.2 Collaboration Design Guidelines

Design Guideline 7

Facilitate collaboration by enabling multimodal communication using a mixture of verbal and nonverbal tools.

During synchronous collaboration with other users in the context of a joint data analysis activity, verbal communication can be an essential tool that allows for the discussion of data discoveries, the sharing of knowledge and perspectives based on the collaborators' expertise, and the coordination of their group efforts. Thus, consider providing the support for verbal communication independent of a user's interface type and their physical location. In addition to their ability to talk with each other, nonverbal tools can also provide important information cues that facilitate the collaboration. Consider the implementation of nonverbal information cues, particularly visual ones, directly integrated as part of the visualization, as they can facilitate aspects of data guidance and identification by supporting the user to detect respective visual differences. For instance, important aspects of the visualized data in the interface may be temporarily highlighted, thus standing out from other parts in the immersive visualization, aiming to catch the user's attention. In practice, rather than using verbal and nonverbal communication cues in isolation, they are often applied in a multimodal manner. The ability to make references is typically accomplished through a nonverbal pointing action and accompanied by a verbal expression. This allows collaborators to focus on the same data artifact, and in turn enables them to establish a common ground during their joint data analysis activity. Consequently, consider providing both verbal and nonverbal tools in order to support collaboration through the design and implementation of powerful multimodal features.

Design Guideline 8

Consider the design of the nonverbal collaborative information cues; modify or add artifacts to a data entity, or modify its environment.

The purpose of nonverbal collaborative information cues is typically to catch the user's attention in order to guide them to an artifact in the virtual space that is of interest. Several aspects should be taken into consideration when designing such nonverbal information cues. The overall composition of the VE is of relevance, providing opportunities or limitations for the design of nonverbal information cues. Does the VE feature rather abstract or more realistic graphical artifacts? Is the VE rich of graphical features or rather minimalistic? The determination of such properties is a good starting point for the subsequent design of nonverbal information cues. Similarly, this also applies to the overall data entity visualization design, i.e., the composition of the visual representation of a data item in the VE. Its complexity may allow or limit the possibilities for the design and implementation of subsequent nonverbal information cues. With respect to a data entity, consider whether to modify its visual appearance, add additional artifacts to it, or modify environmental features in the VE that can be easily associated with the data entity respectively. Consider also the overall characteristics of nonverbal collaborative information cues as well as the implications from the interaction in the immersive VE. Collaborative information cues are typically just temporary. While the immersed user can explore the entire virtual 3D space (360° field of regard), it is not guaranteed that the performed information cue appears in their field of view – it might as well appear behind them. Hence, consider designing information cues to be distinct and recognizable, even if the immersed user does not observe them when they first appear.

Design Guideline 9

Consider the update frequency of the nonverbal collaborative information cues; utilize continuous updates to allow for fluent collaboration, and on demand updates for focused ad hoc group efforts.

When collaborating synchronously with other users across different types of interfaces, consider how often collaborative state updates from one interface to the other are transmitted. Start by considering what information is valuable to the respective collaborator for sharing and display in their interface. Then, consider when and when not to share the information with the collaborator. Finally, consider how and how often the information about a user's state should be shared with the collaborator. Anticipated collaboration styles can be a useful indicator to determine the update frequency of such nonverbal collaborative information cues. For instance, consider continuous updates (automatically and frequently) to allow a more fluent interplay between collaborators during close collaboration. Similarly, consider the utilization of less frequent on demand updates for focused ad hoc group efforts among otherwise mostly individual efforts (when the information about a collaborator's state are not always necessary). Consider the benefits and drawbacks of the applied update frequency of the nonverbal collaborative information cues within the context of the anticipated individual and group efforts – one does not want to distract a user during their more independent work.

Design Guideline 10

Consider the classification of the collaborative data analysis experience; take into account data context, scenario, tasks, technologies, and user roles.

When setting out to design and implement a collaborative data analysis experience, the various aspects of the anticipated collaboration should be unambiguously classified by utilizing established terminologies, frameworks, instruments, questionnaires, and so forth. Consider the context as well as the type of data. Carefully consider also the data analysis scenario and as such the purpose of the collaboration to begin with. Think critically about the overall data analysis workflow. What tools and interfaces are desired, what types of tasks do they aim to support, and how do they contribute to the collaborative data analysis? Consider the display and interaction technologies utilized across potentially multiple different data analysis interfaces, their implications on the collaborative activity, and how to support a smooth interplay between them. Furthermore, do not neglect the roles of the involved users. What is their knowledge and expertise, and how are they expected to partake in and contribute to the joint data analysis. With many different aspects, purposes, objectives, and requirements to consider, the classification can facilitate a critical design process.

Chapter 8

Conclusions and Future Work

Contents

The conduction of various empirical research efforts allowed for the investigation of *interaction* and *collaboration* within the overall context of Immersive Analytics (IA). In response to the thesis' research goal (see Section 1.2), the obtained insights and experiences allowed for discussion and respective reflections. Ultimately, these led to the synthesis of ten design guidelines, as presented throughout Chapter 7, aiming to provide useful considerations and to facilitate the development of immersive data analysis environments. This chapter begins with Section 8.1, providing a conclusive summary of the research presented throughout this thesis by revisiting the three defined research objectives that allowed to reach the posed research goal. Some impulses and directions for future work are presented in Section 8.2, effectively concluding this thesis.

8.1 Conclusions

This thesis set out to investigate interaction and collaboration within the context of IA, an applied and interdisciplinary research area that aims to utilize immersive human-computer interfaces to create engaging virtual spaces that facilitate data exploration, analytical reasoning, and collaborative decision making. To narrow down the scope of this investigation, a focus was set on data analysis environments that implement a Virtual Reality (VR) approach through head-mounted display (HMD) devices and 3D gestural input as means for intuitive interaction in three-dimensional (3D) Virtual Environments (VEs). Next to the focus on the described technological approach, the data analysis scope was set to focus on spatio-temporal data, a data type that is relevant for the measurement and observation of various real-world phenomena. To facilitate the overall research goal of contributing to the research community with empirical insights and the proposal of design guidelines within the presented context, three research objectives were defined and addressed accordingly throughout the thesis.

> **Research Objective 1**
>
> Design and implementation of a system that allows for multivariate data analysis using immersive display and interaction technologies.

Modular System Architecture Immersive data analysis environments can be complex systems that involve various different components, both conceptually as well as technologically. This becomes even more apparent the moment that collaborative components are introduced, noticeably increasing the system's complexity, and thus requiring for conceptual considerations and technological infrastructure. To address this matter, and to create a foundation for the subsequent implementation of various immersive data analysis tools, a general system architecture has been described throughout Chapter 4. The definition of several essential requirements did not only facilitate the system's design overall, but also indirectly assists with the formal description of its various components and capabilities. The architecture is purposefully designed to be modular and approachable, allowing for adaptation and respective extension through other researchers and practitioners by outlining major buildings blocks, providing impulses and considerations for the design of similar systems in the future. Within the scope of this thesis, the presented system architecture was applied to implement three major VE iterations (see Chapter 5), including two collaborative system setups that focused on combining immersive and non-immersive interfaces (see Chapter 6). The successful implementation of these data analysis environments, including the presentation of several complementary use cases that demonstrated the developed VE iterations in alternative data contexts, allowed for subsequent empirical evaluations with human participants. In turn, this indirectly confirms the validity of the described system architecture within the presented context.

> **Research Objective 2**
>
> Investigation of 3D UI design approaches to support immersive interaction with spatio-temporal data.

Virtual Environment Iterations and Task Terminology To investigate the interaction with spatio-temporal data in immersive VEs with a focus on 3D gestural input, i.e., hand interaction, three major VE iterations were developed and presented in Chapter 5. At their core, these VEs were centered around different data entity visualization designs, i.e., the visual representation of data item in the virtual space. Different visual mapping strategies were applied in order to encode respective spatio-temporal data dimensions in these data entities. Next to the design, development, and empirical evaluation of the three VE iterations, Chapter 5 also contributes with the presentation of the adopted terminology for data analysis tasks and their adaptation towards the context of IA and the interaction with spatio-temporal data in VEs.

Interaction Using Spheres The initial VE utilized a *Sphere* design (spatial geolocation encoding, no temporal encoding), and implemented a first set of exploratory interaction features that were mapped onto three different input modalities, i.e., gamepad, 3D gestural input, and physical, tracked controller. The experiences and insights obtained from a comparative evaluation of these input modalities validated not just the technological feasibility for such an immersive data exploration environment, but also the users' ability to make use of it and solve related tasks. Furthermore, the evaluation also confirmed the suitability of hand interaction within the presented context.

Interaction Using Stacked Cuboids The second VE extended the initial concept by implementing a *Stacked Cuboid* design (spatial geolocation encoding, temporal time event encoding), and was investigated within the context of a real-world dataset related to multilingualism on social media within the Nordic region. The designed 3D gestural interface focused on providing features that allow for intuitive interaction across both the spatial and the temporal data context, applying a mixture of hand-based grasping, gestural command, and graphical menu interaction techniques for essential analysis tasks. The obtained empirical feedback based on a case study with linguists, who could hands-on operate the developed VE within the scope of various lab experiments and public demonstrations, provided valuable impulses and considerations for the further development of such immersive data analysis environments – not just in regard to engaging interaction in VEs, but also with respect to reflections on real-world data analysis workflows in practice. Besides highlighting the importance of system capabilities that allow for note taking and insights export for later use in other data analysis tools, the experiences with the linguists also encouraged the further exploration of collaborative aspects (as covered within the scope of the third research objective).

Interaction Using 3D Radar Charts To further advance the immersive interaction with spatio-temporal data, a third and final VE iteration was developed, centered around a *3D Radar Chart* design (spatial geolocation encoding, temporal time series encoding). The empirical efforts within the scope of this VE iteration include (1) the validation of the overall data entity visualization design, (2) the exploration of a proof-of-concept prototype to allow virtual note taking, and (3) the investigation of an extended feature set to support a variety of data analysis tasks using 3D gestural input. The obtained insights allow for critical reflections with respect to hand-based grasping, gestural commands, and unintentional commands, providing considerations for the design and implementation of future 3D gestural interfaces for immersive data analysis tools.

Research Objective 3

Extension of the immersive data analysis system to support collaboration using heterogeneous interfaces and user roles.

Hybrid Asymmetric Collaboration The investigation of Collaborative Immersive Analytics (CIA) within the scope of this thesis was driven by the motivation of bridging immersive and non-immersive data analysis interfaces. For this purpose, various aspects of cross-platform collaboration have been critically examined, resulting in the proposal of *Hybrid Asymmetric Collaboration* – a concept that aims to distinctly differentiate between heterogeneous technologies and user roles in collaborative data analysis contexts. This concept has been empirically investigated in several evaluations throughout Chapter 6, thematically aligned with the second and third VE iterations.

Collaboration Using Stacked Cuboids The collaborative setup within the scope of the second VE iteration and the sociolinguistic data context was used to obtain first practical experiences, investigating means that would allow the two collaborators, one immersed *inside* the VE and one *outside* using a non-immersive desktop interface, to be aware of each other, to establish a common ground, and to make visual references to synchronously explore the spatial data context together. For this purpose, a first set of nonverbal collaborative information cues was implemented across their interfaces, enabling the subsequent transfer of user signals. Based on an evaluation with pairs of linguistics students, who conducted an explorative analysis task, empirical feedback was obtained, among others, with respect to (1) their ability make joint data interpretations and assessments, (2) their data exploration strategies, (3) the deictic terminology they adopted, as well as (4) the collaborative system's usability.

Collaboration Using 3D Radar Charts and Spatio-Temporal Collaboration Questionnaire The insights and constructive feedback obtained from the first evaluation encouraged further iteration to improve on the collaboration design within the scope of the third VE iteration. The collaborative setup and confirmative analysis task required pairs of users to analyze a spatio-temporal dataset not just in regard to the spatial context, but also the temporal one. To further explore the design of nonverbal information cues that can be utilized as visual references in a Collaborative Virtual Environment (CVE), several spatial and temporal reference designs were evaluated. Their visual appearance was informed through the definition of three generalized design options that aim to facilitate the conceptual design of similar collaborative information cues on a general level. The evaluation results of the implemented visual references, comparing their aesthetics and legibility as well as general user preference, indicate that different reference designs may be preferable based on their task. The actual evaluation of the collaborative interplay between an immersed and a non-immersed user was, by comparison to the first collaborative investigation in the prior VE iteration, more thorough, additionally exploring aspects of user engagement as well as utilizing the self-constructed *Spatio-Temporal Collaboration Questionnaire* (STCQ). The questionnaire in particular allowed the participant pairs to self-assess their own collaboration in a more structured and formal manner. Generally, the obtained empirical feedback indicates that the iterated collaborative features across their interfaces enabled the participant pairs to closely and organically collaborate, allowing them to equally contribute to their joint data analysis activity, and resulting in an engaged and rewarding experience for both collaborators. The results of this evaluation, in combination with the discussion and reflections, are encouraging for the design and development of future data analysis environments that utilize heterogeneous display and interaction technologies and distinct user roles to allow synchronous cross-platform collaboration.

8.2 Future Work

Based on the research presented throughout this thesis, several directions for future work are conceivable.

One starting point could be the further investigation of the presented design space (see Section 2.6). The design space reflects on the interdisciplinary characteristics of the presented empirical work through its modular composition. Thus, either similar investigations could be conducted by adopting the design space as is, or through adaptation of individual modules that would create an alternative design space. For instance, while maintaining the overall theme of IA and interaction with spatio-temporal data, an immersive approach other than VR could be applied to investigate similar data interaction aspects using different types of immersive display and interaction technologies – thus adapting the respective left part of the design space accordingly. With similar rapid advances in regard to Augmented Reality technologies in recent years, an arguably obvious alternative would be to explore the interaction with spatio-temporal data using such an approach instead of VR. Due to the capabilities of modern cross-platform engines, developed artifacts may even be transferred into a new visualization and interaction modality, such as illustrated in Figure 8.1, presenting a running prototype of the implemented 3D radar chart visualization in augmented reality. Similarly, alternative collaboration setups with respect to both technology and user role are conceivable, exploring other dimensions in alignment with Milgram

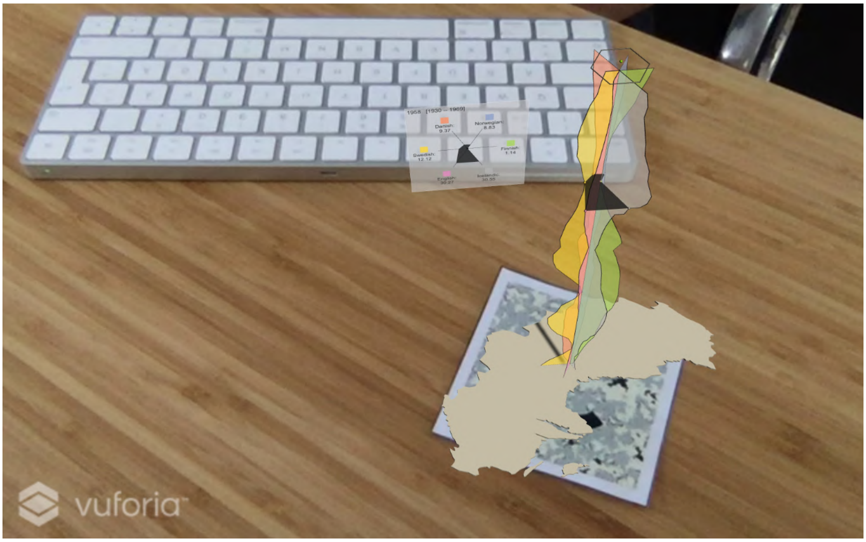


Figure 8.1: An impression of a prototype, illustrating the developed 3D Radar Chart design (see Section 5.5) under utilization of an Augmented Reality approach instead of a VR one. The prototype utilizes a marker-based tracking approach to align the visualization in 3D and in real-time with a physical marker that is placed in the real-world environment. Some of the interactive features were even mapped to a keyboard as input modality, for instance to enable interaction with the 3D radar chart's time slice to select new time events.

et al.'s (1995) Reality-Virtuality Continuum, for instance as recently surveyed by Fröhler et al. (2022).

The thesis' focus on spatio-temporal data was chosen due to the relevance of measuring and observing various real-world phenomena across many different contexts. Naturally, besides the data contexts presented throughout this thesis (elections, social media, forestry, weather and climate), many other datasets and analysis scenarios exist that invite to be explored through immersive display and interaction technologies.

The implemented logging system as part of the overall system architecture was useful within the scope of the conducted empirical evaluations, among others, for the retrospective visualization of a user's spatial data exploration over time (see Appendix C). Rather than using a non-immersive medium for the visualization of a user's immersive data analysis, it could be intriguing to display such information directly in the immersive VE itself, for instance similar as described by Büschel et al. (2021). This could arguably also be useful for other types of collaboration, such as asynchronous collaboration, where an analyst traces back the prior interactions of a peer to make similar observations and interpretations.

Throughout the various empirical evaluations presented in this thesis, the anticipated duration that a user would spent immersed in the VE was between 15 to 35 minutes, which naturally varied from participant to participant. However, under assumption of being immersed in the VE for a longer duration, it would then arguably also make sense to explicitly investigate aspects about the 3D gestural interface's comfort and physical fatigue. Furthermore, even though the participants throughout the evaluations were generally able to quickly learn the operation of the interface through hand interaction, usually within a 5 minutes warm-up, it would also be intriguing to investigate more specifically aspects of learnability – a topic that is according to Rempel et al. (2014) often disregarded and underexplored.

Some opportunities to combine abstract data visualization with graphically more realistic virtual environments have been discussed as part of the presented urban climate data use case (see Section 5.5.8), further exploring the concept of immersive situated analytics (Bressa et al., 2022; Thomas et al., 2018). Similar to the previously suggested retrospective visualization of user interaction data in immersive environments, a similar conceptual approach could be applied within the context of situated visualizations, aiming to address not just explorative and confirmative data analysis tasks, but also facilitate the presentation of analysis results, i.e., the analyst's third task category (see Section 5.2) – a topic that was not further explored within the scope of this thesis. An immersive presentation of data analysis observations and interpretations, similar to a "guided tour", could align well with concepts such as immersive visual data stories (Isenberg et al., 2018), aiming to integrate narrative components in the results presentation. In combination with intuitive immersive interaction, these could potentially facilitate the approachable dissemination of data observations, even to data novices.

As part of the third VE iteration, a proof-of-concept feature to allow virtual note taking was explored. Note taking is a fundamental part of analysis workflows (Willett and Isenberg, 2015), but unfortunately another topic that is not yet very thoroughly explored within the context of IA (Fonnet and Prié, 2021). Under the umbrella of the presented *User Session Data Transfer* building block of the system architecture and the implemented virtual note taking (see Sections 4.2.3 and 5.5.3.2), this is another opportunity that is worth of further investigation in the future. Potentially, such note taking could be useful not just for the analyst themself who is creating the note, but also for other collaborators when utilized as virtual annotation directly in the immersive space.

The assessment of collaborative work within the context of CIA has, among others, also been deemed a major topic with respect to the current IA research agenda (Ens et al., 2021). The construction and application of the STCQ (see Section 6.2) facilitated the evaluation of a user pair's collaboration as subjective self-assessments. As its individual items are held purposefully generic, it would be intriguing to re-use and potentially further adapt the questionnaire for similar collaboration evaluations in the future.

Naturally, the presented empirical results, discussion, and reflections should be interpreted within the thesis' overall motivation, scope, goal, objectives, and design space in mind (see Chapter 1 and Section 2.6). Further investigations based on the described and documented empirical evaluations may provide additional useful insights and impulses for the interaction and collaboration around spatio-temporal data in immersive VEs.

Appendix A

Online Media

Videos Video demonstrations were recorded to present the various developed interfaces and their functionalities at various stages throughout this thesis. The following collection provides brief descriptions and links to some of these videos in thematically chronological order.

- **VE Iteration 1 – System Demonstration**
  Demonstration of the implemented Sphere design, enabling a user to explore open data in an immersive VE (see Section 5.3). Note that the system even included an early collaboration prototype (that at that stage is not further discussed within the scope of this thesis).
  `https://vimeo.com/230053828`
- **VE Iteration 1 – Input Technology Comparison**
  Demonstration of the implemented Sphere design as three pre-recorded instruction videos, each dedicated to demonstrate one input modality, and shown to the participants as part of the respective evaluation (see Section 5.3.3).
  Gamepad: `https://vrxar.lnu.se/odxvr/gamepad.mp4`
  3D Gestural Input: `https://vrxar.lnu.se/odxvr/vbmc.mp4`
  Physical, Tracked Controller: `https://vrxar.lnu.se/odxvr/rsvr.mp4`
- **VE Iteration 2 – Work-In-Progress Prototype**
  Demonstration of the implemented Stacked Cuboid design at an early stage, implementing the physical, tracked controller input modality, and utilized to demonstrate the system to linguists in May 2018 (see Section 5.4.3).
  `https://vimeo.com/270379265`
- **VE Iteration 2 – Linguistics**
  Demonstration of the implemented Stacked Cuboid design within the overall linguistics case study (see Section 5.4.3).
  `https://varieng.helsinki.fi/series/volumes/20/alissandrakis_et_al/2019-VARIENG_odxvrxnts-video.mp4`
- **VE Iteration 2 – Swedish Election**
  Demonstration of the implemented proof-of-concept prototype that illustrates the Stacked Cuboid design in the Swedish Election use case (see Section 5.4.4).
  `https://vimeo.com/317958351`

- **Virtual Environment Demo Reel – Spring 2017 to Spring 2019**
  A demo reel presenting various Virtual Environment prototypes that were developed throughout the time from spring 2017 to spring 2019. This was a complementary video shown at the thesis author's progression seminar on September 18, 2019.
  `https://vimeo.com/361004428`
- **VE Iteration 3 – Visualization Design Validation**
  Demonstration of the implemented 3D Radar Chart design in the initial iteration as used to validate its overall visualization design concept and first set of interactive features (see Section 5.5.3).
  `https://vimeo.com/393378221`
- **VE Iteration 3 – Uniform 3D Gestural Interface Design**
  Demonstration of the implemented 3D Radar Chart design in the second iteration as used to evaluate its 3D gestural interface and extended data analysis feature set (see Section 5.5.5).
  `https://vrxar.lnu.se/tdrc/tdrc-v2.mp4`
- **Collaboration in VE Iteration 2 – Collaborative Explorative Analysis**
  Demonstration of the implemented collaborative system, within the context of exploring multilingualism in tweets and hashtags on Twitter, and shown to the participants as part of evaluating their collaborative explorative analysis (see Section 6.3.3).
  `https://vimeo.com/451482987`
- **Collaboration in VE Iteration 3 – Collaborative Confirmative Analysis**
  Demonstration of the implemented collaborative system, within the context of exploring the Plant-Weather timelines dataset, illustrating all system features that were available to the participant pairs for the evaluation of their collaborative confirmative analysis (see Section 6.4.6).
  `https://vimeo.com/623459537`
- **3D Radar Chart Using Augmented Reality**
  Demonstration of an implemented proof-of-concept prototype that illustrates the 3D Radar Chart design using an Augmented Reality approach instead of a Virtual Reality one (see Section 8.2).
  `https://vrxar.lnu.se/tdrc/ar-demo.mp4`

360° Interactive, Annotated Web Viewer For additional illustration of the developed immersive interfaces using a non-immersive 2D display medium, in-engine 360° screenshots were taken and then displayed using a panoramic web viewer. This allows viewers to get a representative impression of the interface from the immersed user's field of view, even without a head-mounted display. Some of these views feature annotations that can be hovered with the pointer (mouse) to display additional information. The following collection provides brief descriptions and links to some of these panoramic views in thematically chronological order.

- **VE Iteration 2 – Linguistics**
  Demonstration of the implemented Stacked Cuboid design within the overall linguistics case study (see Section 5.4.3).
  `https://vrxar.lnu.se/apps/2019-VARIENG-odxvrxnts-360/`
- **VE Iteration 3 – Visualization Design Validation**
  Demonstration of the implemented 3D Radar Chart design in the initial iteration as used to validate its overall visualization design concept and first set of interactive features (see Section 5.5.3).
  `https://vrxar.lnu.se/apps/2020-nordichi-3drc/`
- **VE Iteration 3 – Uniform 3D Gestural Interface Design**
  Demonstration of the implemented 3D Radar Chart design in the second iteration as used to evaluate its 3D gestural interface and extended data analysis feature set (see Section 5.5.5).
  `https://vrxar.lnu.se/apps/radartimeui_v2-360/`
- **Collaboration in VE Iteration 2 – Collaborative Explorative Analysis**
  Demonstration of the implemented collaborative system, within the context of exploring multilingualism in tweets and hashtags on Twitter, and as used by the immersed participant for the evaluation of their collaborative explorative analysis (see Section 6.3.3).
  `https://vrxar.lnu.se/apps/2020-nordichi-hcia/`
- **Collaboration in VE Iteration 3 – User Preferences of Reference Designs**
  Demonstration of all the implemented spatio-temporal reference designs as collaborative information cues, within the context of the 3D Radar Chart design, and shown in the Virtual Environment to the participants as part of evaluating the user preferences of these reference designs (see Section 6.4.4). Note that the reference design descriptors slightly differ from their presentation in the thesis, for instance using sAe to indicate spatial reference variant A (pillar) as elementary task configuration, and so on.
  `https://vrxar.lnu.se/apps/tdrc-ref-360/`

Interactive Pathway Visualization Web Viewer The implemented logging system as part of the developed immersive data analysis environments allowed, among others, for the retrospective visualization of a user's spatial data exploration over time. Static impressions of these pathway visualizations are presented in Appendix C. Additionally, the pathway visualizations of the two collaborative evaluations presented in Chapter 6 can also be viewed online, interactive, in 3D.

- **Collaboration in VE Iteration 2 – Collaborative Explorative Analysis**
  The pathway visualizations of the linguist pairs, illustrating their spatial data exploration as part of evaluating their collaborative explorative analysis (see Section 6.3.3).
  `https://vrxar.lnu.se/apps/2020-nordichi-hcia/pathwayvis/`
- **Collaboration in VE Iteration 3 – Collaborative Confirmative Analysis**
  The pathway visualizations of the participant pairs, illustrating their spatial data exploration as part of evaluating their collaborative confirmative analysis (see Section 6.4.6).
  `https://vrxar.lnu.se/apps/2021-frivr/`

Appendix B

Software Modules

Open Source Statement Various software modules have been developed over time to practically facilitate the implementation of the interfaces and systems presented throughout this thesis. Some of these modules, those deemed useful and relevant, are freely available on the Internet as open source for other researchers, practitioners, students, and so on, to use and adapt. The following collection provides brief descriptions and links to these modules.

- **Unity – PolyExtruder**
  This module provides the functionality to create custom meshes (polygons) in Unity based on a collection of vertices directly at runtime. These 2D meshes are created along the x- and z-dimensions in the 3D space. The meshes can be extruded into 3D prisms along the y-dimension in the 3D space. Among others, this module was utilized to render the various geographic features on the floor in the developed Virtual Environments, for instance the Nordic countries (see Section 5.4.1), the municipalities of Sweden (see Section 5.4.4), the European countries (see Section 5.5.1), the counties of Sweden (see Section 5.5.7), and the districts of the city of Norrköping (see Section 5.5.8).
  `https://github.com/nicoversity/unity_polyextruder`
- **Unity – rworldmap import**
  This module illustrates a workflow of exporting vector data that represent the Earth's countries from the R package *rworldmap*,[1] generate respective C# classes that can be used in Unity, and finally visualize the exported vector data in Unity as extruded polygons (using the prior listed *Unity – PolyExtruder* module). This workflow has been utilized to visualize the Nordic (see Section 5.4.1) and European countries (see Section 5.5.1) in the developed Virtual Environments.
  `https://github.com/nicoversity/unity_rworldmap`
- **Unity – Log2CSV**
  This module provides the functionality to easily integrate a simple logging system in a Unity application, and has been used as part of the immersive Virtual Environment's architecture (see Section 4.2.2) throughout all iterations

[1] Andy South. rworldmap: Mapping Global Data. Retrieved June 1, 2022, from `https://cran.r-project.org/web/packages/rworldmap/`

presented in this thesis.
`https://github.com/nicoversity/unity_log2csv`

- **Unity – 3D Radar Chart**
  This module provides the initial implementation of the 3D Radar Chart design (see Section 5.5.3) as a data-agnostic and interactive data visualization for utilization in Unity.
  `https://github.com/nicoversity/unity_3dradarchart`
- **Unity – MediaUpload**
  This module provides a simple workflow that illustrates (1) the capture of media data within a Unity application, and (2) the upload of the media data as binary data via HTTP to a Node.js server. The module was used to implement the proof-of-concept virtual note taking feature as part of the third Virtual Environment iteration (see Section 5.5.3.2).
  `https://github.com/nicoversity/unity_mediaupload`
- **Unity – Connect via WebSocket server to JavaScript client**
  This module provides a simple and minimalistic template to illustrate communication and data transfer between a Unity client application and a JavaScript web client application via WebSocket connection implemented as a Node.js server. The module was used to implement the Collaboration Infrastructure (see Section 4.2.4) as part of the third Virtual Environment iteration (see Section 6.4).
  `https://github.com/nicoversity/unity_wss_js`

Appendix C

Pathway Visualizations

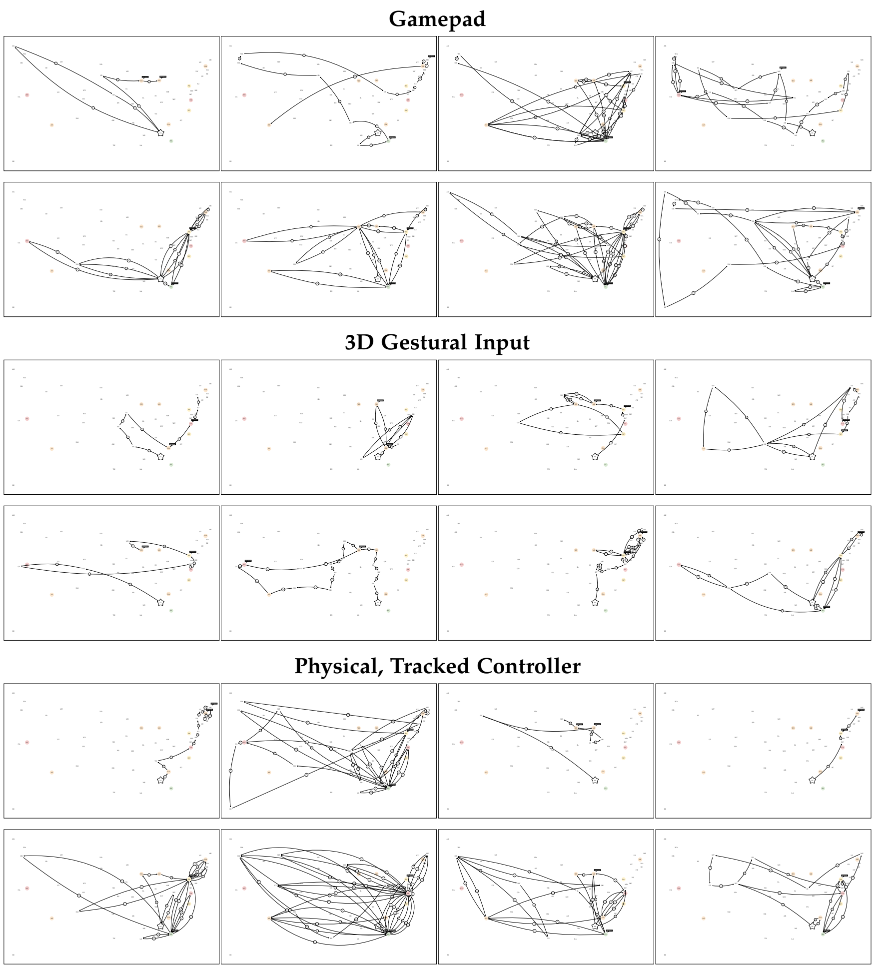


Figure C.1: Pathway visualizations, representing a user's spatial data exploration over time, based on the data collected within the comparative input technology evaluation (see Section 5.3.4.4).

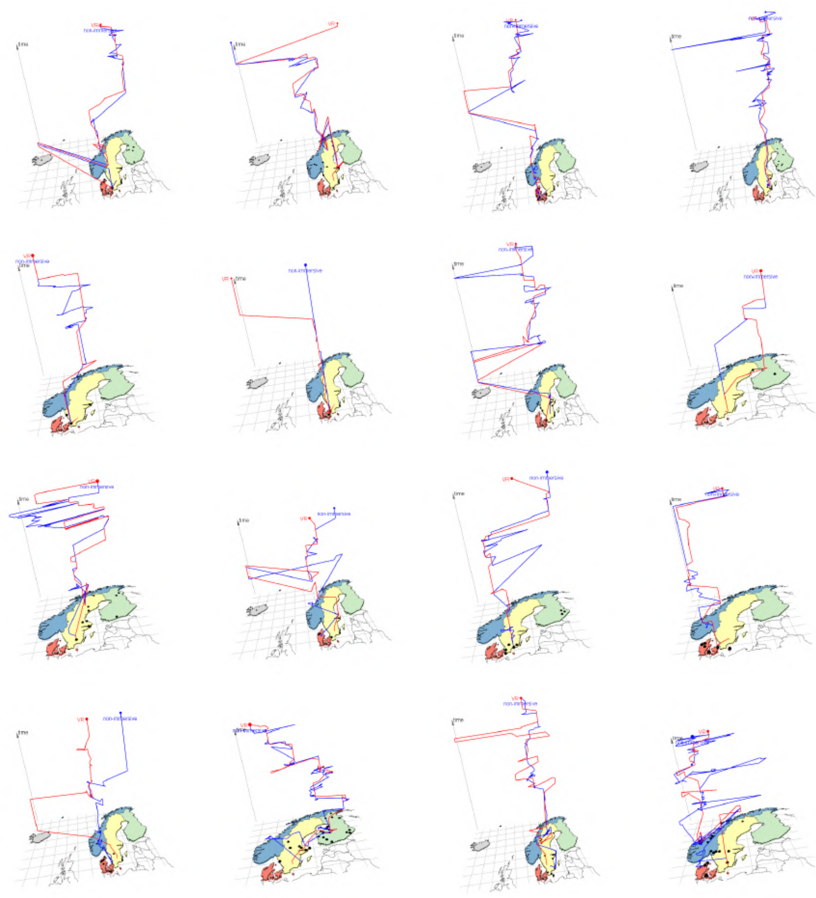


Figure C.2: Pathway visualizations, representing a participant pair's spatial data exploration over time, based on the data collected within the collaborative explorative analysis evaluation (see Section 6.3.4.4). *Visualizations created by Aris Alissandrakis.* **Note**: A link to an interactive 3D online viewer of these pathway visualizations is listed in Appendix A.

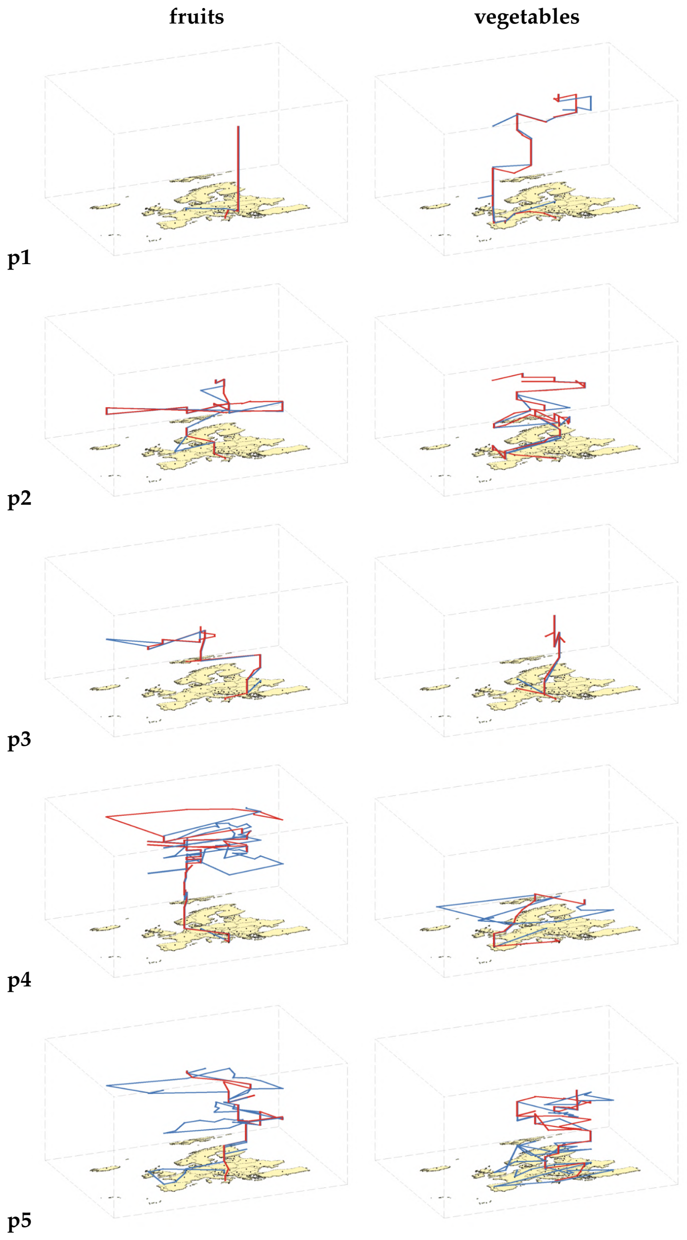


Figure C.3: Pathway visualizations, representing a participant pair's spatial data exploration over time, based on the data collected within the collaborative confirmative analysis evaluation (see Section 6.4.7.4). **Note**: A link to an interactive 3D online viewer of these pathway visualizations is listed in Appendix A.

Appendix D

Study Material

Collaboration in VE Iteration 3 – Collaborative Confirmative Analysis With respect to the conducted evaluation of investigating a participant pair's collaboration during a confirmative analysis task within the overall context of *Hybrid Asymmetric Collaboration* (see Section 6.4.6), the following collection provides some supplementary material that was utilized in the evaluation.

- **Spatio-Temporal Collaboration Questionnaire (STCQ)**
  An online repository that contains overall information about the self-constructed STCQ (see Section 6.2), including instructions on how to administer it, a generalized PDF template for re-use as is, as well as the respective LaTeX source files to facilitate remix and further adaptation.
  `https://github.com/nicoversity/stcq`
- **correlated-timelines** by *Aris Alissandrakis*
  An online repository that contains the Plant-Weather timelines (PWt) dataset as utilized throughout Sections 5.5 and 6.4, including a detailed description on how the dataset was generated.
  `https://github.com/arisalissandrakis/correlated-timelines`
- Figures D.1 and D.2 exhibit the prepared Spatio-Temporal Collaboration Questionnaire (STCQ) that was administered to each collaborator.
- Figure D.3 exhibits the "science fictional" task scenario that was presented to the participant pairs within the scope of the evaluation (see Section 6.4.6.3).
- Figure D.4 exhibits the prepared physical task answer sheet for the *fruits* scenario (see Section 6.4.6.3).
- Figure D.4 exhibits the prepared physical task answer sheet for the *vegetables* scenario (see Section 6.4.6.3).
- Individual visualizations that present a pair's verbal activity (based on the conducted audio analysis) in combination with their shared spatial context (based on the logging data) are included in the supplementary material of the publication by Reski et al. (2022), available at: `https://www.frontiersin.org/articles/10.3389/frvir.2021.743445/full#supplementary-material`.

Synchronous Asymmetric Interaction within the Context of Collaborative Immersive Analytics

Questionnaire: Collaboration

Instructions: For each of the following dimensions **[TSIA, NC, SC, AO]**, read carefully its definition, and for the questions / statements, mark *one* box that best describes your reactions to the tested application today.

Application
- ☐ Virtual Reality Application.
- ☐ Desktop Application.

Session

Date/Time: ______________________

Task: ☐ Fruits ☐ Vegetables

[TSIA] Transitions between Shared and Individual Activities: The interplay between individual and group efforts, including the ability to switch between these, within the scope of collaborative work.

TSIA.1	How many of your efforts during this task would you consider to have been *individual* efforts?	none	a few	some	a lot	every
TSIA.2	How many of your efforts during this task would you consider to have been *group* efforts?	none	a few	some	a lot	every
TSIA.3	According to your impression, who was more in a leading / directing role during the *group* efforts?	mostly other	more other, some me	both equally	more me, some other	mostly me

---

[NC] Negotiation and Communication: Verbal conversation (i.e., talk) facilitated through the ability of utilizing nonverbal information cues in order to discuss and interpret any task-related aspects of the activity (e.g., findings in the data, roles and structure of task approach, and so on).

NC.1	According to your impression, how often did you communicate *verbally* to your partner?	never	rarely	sometimes	often	constantly
NC.2	According to your impression, how often did you communicate *nonverbally* to your partner?	never	rarely	sometimes	often	constantly
NC.3	How often would you consider did *dialog* take place?	never	rarely	sometimes	often	constantly
NC.4	How often would you consider did *negotiation* take place?	never	rarely	sometimes	often	constantly
NC.5	Who would you say mostly initiated the *negotiations*?	mostly other	more other, some me	both equally	more me, some other	mostly me

---

Please continue on the next page.

1

Figure D.1: The first page (of two) of the prepared Spatio-Temporal Collaboration Questionnaire (STCQ) that was administered to each collaborator within the scope of the collaborative confirmative analysis evaluation (see Section 6.4.6.4).

[SC] Sharing Context: Characteristics and features of the shared space that facilitate and support focused and unfocused collaborative work, leading to shared understandings.

		strongly disagree	disagree	neutral	agree	strongly agree
SC.1	The collaborative features of the system allowed me to focus on the same subject as my partner.	○	○	○	○	○
SC.2	The collaborative features of the system allowed me to establish a dialog with my partner.	○	○	○	○	○
SC.3	The collaborative features of the system distracted me from my *individual* efforts.	○	○	○	○	○

[AO] Awareness of Others: The ability to understand your partner's activity during times of (1) focused collaboration and active communication (i.e., *group* efforts), as well as (2) more independent and *individual* work.

		not at all	a bit	some	a lot	always
AO.1	During your *group* efforts, how much were you aware of your partner's activities?	○	○	○	○	○
AO.2	During your *group* efforts, how much were you aware of your partner's location in space?	○	○	○	○	○
AO.3	During your *group* efforts, how much were you aware of your partner's time reference (time point / interval)?	○	○	○	○	○
AO.4	During your *individual* efforts, how much were you aware of your partner's activities?	○	○	○	○	○
AO.5	During your *individual* efforts, how much were you aware of your partner's location in space?	○	○	○	○	○
AO.6	During your *individual* efforts, how much were you aware of your partner's time reference (time point / interval)?	○	○	○	○	○

2

Figure D.2: The second page (of two) of the prepared Spatio-Temporal Collaboration Questionnaire (STCQ) that was administered to each collaborator within the scope of the collaborative confirmative analysis evaluation (see Section 6.4.6.4).

Disclaimer: The presented scenario and task are fictional, and have been exclusively created for the study you are participating in.

Scenario: It is the year 2X42. A series of scientific and technological advances made it possible to travel through the quantum realm. The exploration of many different variants of our dear Mother Earth followed in the years after. You are a two-person science team responsible for one such expedition. While one of you specializes on the collection and analysis of weather data, such as for instance sunlight and humidity levels, the other is an expert in the study and observation of plants, such as different types of fruits and vegetables.

After a joint excursion through the quantum realm during which you collected 150 days worth of data from different locations all over, what appears to be, the European landmass, you are now back in your research lab. Using the (non-immersive) weather terminal as well as the (immersive) plant exploration environment, you are ready to together take a closer look and make sense of your collected data.

Task: Your superintendent asked you for a report on the collected data. Collaboratively explore the collected weather and plant data in space and time, and use the provided tools to make assessments that describe the relationship between each plant and the two weather variables (sunlight and humidity). In short, based on your observations, determine the type of correlation between each weather and plant data, and additionally indicate how confident you are with those assessments. To support your conclusions, you should better write down noteworthy observations along the way.

Further Information:

- A correlation refers to the relationship between two variables.
- A positive correlation indicates that when one variable is increasing, the other variable is increasing as well. Or, when one variable is decreasing, the other variable is decreasing as well.
- A negative correlation indicates that when one variable is increasing, the other variable is decreasing (and vice versa).
- No correlation would indicate that when one variable is increasing, the other might be increasing, decreasing, or remain unchanged with equal probability.
- If you cannot determine the type of correlation based on your observations, please indicate so.
- You can assume that the location does not affect the correlations. A relationship between a weather variable and a plant would be the same across the planet, no matter the specific geographic location.

1

Figure D.3: The "science fictional" task scenario that was presented to the participant pairs within the scope of the collaborative confirmative analysis evaluation (see Section 6.4.6.3).

Synchronous Asymmetric Interaction within the context of Collaborative Immersive Analytics

Session - Date / Time: ____________________

Correlation: Based on your joint data exploration, please make assessments that describe the relationship between *fruit and sunlight*, as well as *fruit and humidity*.
Confidence: How *sure / confident* are you with your correlation assessment?

Fruit	**Sunlight**		**Humidity**	
	Correlation	*Confidence*	*Correlation*	*Confidence*
Apples	▫ Positive ▫ None ▫ Negative	▫ Do not know ▫ Low ▫ Medium ▫ High	▫ Positive ▫ None ▫ Negative	▫ Do not know ▫ Low ▫ Medium ▫ High
Oranges	▫ Positive ▫ None ▫ Negative	▫ Do not know ▫ Low ▫ Medium ▫ High	▫ Positive ▫ None ▫ Negative	▫ Do not know ▫ Low ▫ Medium ▫ High
Bananas	▫ Positive ▫ None ▫ Negative	▫ Do not know ▫ Low ▫ Medium ▫ High	▫ Positive ▫ None ▫ Negative	▫ Do not know ▫ Low ▫ Medium ▫ High
Berries	▫ Positive ▫ None ▫ Negative	▫ Do not know ▫ Low ▫ Medium ▫ High	▫ Positive ▫ None ▫ Negative	▫ Do not know ▫ Low ▫ Medium ▫ High
Grapes	▫ Positive ▫ None ▫ Negative	▫ Do not know ▫ Low ▫ Medium ▫ High	▫ Positive ▫ None ▫ Negative	▫ Do not know ▫ Low ▫ Medium ▫ High

Note: Once both of you agree that you have finished your joint data exploration, please say aloud "*We are done with the data exploration.*"

Figure D.4: The prepared physical task answer sheet for the *fruits* scenario that the non-immersed collaborator was in charge of within the scope of the collaborative confirmative analysis evaluation (see Section 6.4.6.3).

Synchronous Asymmetric Interaction within the context of Collaborative Immersive Analytics

Session - Date / Time: ____________________

Correlation: Based on your joint data exploration, please make assessments that describe the relationship between *vegetable and sunlight*, as well as *vegetable and humidity*.
Confidence: How *sure / confident* are you with your correlation assessment?

Vegetable	**Sunlight**		**Humidity**	
	Correlation	*Confidence*	*Correlation*	*Confidence*
Tomatoes	▫ Positive ▫ None ▫ Negative	▫ Do not know ▫ Low ▫ Medium ▫ High	▫ Positive ▫ None ▫ Negative	▫ Do not know ▫ Low ▫ Medium ▫ High
Carrots	▫ Positive ▫ None ▫ Negative	▫ Do not know ▫ Low ▫ Medium ▫ High	▫ Positive ▫ None ▫ Negative	▫ Do not know ▫ Low ▫ Medium ▫ High
Potatoes	▫ Positive ▫ None ▫ Negative	▫ Do not know ▫ Low ▫ Medium ▫ High	▫ Positive ▫ None ▫ Negative	▫ Do not know ▫ Low ▫ Medium ▫ High
Cabbages	▫ Positive ▫ None ▫ Negative	▫ Do not know ▫ Low ▫ Medium ▫ High	▫ Positive ▫ None ▫ Negative	▫ Do not know ▫ Low ▫ Medium ▫ High
Lettuces	▫ Positive ▫ None ▫ Negative	▫ Do not know ▫ Low ▫ Medium ▫ High	▫ Positive ▫ None ▫ Negative	▫ Do not know ▫ Low ▫ Medium ▫ High

Note: Once both of you agree that you have finished your joint data exploration, please say aloud "*We are done with the data exploration.*"

Figure D.5: The prepared physical task answer sheet for the *vegetables* scenario that the non-immersed collaborator was in charge of within the scope of the collaborative confirmative analysis evaluation (see Section 6.4.6.3).

Bibliography

Rossella Aiello, Gennaro Costagliola, Giancarlo Nota, and Fabrizio Torre. 3DRC Visualizations to Support the Reconciliation of Diverging Project Views. *International Journal of Software and Informatics (IJSI)*, 9(3):331–354, 2015. ISSN 1673-7288. URL http://www.ijsi.org/ijsi/article/abstract/i224.

Wolfgang Aigner, Silvia Miksch, Heidrun Schumann, and Christian Tominski. *Visualization of Time-Oriented Data*. Human-Computer Interaction Series (HCIS). Springer London, 1st edition, 2011. ISBN 9780857290786. doi:10.1007/978-0-85729-079-3. URL https://doi.org/10.1007/978-0-85729-079-3.

Mina Akaishi and Yoshihiro Okada. Time-tunnel: Visual Analysis Tool for Time-series Numerical Data and Its Aspects as Multimedia Presentation Tool. In *Proceedings of the Eigth International Conference on Information Visualization (IV 2004)*, pages 456–461, London, UK, 16 July 2004. Institute of Electrical and Electronics Engineers (IEEE). ISBN 0-7695-2177-0. doi:10.1109/IV.2004.1320184. URL https://doi.org/10.1109/IV.2004.1320184.

Toni Alatalo, Timo Koskela, Matti Pouke, Paula Alavesa, and Timo Ojala. VirtualOulu: Collaborative, Immersive and Extensible 3D City Model on the Web. In *Proceedings of the 21st International Conference on Web3D Technology (Web3D 2016)*, pages 95–103, Anaheim, California, USA, 22–24 July 2016. Association for Computing Machinery (ACM). ISBN 978-1-4503-4428-9. doi:10.1145/2945292.2945305. URL https://doi.org/10.1145/2945292.2945305.

Aris Alissandrakis, Nico Reski, Mikko Laitinen, Jukka Tyrkkö, Magnus Levin, and Jonas Lundberg. Visualizing dynamic text corpora using Virtual Reality. In *The 39th Annual Conference of the International Computer Archive for Modern and Medieval English (ICAME 39): Corpus Linguistics and Changing Society*, ICAME 39 Book of Abstracts, page 205, Tampere, Finland, 30 May–3 June 2018. International Computer Archive of Modern and Medieval English (ICAME). URL http://urn.kb.se/resolve?urn=urn:nbn:se:lnu:diva-75064.

Aris Alissandrakis, Nico Reski, Mikko Laitinen, Jukka Tyrkkö, Jonas Lundberg, and Magnus Levin. Visualizing rich corpus data using virtual reality. *Studies in Variation, Contacts and Change in English: Corpus Approaches into World Englishes and Language Contrasts*, 20, December 2019. ISSN 1797-4453. URL https://varieng.helsinki.fi/series/volumes/20/alissandrakis_et_al/.

Natalia Andrienko and Gennady Andrienko. *Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach*. Springer, Berlin, Heidelberg, 1st edition, 2006. ISBN 978-3-540-31190-4. doi:10.1007/3-540-31190-4. URL https://doi.org/10.1007/3-540-31190-4.

J. H. Erik Andriessen. Group Processes. In *Working with Groupware: Understanding and Evaluating Collaboration Technology*, Computer Supported Cooperative Work (CSCW), pages 89–124. Springer, London, 2003. ISBN 978-1-4471-0067-6. doi:10.1007/978-1-4471-0067-6_6. URL https://doi.org/10.1007/978-1-4471-0067-6_6.

Christopher R. Austin, Barrett Ens, Kadek Ananta Satriadi, and Bernhard Jenny. Elicitation study investigating hand and foot gesture interaction for immersive maps in augmented reality. *Cartography and Geographic Information Science*, 47(3):214–228, 7 January 2020. doi:10.1080/15230406.2019.1696232. URL https://doi.org/10.1080/15230406.2019.1696232.

Daniel Bachmann, Frank Weichert, and Gerhard Rinkenauer. Review of Three-Dimensional Human-Computer Interaction with Focus on the Leap Motion Controller. *Sensors*, 18(7):2194:1–39, 2018. doi:10.3390/s18072194. URL `https://doi.org/10.3390/s18072194`.

Ronald M. Baecker, Jonathan Grudin, William A. S. Buxton, and Saul Greenberg. Groupware and Computer-Supported Cooperative Work. In Ronald M. Baecker, Jonathan Grudin, William A. S. Buxton, and Saul Greenberg, editors, *Readings in Human-Computer Interaction: Toward the Year 2000*, Interactive Technologies, pages 741–753. Morgan Kaufmann, 2nd edition, 1 April 1995. ISBN 978-0-08-051574-8. doi:10.1016/B978-0-08-051574-8.50077-7. URL `https://doi.org/10.1016/B978-0-08-051574-8.50077-7`.

Muhammad Zeeshan Baig and Manolya Kavakli. A Survey on Psycho-Physiological Analysis & Measurement Methods in Multimodal Systems. *Multimodal Technologies and Interaction*, 3(2):37:1–20, 28 May 2019. doi:10.3390/mti3020037. URL `https://doi.org/10.3390/mti3020037`.

Aaron Bangor, Philip Kortum, and James Miller. Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. *Journal of Usability Studies*, 4(3):114–123, May 2009. URL `http://uxpajournal.org/determining-what-individual-sus-scores-mean-adding-an-adjective-rating-scale/`.

Martin S. Banks, Joohwan Kim, and Takashi Shibata. Insight into Vergence-Accommodation Mismatch. In *Proceedings of Head- and Helmet-Mounted Displays XVIII: Design and Applications (SPIE Vol. 8735)*, pages 873509:1–12, Baltimore, Maryland, USA, 16 May 2013. Society of Photo-Optical Instrumentation Engineers (SPIE). doi:10.1117/12.2019866. URL `https://doi.org/10.1117/12.2019866`.

Louise Barkhuus and Jennifer A. Rode. From Mice to Men - 24 Years of Evaluation in CHI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2007)*, pages 1–16, San Jose, California, USA, 28 April–3 May 2007. Association for Computing Machinery (ACM). doi:10.1145/1240624.2180963. URL `https://doi.org/10.1145/1240624.2180963`.

Victor R. Basili, Gianluigi Caldiera, and H. Dieter Rombach. Goal Question Metric Paradigm. In John J. Marciniak, editor, *Encyclopedia of Software Engineering*, pages 528–532. John Wiley & Sons, Inc., 1994.

Steve Benford and Lennart E. Fahlén. A Spatial Model of Interaction in Large Virtual Environments. In *Proceedings of the Third European Conference on Computer-Supported Cooperative Work (ECSCW 1993)*, pages 109–124, Milan, Italy, 13–17 September 1993. Springer, Dordrecht. ISBN 978-94-011-2094-4. doi:10.1007/978-94-011-2094-4_8. URL `https:doi.org/10.1007/978-94-011-2094-4_8`.

Steve Benford, John Bowers, Lennart E. Fahlén, and Chris Greenhalgh. Managing mutual awareness in collaborative virtual environments. In *Proceedings of the Conference on Virtual Reality Software and Technology (VRST 1994)*, pages 223–236, Singapore, 23–26 August 1994. World Scientific Publishing Co., Inc. ISBN 9810218672. URL `https://dl.acm.org/doi/10.5555/207072.207146`.

Alberto Betella, Enrique Martínez Bueno, Wipawee Kongsantad, Riccardo Zucca, Xerxes D. Arsiwalla, Pedro Omedas, and Paul F. M. J. Verschure. Understanding Large Network Datasets through Embodied Interaction in Virtual Reality. In *Proceedings of the 2014 Virtual Reality International Conference (VRIC 2014)*, pages 1–7, Laval, France, 9–11 April 2014. Association for Computing Machinery (ACM). ISBN 978-1-4503-2626-1. doi:10.1145/2617841.2620711. URL `https://doi.org/10.1145/2617841.2620711`.

Mark Billinghurst, Maxime Cordeil, Anastasia Bezerianos, and Todd Margolis. Collaborative Immersive Analytics. In Kim Marriott, Falk Schreiber, Tim Dwyer, Karsten Klein, Nathalie Henry Riche, Takayuki Itoh, Wolfgang Stuerzlinger, and Bruce H. Thomas, editors, *Immersive Analytics*, Lecture

Notes in Computer Science (LNCS, volume 11190), pages 221–257. Springer, Cham, first online edition, 16 October 2018. ISBN 978-3-030-01388-2. doi:10.1007/978-3-030-01388-2_8. URL `https://doi.org/10.1007/978-3-030-01388-2_8`.

Pauline Bimberg, Tim Weissker, and Alexander Kulik. On the Usage of the Simulator Sickness Questionnaire for Virtual Reality Research. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (IEEE VRW 2020)*, pages 464–467, Atlanta, Georgia, USA, 22–26 March 2020. Institute of Electrical and Electronics Engineers (IEEE). ISBN 978-1-7281-6532-5. doi:10.1109/VRW50115.2020.00098. URL `https://doi.org/10.1109/VRW50115.2020.00098`.

Alexander Bogner, Beate Littig, and Wolfgang Menz. Introduction: Expert Interviews - An Introduction to a New Methodological Debate. In Alexander Bogner, Beate Littig, and Wolfgang Menz, editors, *Interviewing Experts*, Research Methods Series (REMES), pages 1–13. Palgrave Macmillan, London, 1st edition, 2010. ISBN 978-0-230-24427-6. doi:10.1057/9780230244276_1. URL `https://doi.org/10.1057/9780230244276_1`.

Richard A. Bolt. "Put-that-there": Voice and Gesture at the Graphics Interface. *ACM SIGGRAPH Computer Graphics*, 14(3):262–270, 1 July 1980. doi:10.1145/965105.807503. URL `https://doi.org/10.1145/965105.807503`.

Stéphane Bouchard, Geneviève Robillard, and Patrice Renaud. Revising the factor structure of the Simulator Sickness Questionnaire. *Annual Review of CyberTherapy and Telemedicine*, 5(Summer 2007): 117–122, 2007. URL `https://www.arctt.info/volume-5-summer-2007`.

Doug A. Bowman and Ryan P. McMahan. Virtual Reality: How Much Immersion Is Enough? *Computer*, 40(7):36–43, 13 August 2007. doi:10.1109/MC.2007.257. URL `https://doi.org/10.1109/MC.2007.257`.

Nora Mills Boyd and James Bogen. Theory and Observation in Science. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2021 edition, 2021. URL `https://plato.stanford.edu/archives/win2021/entries/science-theory-observation/`.

Matthew Brehmer and Tamara Munzner. A Multi-Level Typology of Abstract Visualization Tasks. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2376–2885, 16 October 2013. doi:10.1109/TVCG.2013.124. URL `https://doi.org/10.1109/TVCG.2013.124`.

Nathalie Bressa, Henrik Korsgaard, Aurélien Tabard, Steven Houben, and Jo Vermeulen. What's the Situation with Situated Visualization? A Survey and Perspectives on Situatedness. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):107–117, January 2022. doi:10.1109/TVCG.2021.3114835. URL `https://doi.org/10.1109/TVCG.2021.3114835`.

John Brooke. SUS: A 'Quick and Dirty' Usability Scale. In Patrick W. Jordan, B. Thomas, Ian Lyall McClelland, and Bernard Weerdmeester, editors, *Usability Evaluation In Industry*, pages 189–194. CRC Press, 11 June 1996. ISBN 978-3-030-01388-2. doi:10.1201/9781498710411. URL `https://doi.org/10.1201/9781498710411`.

John Brooke. SUS: A Retrospective. *Journal of Usability Studies*, 8(2):29–40, February 2013. URL `http://uxpajournal.org/sus-a-retrospective/`.

Grigore Burdea, Paul Richard, and Philippe Coiffet. Multimodal Virtual Reality: Input-Output Devices, System Integration, and Human Factors. *International Journal of Human-Computer Interaction*, 8(1):5–24, 1996. doi:10.1080/10447319609526138. URL `https://doi.org/10.1080/10447319609526138`.

Peter W. S. Butcher, Nigel W. John, and Panagiotis D. Ritsos. VRIA - A Framework for Immersive Analytics on the Web. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems (CHI EA 2019)*, pages LBW2615:1–6, Glasgow, Schotland, UK, 4–9 May 2019. Association for Computing Machinery (ACM). ISBN 978-1-4503-5971-9. doi:10.1145/3290607.3312798. URL `https://doi.org/10.1145/3290607.3312798`.

Wolfgang Büschel, Jian Chen, Raimund Dachselt, Steven Drucker, Tim Dwyer, Carsten Görg, Tobias Isenberg, Andreas Kerren, Chris North, and Wolfgang Stuerzlinger. Interaction for Immersive Analytics. In Kim Marriott, Falk Schreiber, Tim Dwyer, Karsten Klein, Nathalie Henry Riche, Takayuki Itoh, Wolfgang Stuerzlinger, and Bruce H. Thomas, editors, *Immersive Analytics*, Lecture Notes in Computer Science (LNCS, volume 11190), pages 95–138. Springer, Cham, first online edition, 16 October 2018. ISBN 978-3-030-01388-2. doi:10.1007/978-3-030-01388-2_4. URL `https://doi.org/10.1007/978-3-030-01388-2_4`.

Wolfgang Büschel, Anke Lehmann, and Raimund Dachselt. MIRIA: A Mixed Reality Toolkit for the In-Situ Visualization and Analysis of Spatio-Temporal Interaction Data. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI 2021)*, pages 470:1–15, Yokohama, Japan, 8–13 May 2021. Association for Computing Machinery (ACM). doi:10.1145/3411764.3445651. URL `https://doi.org/10.1145/3411764.3445651`.

Giuseppe Caggianese, Luigi Gallo, and Pietro Neroni. The Vive Controllers vs. Leap Motion for Interactions in Virtual Environments: A Comparative Evaluation. In Giuseppe De Pietro, Luigi Gallo, Robert J. Howlett, Lakhmi C. Jain, and Ljubo Vlacic, editors, *Intelligent Interactive Multimedia Systems and Services (KES-IIMSS-18) 2018*, Smart Innovation, Systems and Technologies (SIST, volume 98), pages 24–33. Springer, Cham, first online edition, 12 June 2018. ISBN 978-3-319-92231-7. doi:10.1007/978-3-319-92231-7_3. URL `https://doi.org/10.1007/978-3-319-92231-7_3`.

Jorge C. S. Cardoso. Comparison of Gesture, Gamepad, and Gaze-based Locomotion for VR Worlds. In *Proceedings of the 22nd ACM Symposium on Virtual Reality Software and Technology (VRST 2016)*, pages 319–-320, Munich, Germany, 2–4 November 2016. Association for Computing Machinery (ACM). ISBN 978-1-4503-4491-3. doi:10.1145/2993369.2996327. URL `https://doi.org/10.1145/2993369.2996327`.

Julien Casarin, Nicolas Pacqueriaud, and Dominique Bechmann. UMI3D: A Unity3D Toolbox to Support CSCW Systems Properties in Generic 3D User Interfaces. *Proceedings of the ACM on Human-Computer Interaction (CSCW)*, 2(CSCW):29:1–20, November 2018. doi:10.1145/3274298. URL `https://doi.org/10.1145/3274298`.

Marco Cavallo, Mishal Dholakia, Matous Havlena, Kenneth Ocheltree, and Mark Podlaseck. Immersive Insights: A Hybrid Analytics System for Collaborative Exploratory Data Analysis. In *Proceedings of the 25th ACM Symposium on Virtual Reality Software and Technology (VRST 2019)*, pages 9:1–12, Parramatta, NSW, Australia, 12–15 November 2019. Association for Computing Machinery (ACM). ISBN 978-1-4503-7001-1. doi:10.1145/3359996.3364242. URL `https://doi.org/10.1145/3359996.3364242`.

Daniel Cernea. *User-Centered Collaborative Visualization*. PhD thesis, Linnaeus University, Växjö, Sweden, 11 June 2015. URL `http://urn.kb.se/resolve?urn=urn:nbn:se:lnu:diva-42976`.

Tom Chandler, Maxime Cordeil, Tobias Czauderna, Tim Dwyer, Jaroslaw Glowacki, Cagatay Goncu, Matthias Klapperstueck, Karsten Klein, Kim Marriott, Falk Schreiber, and Elliot Wilson. Immersive Analytics. In *2015 Big Data Visual Analytics (BDVA)*, volume 1, pages 1–8, Hobart, Tasmania, Australia, 2 June 2015. Institute of Electrical and Electronics Engineers (IEEE). doi:10.1109/BDVA.2015.7314296. URL `https://doi.org/10.1109/BDVA.2015.7314296`.

Janglin Chen, Wayne Cranton, and Mark Fihn, editors. *Handbook of Visual Display Technology*. Springer, Berlin, Heidelberg, 1st edition, 2012. ISBN 978-3-540-79567-4. doi:10.1007/978-3-540-79567-4. URL https://doi.org/10.1007/978-3-540-79567-4.

Taizhou Chen, Yi-Shiun Wu, and Kening Zhu. Investigating Different Modalities of Directional Cues for Multi-task Visual-Searching Scenario in Virtual Reality. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology (VRST 2018)*, pages 41:1–5, Tokyo, Japan, 28 November–1 December 2018. Association for Computing Machinery (ACM). ISBN 978-1-4503-6086-9. doi:10.1145/3281505.3281516. URL https://doi.org/10.1145/3281505.3281516.

Elizabeth F. Churchill and David Snowdon. Collaborative Virtual Environments: An Introductory Review of Issues and Systems. *Virtual Reality*, 3(1):3–15, March 1998. doi:10.1007/BF01409793. URL https://doi.org/10.1007/BF01409793.

Herbert H. Clark and Susan E. Brennan. Grounding in communication. In Lauren Resnick, John M. Levine, and Stephanie D. Teasley, editors, *Perspectives on socially shared cognition*, pages 127–149. American Psychological Association, 1991. ISBN 0-8058-4014-1. doi:10.1037/10096-006. URL https://doi.org/10.1037/10096-006.

Damien Clergeaud, Joan Sol Roo, Martin Hachet, and Pascal Guitton. Towards Seamless Interaction between Physical and Virtual Locations for Asymmetric Collaboration. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology (VRST 2017)*, pages 17:1–4, Gothenburg, Sweden, 8–10 November 2017. Association for Computing Machinery (ACM). ISBN 978-1-4503-5548-3. doi:10.1145/3139131.3139165. URL https://doi.org/10.1145/3139131.3139165.

Ernest Cline. *Ready Player One*. The Crown Publishing Group, 1st edition, 16 August 2011. ISBN 9780307887436.

Maxime Cordeil, Andrew Cunningham, Tim Dwyer, Bruce H. Thomas, and Kim Marriott. ImAxes: Immersive Axes as Embodied Affordances for Interactive Multivariate Data Visualisation. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (UIST 2017)*, pages 71–83, Quebec City, Quebec, Canada, 22–25 October 2017a. Association for Computing Machinery (ACM). ISBN 978-1-4503-4981-9. doi:10.1145/3126594.3126613. URL https://doi.org/10.1145/3126594.3126613.

Maxime Cordeil, Tim Dwyer, Karsten Klein, Bireswar Laha, Kim Marriott, and Bruce H. Thomas. Immersive Collaborative Analysis of Network Connectivity: CAVE-style or Head-Mounted Display? *IEEE Transactions on Visualization and Computer Graphics*, 23(1):441–450, 1 January 2017b. doi:10.1109/TVCG.2016.2599107. URL https://doi.org/10.1109/TVCG.2016.2599107.

Maxime Cordeil, Andrew Cunningham, Benjamin Bach, Christophe Hurter, Bruce H. Thomas, Kim Marriott, and Tim Dwyer. IATK: An Immersive Analytics Toolkit. In *Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces (VR 2019)*, pages 200–209, Osaka, Japan, 23–27 March 2019. Institute of Electrical and Electronics Engineers (IEEE). ISBN 978-1-7281-1377-7. doi:10.1109/VR.2019.8797978. URL https://doi.org/10.1109/VR.2019.8797978.

Marcos Túlio Silva Costa, Lanna Pinheiro Vieira, Elizabete de Oliveira Barbosa, Luciana Mendes Oliveira, Pauline Maillot, César Augusto Ottero Vaghetti, Mauro Giovani Carta, Sérgio Machado, Valeska Gatica-Rojas, and Renato Sobral Monteiro-Junior. Virtual Reality-Based Exercise with Exergames as Medicine in Different Contexts: A Short Review. *Clinical Practice & Epidemiology in Mental Health*, 15:15–20, 31 January 2019. doi:10.2174/1745017901915010015. URL https://doi.org/10.2174/1745017901915010015.

Armando Cruz, Leonel Morgado, Hugo Paredes, Benjamim Fonseca, and Paulo Martins. Fitting Three Dimensional Virtual Worlds into CSCW. In *Proceedings of the 19th International Conference* *on Computer Supported Cooperative Work in Design (CSCWD 2015)*, pages 419–424, Calabria, Italy, 6–8 May 2015. Institute of Electrical and Electronics Engineers (IEEE). ISBN 978-1-4799-2002-0. doi:10.1109/CSCWD.2015.7230996. URL `https://doi.org/10.1109/CSCWD.2015.7230996`.

Carolina Cruz-Neira, Daniel J. Sandin, Thomas A. DeFanti, Robert V. Kenyon, and John C. Hart. The CAVE: Audio Visual Experience Automatic Virtual Environment. *Communications of the ACM*, 35(6): 64–72, 1 June 1992. doi:10.1145/129888.129892. URL `https://doi.org/10.1145/129888.129892`.

Carolina Cruz-Neira, Jason Leigh, Michael Papka, Craig Barnes, Steven M. Cohen, Sumit Das, Roger Engelmann, Randy Hudson, Trina Roy, Lewis Siegel, Christina Vasilakis, Thomas A. DeFanti, and Daniel J. Sandin. Scientists in Wonderland: A Report on Visualization Application the CAVE Virtual Reality Environment. In *Proceedings of the 1993 IEEE Research Properties in Virtual Reality Symposium*, pages 59–66, San Jose, California, USA, 25–26 October 1993. Institute of Electrical and Electronics Engineers (IEEE). doi:10.1109/VRAIS.1993.378262. URL `https://doi.org/10.1109/VRAIS.1993.378262`.

Mihaly Csikszentmihalyi and Isabella Selega Csikszentmihalyi, editors. *Optimal Experience: Psychological Studies of Flow in Consciousness*. Cambridge University Press, 31 July 1992. ISBN 9780521438094.

Raimund Dachselt and Anett Hübner. Three-dimensional menus: A survey and taxonomy. *Computers & Graphics*, 31:53–65, 18 October 2007. doi:10.1016/j.cag.2006.09.006. URL `https://doi.org/10.1016/j.cag.2006.09.006`.

Ryan Anthony J. de Belen, Huyen Nguyen, Daniel Filonik, Dennis Del Favero, and Tomasz Bednarz. A systematic review of the current state of collaborative mixed reality technologies: 2013–2018. *AIMS Electronics and Electrical Engineering*, 3, Special Issue: Augmented and Virtual Reality for Industry 4.0(2):181–223, 17 May 2019. doi:10.3934/ElectrEng.2019.2.181. URL `https://doi.org/10.3934/ElectrEng.2019.2.181`.

William Delamare, Thomas Janssoone Céline Coutrix, and Laurence Nigay. Designing 3D Gesture Guidance: Visual Feedback and Feedforward Design Options. In *Proceedings of the the International Working Conference on Advanced Visual Interfaces (AVI 2016)*, pages 152–159, Bari, Italy, 7–10 June 2016. Association for Computing Machinery (ACM). ISBN 9781450341318. doi:10.1145/2909132.2909260. URL `https://doi.org/10.1145/2909132.2909260`.

Alan Dix. Computer Supported Cooperative Work: A Framework. In Duska Rosenberg and Chris Hutchison, editors, *Design Issues in CSCW*, Computer Supported Cooperative Work, pages 9–26. Springer, London, 1994. ISBN 978-1-4471-2029-2. doi:10.1007/978-3-030-01388-2_2. URL `https://doi.org/10.1007/978-1-4471-2029-2_2`.

Alan Dix. *Statistics for HCI: Making Sense of Quantitative Data*. Synthesis Lectures on Human-Centered Informatics. Morgan & Claypool, 1st edition, April 2020. ISBN 9781681737430. doi:10.2200/S00974ED1V01Y201912HCI044. URL `https://doi.org/10.2200/S00974ED1V01Y201912HCI044`.

Alan Dix, Janet Finlay, Gregory D. Abowd, and Russel Beale. *Human-Computer Interaction*. Pearson, 3rd edition, 2004. ISBN 9780130461094.

Ciro Donalek, S. George Djorgovski, Alex Cioc, Anwell Wang, Jerry Zhang, Elizabeth Lawler, Stacy Yeh, Ashish Mahabal, Matthew Graham, Andrew Drake, Scott Davidoff, Jeffrey S. Norris, and Giuseppe Longo. Immersive and Collaborative Data Visualization Using Virtual Reality Platforms. In *2014 IEEE International Conference on Big Data (Big Data)*, pages 609–614, Washington, District of Columbia, USA, 27–30 October 2014. Institute of Electrical and Electronics Engineers (IEEE). doi:10.1109/BigData.2014.7004282. URL `https://doi.org/10.1109/BigData.2014.7004282`.

Geoffrey M. Draper, Yarden Livnat, and Richard F. Riesenfeld. A Survey of Radial Methods for Information Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 15(5):759–776, September–October 2009. doi:10.1109/TVCG.2009.23. URL `https://doi.org/10.1109/TVCG.2009.23`.

Aaron Duane and Björn Þór Jónsson. ViRMA: Virtual Reality Multimedia Analytics at LSC 2021. In *Proceedings of the 4th Annual on Lifelog Search Challenge (LSC 2021)*, pages 29–34, Taipei, Taiwan, 21 August 2021. Association for Computing Machinery (ACM). doi:10.1145/3463948.3469067. URL `https://doi.org/10.1145/3463948.3469067`.

Tim Dwyer, Kim Marriott, Tobias Isenberg, Karsten Klein, Nathalie Henry Riche, Flak Schreiber, and Wolfgang Stuerzlinger. Immersive Analytics: An Introduction. In Kim Marriott, Falk Schreiber, Tim Dwyer, Karsten Klein, Nathalie Henry Riche, Takayuki Itoh, Wolfgang Stuerzlinger, and Bruce H. Thomas, editors, *Immersive Analytics*, Lecture Notes in Computer Science (LNCS, volume 11190), pages 1–23. Springer, Cham, first online edition, 16 October 2018. ISBN 978-3-030-01388-2. doi:10.1007/978-3-030-01388-2_1. URL `https://doi.org/10.1007/978-3-030-01388-2_1`.

Barrett Ens, Joel Lanir, Anthony Tang, Scott Bateman, Gun Lee, Thammathip Piumsomboon, and Mark Billinghurst. Revisiting collaboration through Mixed Reality: The evolution of Groupware. *International Journal of Human-Computer Studies*, 131:81–98, November 2019. doi:10.1016/j.ijhcs.2019.05.011. URL `https://doi.org/10.1016/j.ijhcs.2019.05.011`.

Barrett Ens, Benjamin Bach, Maxime Cordeil, Ulrich Engelke, Marcos Serrano, Wesley Willett, Arnaud Prouzeau, Christoph Anthes, Wolfgang Büschel, Cody Dunne, Tim Dwyer, Jens Grubert, Jason H. Haga, Nurit Kirshenbaum, Dylan Kobayashi, Tica Lin, Monsurat Olaosebikan, Fabian Pointecker, David Saffo, Nazmus Saquib, Dieter Schmalstieg, Danielle Albers Szafir, Matt Whitlock, and Yalong Yang. Grand Challenges in Immersive Analytics. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI 2021)*, pages 459:1–17, Yokohama, Japan, 8–13 May 2021. Association for Computing Machinery (ACM). doi:10.1145/3411764.3446866. URL `https://doi.org/10.1145/3411764.3446866`.

Alessandro Febretti, Arthur Nishimoto, Terrance Thigpen, Jonas Talandis, Lance Long, J. D. Pirtle, Tom Peterka, Alan Verlo, Maxine Brown, Dana Plepys, Dan Sandin, Luc Renambot, Andrew Johnson, and Jason Leigh. CAVE2: A Hybrid Reality Environment for Immersive Simulation and Information Analytics. In *Proceedings of The Engineering Reality of Virtual Reality (SPIE Vol. 8649)*, pages 864903:1–12, Burlingame, California, USA, 4 March 2013. Society of Photo-Optical Instrumentation Engineers (SPIE). doi:10.1117/12.2005484. URL `https://doi.org/10.1117/12.2005484`.

Lucas Figueiredo, Eduardo Rodrigues, João Teixeira, and Veronica Teichrieb. A comparative evaluation of direct hand and wand interactions on consumer devices. *Computers & Graphics*, 77:108–121, 18 October 2018. doi:10.1016/j.cag.2018.10.006. URL `https://doi.org/10.1016/j.cag.2018.10.006`.

Wim Fikkert, Marco D'Ambros, Torsten Bierz, and T. J. Jankun-Kelly. Interacting with Visualizations. In Andreas Kerren, Achim Ebert, and Jörg Meyer, editors, *Human-Centered Visualization Environments: GI-Dagstuhl Research Seminar, Dagstuhl Castle, Germany, March 5-8, 2006, Revised Lectures*, Lecture Notes in Computer Science (LNCS, volume 4417), pages 77–162. Springer, Berlin, Heidelberg, 2007. ISBN 978-3-540-71949-6. doi:10.1007/978-3-540-71949-6_3. URL `https://doi.org/10.1007/978-3-540-71949-6_3`.

Florian Fittkau, Alexander Krause, and Wilhelm Hasselbring. Exploring Software Cities in Virtual Reality. In *Proceedings of the 3rd Working Conference on Software Visualization (VISSOFT 2015)*, pages 130–134, Bremen, Germany, 23 November 2015. Institute of Electrical and Electronics Engineers (IEEE). ISBN 978-1-4673-7526-9. doi:10.1109/VISSOFT.2015.7332423. URL `https://doi.org/10.1109/VISSOFT.2015.7332423`.

Adrien Fonnet and Yannick Prié. Survey of Immersive Analytics. *IEEE Transactions on Visualization and Computer Graphics*, 27(3):2101–2122, 1 March 2021. doi:10.1109/TVCG.2019.2929033. URL `https://doi.org/10.1109/TVCG.2019.2929033`.

Marsha E. Fonteyn, Benjamin Kuipers, and Susan J. Grobe. A Description of Think Aloud Method and Protocol Analysis. *Qualitative Health Research*, 3(4):430–441, 1 November 1993. doi:10.1177/104973239300300403. URL `https://doi.org/10.1177/104973239300300403`.

Clifton Forlines and Kent Wittenburg. Wakame: Sense Making of Multi-Dimensional Spatial-Temporal Data. In *Proceedings of the the International Conference on Advanced Visual Interfaces (AVI 2010)*, pages 33–40, Roma, Italy, 26–28 May 2010. Association for Computing Machinery (ACM). ISBN 9781450300766. doi:10.1145/1842993.1843000. URL `https://doi.org/10.1145/1842993.1843000`.

Maarten Franssen, Gert-Jan Lokhorst, and Ibo van de Poel. Philosophy of Technology. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2018 edition, 2018. URL `https://plato.stanford.edu/archives/fall2018/entries/technology/`.

Bernhard Fröhler, Christoph Anthes, Fabian Pointecker, Judith Friedl, Daniel Schwajda, Andreas Riegler, Shailesh Tripathi, Clemens Holzmann, Manuel Brunner, Herbert Jodlbauer, Hans-Christian Jetter, and Christoph Heinzl. A Survey on Cross-Virtuality Analytics. *Computer Graphics Forum*, 41 (1):465–494, February 2022. doi:10.1111/cgf.14447. URL `https://doi.org/10.1111/cgf.14447`.

William Gibson. *Neuromancer*. New Ace SF Special. Ace Science Fiction Books, 1st edition, 1 July 1984. ISBN 9780441569564.

Elizabeth Goodman, Mike Kuniavsky, and Andrea Moed. *Observing the User Experience: A Practitioner's Guide to User Research*. Elsevier Inc., 2nd edition, 2012. ISBN 978-0-12-384869-7. doi:10.1016/C2010-0-64844-9. URL `https://doi.org/10.1016/C2010-0-64844-9`.

Michael J. Gourlay and Robert T. Held. Head-Mounted-Display Tracking for Augmented and Virtual Reality. *Information Display*, 33(1):6–10, 1 January 2017. doi:10.1002/j.2637-496X.2017.tb00962.x. URL `https://doi.org/10.1002/j.2637-496X.2017.tb00962.x`.

Jerônimo Gustavo Grandi, Henrique Galvan Debarba, and Anderson Maciel. Characterizing Asymmetric Collaborative Interactions in Virtual and Augmented Realities. In *Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces (VR 2019)*, pages 127–135, Osaka, Japan, 23–27 March 2019. Institute of Electrical and Electronics Engineers (IEEE). ISBN 978-1-7281-1377-7. doi:10.1109/VR.2019.8798080. URL `https://doi.org/10.1109/VR.2019.8798080`.

Jan Gugenheimer, Evgeny Stemasov, Julian Frommel, and Enrico Rukzio. ShareVR: Enabling Co-Located Experiences for Virtual Reality between HMD and Non-HMD Users. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI 2017)*, pages 4021–4033, Denver, Colorado, USA, 6–11 May 2017. Association for Computing Machinery (ACM). doi:10.1145/3025453.3025683. URL `https://doi.org/10.1145/3025453.3025683`.

Elisa Gusai, Chiara Bassano, Fabio Solari, and Manuela Chessa. Interaction in an Immersive Collaborative Virtual Reality Environment: A Comparison Between Leap Motion and HTC Controllers. In Sebastiano Battiato, Giovanni Maria Farinella, Marco Leo, and Giovanni Gallo, editors, *New Trends in Image Analysis and Processing – International Conference on Image Analysis and Processing (ICIAP) 2017*, Lecture Notes in Computer Science (LNIP, volume 10590), pages 290–300. Springer, Cham, first online edition, 31 December 2017. ISBN 978-3-319-70742-6. doi:10.1007/978-3-319-70742-6_27. URL `https://doi.org/10.1007/978-3-319-70742-6_27`.

Carl Gutwin and Saul Greenberg. A Descriptive Framework of Workspace Awareness for Real-Time Groupware. *Computer Supported Cooperative Work (CSCW)*, 11(3):411–446, 1 September 2002. doi:10.1023/A:1021271517844. URL `https://doi.org/10.1023/A:1021271517844`.

Richard Hackathorn and Todd Margolis. Immersive Analytics: Building Virtual Data Worlds for Collaborative Decision Support. In *2016 Workshop on Immersive Analytics (IA)*, pages 44–47, Greenville, South Carolina, USA, 20 March 2016. Institute of Electrical and Electronics Engineers (IEEE). doi:10.1109/IMMERSIVE.2016.7932382. URL `https://doi.org/10.1109/IMMERSIVE.2016.7932382`.

Steven T. Hackstadt and Allen D. Malony. Visualizing Parallel Programs and Performance. *IEEE Computer Graphics and Applications*, 15(4):12–14, July 1995. doi:10.1109/38.391482. URL `https://doi.org/10.1109/38.391482`.

Sven Ove Hansson. Science and Pseudo-Science. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2021 edition, 2021. URL `https://plato.stanford.edu/archives/fall2021/entries/pseudo-science/`.

Sandra G. Hart. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(9):904–908, October 2006. doi:10.1177/154193120605000909. URL `https://doi.org/10.1177/154193120605000909`.

Sandra G. Hart and Lowell E. Staveland. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Advances in Psychology*, 52:139–183, 1988. ISSN 0166-4115. doi:10.1016/S0166-4115(08)62386-9. URL `https://doi.org/10.1016/S0166-4115(08)62386-9`.

Michael T. Heath, Allen D. Malony, and Diane T. Rover. Parallel Performance Visualization: From Practice to Theory. *IEEE Parallel & Distributed Technology: Systems & Applications*, 3(4):44–60, 1995. doi:10.1109/88.473613. URL `https://doi.org/10.1109/88.473613`.

Jeffrey Heer and Maneesh Agrawala. Design Considerations for Collaborative Visual Analytics. In *2007 IEEE Symposium on Visual Analytics Science and Technology*, pages 171–178, Sacramento, California, USA, 30 October–1 November 2008. Institute of Electrical and Electronics Engineers (IEEE). ISBN 978-1-4244-1659-2. doi:10.1109/VAST.2007.4389011. URL `https://doi.org/10.1109/VAST.2007.4389011`.

Paul Heidicker, Eike Langbehn, and Frank Steinicke. Influence of Avatar Appearance on Presence in Social VR. In *Proceedings of the IEEE Symposium on 3D User Interfaces (3DUI 2017)*, pages 233–234, Los Angeles, California, USA, 18–19 March 2017. Institute of Electrical and Electronics Engineers (IEEE). ISBN 978-1-5090-6716-9. doi:10.1109/3DUI.2017.7893357. URL `https://doi.org/10.1109/3DUI.2017.7893357`.

Teresa Hirzle, Maurice Cordts, Enrico Rukzio, Jan Gugenheimer, and Andreas Bulling. A Critical Assessment of the Use of SSQ as a Measure of General Discomfort in VR Head-Mounted Displays. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI 2021)*, pages 530:1–14, Yokohama, Japan, 8–13 May 2021. Association for Computing Machinery (ACM). doi:10.1145/3411764.3445361. URL `https://doi.org/10.1145/3411764.3445361`.

Harri Holappa, Johanna Ylipulli, Sami Rautiainen, Ilya Minyaev, Matti Pouke, and Timo Ojala. VR Application for Technology Education in a Public Library. In *Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia (MUM 2018)*, pages 521–527, Cairo, Egypt, 25–28 November 2018. Association for Computing Machinery (ACM). ISBN 978-1-4503-6594-9. doi:10.1145/3282894.3289719. URL `https://doi.org/10.1145/3282894.3289719`.

Yi-Jheng Huang, Takanori Fujiwara, Yun-Xuan Lin, Wen-Chieh Lin, and Kwan-Liu Ma. A Gesture System for Graph Visualization in Virtual Reality Environments. In *2017 IEEE Pacific Visualization Symposium (PacificVis)*, pages 41–45, Seoul, Korea, 18–21 April 2017. Institute of Electrical and Electronics Engineers (IEEE). doi:10.1109/PACIFICVIS.2017.8031577. URL `https://doi.org/10.1109/PACIFICVIS.2017.8031577`.

Wijnand A. IJsselsteijn, Yvonne A. W. de Kort, and Karolin Poels. The Game Experience Questionnaire. Technical report, Technische Universiteit Eindhoven, 1 January 2013. URL https://research.tue.nl/en/publications/the-game-experience-questionnaire.

Petra Isenberg, Niklas Elmqvist, Jean Scholtz, Daniel Cernea, Kwan-Liu Ma, and Hans Hagen. Collaborative Visualization: Definition, challenges, and research agenda. *Information Visualization*, 10(4):310–326, 1 October 2011. doi:10.1177/1473871611412817. URL https://doi.org/10.1177/1473871611412817.

Petra Isenberg, Danyel Fisher, Sharoda A. Paul, Meredith Ringel Morris, Kori Inkpen, and Mary Czerwinski. Co-Located Collaborative Visual Analytics around a Tabletop Display. *IEEE Transactions on Visualization and Computer Graphics*, 18(5):689–792, May 2012. doi:10.1109/TVCG.2011.287. URL https://doi.org/10.1109/TVCG.2011.287.

Petra Isenberg, Bongshin Lee, Huamin Qu, and Maxime Cordeil. Immersive Visual Data Stories. In Kim Marriott, Falk Schreiber, Tim Dwyer, Karsten Klein, Nathalie Henry Riche, Takayuki Itoh, Wolfgang Stuerzlinger, and Bruce H. Thomas, editors, *Immersive Analytics*, Lecture Notes in Computer Science (LNCS, volume 11190), pages 165–184. Springer, Cham, first online edition, 16 October 2018. ISBN 978-3-030-01388-2. doi:10.1007/978-3-030-01388-2_6. URL https://doi.org/10.1007/978-3-030-01388-2_6.

Tobias Isenberg. An Interaction Continuum for Visualization. In Yvonne Jansen, Petra Isenberg, Jason Dykes, Sheelagh Carpendale, and Dan Keefe, editors, *Proceedings of the VIS Workshop on "Death of the Desktop: Envisioning Visualization without Desktop Computing"*, pages 1–3, Paris, France, 9 November 2014. URL https://hal.inria.fr/hal-01095454.

Alexander Ivanov, Kurtis Danyluk, Christian Jacob, and Wesley Willett. A Walk Among the Data: Exploration and Anthropomorphism in Immersive Unit Visualizations. *IEEE Computer Graphics and Applications*, 39:19–28, May–June 2019. doi:10.1109/MCG.2019.2898941. URL https://doi.org/10.1109/MCG.2019.2898941.

Robert Johansen, Jeff Charles, Robert Mittman, and Paul Saffo. *GroupWare: Computer Support for Business Teams*. The Free Press, 1 October 1988. ISBN 0029164915.

Jari Kangas, Sriram Kishore Kumar, Helena Mehtonen, Jorma Järnstedt, and Roope Raisamo. Trade-Off between Task Accuracy, Task Completion Time and Naturalness for Direct Object Manipulation in Virtual Reality. *Multimodal Technologies and Interaction*, 6(1):6:1–15, 10 January 2022. doi:10.3390/mti6010006. URL https://doi.org/10.3390/mti6010006.

Robert S. Kennedy, Norman E. Lane, Kevin S. Berbaum, and Michael G. Lilienthal. Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology*, 3(3):203–220, 1993. doi:10.1207/s15327108ijap0303_3. URL https://doi.org/10.1207/s15327108ijap0303_3.

Andreas Kerren and Ilir Jusufi. Novel Visual Representations for Software Metrics Using 3D and Animation. In J. Münich and P. Liggesmeyer, editors, *Software Engineering 2009 - Workshopband (GI-Edition)*, Lecture Notes in Informatics (LNI), pages 147–154, Kaiserslautern, Germany, 2009. Gesellschaft für Informatik (GI). ISBN 9783885792444. URL http://urn.kb.se/resolve?urn=urn:nbn:se:vxu:diva-4228.

KyungTae Kim, Waqas Javed, Cary Williams, Niklas Elmqvist, and Pourang Irani. Hugin: A Framework for Awareness and Coordination in Mixed-Presence Collaborative Information Visualization. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces (ITS 2010)*, pages 231–240, Saarbrücken, Germany, 7–10 November 2010. Association for Computing Machinery (ACM). ISBN 9781450303996. doi:10.1145/1936652.1936694. URL https://doi.org/10.1145/1936652.1936694.

Oona Kivelä, Paula Alavesa, Aku Visuri, and Timo Ojala. Study on the Motivational and Physical Effects of Two VR Exergames. In *11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games 2019)*, pages 1–2, Vienna, Austria, 4–6 September 2019. Institute of Electrical and Electronics Engineers (IEEE). ISBN 978-1-7281-4540-2. doi:10.1109/VS-Games.2019.8864544. URL `https://doi.org/10.1109/VS-Games.2019.8864544`.

Kenneth W. Kolence. The Software Empiricist. *ACM SIGMETRICS Performance Evaluation Review*, 2(2): 31–36, June 1973. ISSN 0163-5999. doi:10.1145/1113644.1113647. URL `https://doi.org/10.1145/1113644.1113647`.

Kenneth W. Kolence and Philip J. Kiviat. Software Unit Profiles & Kiviat Figures. *ACM SIGMETRICS Performance Evaluation Review*, 2(3):2–12, September 1973. ISSN 0163-5999. doi:10.1145/1041613.1041614. URL `https://doi.org/10.1145/1041613.1041614`.

Panayiotis Koutsabasis and Panagiotis Vogiatzidakis. Empirical Research in Mid-Air Interaction: A Systematic Review. *International Journal of Human-Computer Interaction*, 35(18):1747–1768, 5 February 2019. doi:10.1080/10447318.2019.1572352. URL `https://doi.org/10.1080/10447318.2019.1572352`.

Matthias Kraus, Johannes Fuchs, Björn Sommer, Karsten Klein, Ulrich Engelke, Daniel Keim, and Falk Schreiber. Immersive Analytics with Abstract 3D Visualizations: A Survey. *Computer Graphics Forum*, 41(1):201–229, February 2022. doi:10.1111/cgf.14430. URL `https://doi.org/10.1111/cgf.14430`.

Steinar Kristoffersen and Fredrik Ljungberg. "Making Place" to Make IT work: Empirical Explorations of HCI for Mobile CSCW. In *Proceedings of the international ACM SIGGROUP conference on Supporting group work (GROUP 1999)*, pages 276–285, Phoenix, Arizona, USA, November 1999. Association for Computing Machinery (ACM). doi:10.1145/320297.320330. URL `https://doi.org/10.1145/320297.320330`.

James J. Kuch and Thomas S. Huang. Vision Based Hand Modeling and Tracking for Virtual Teleconferencing and Telecollaboration. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV 1995)*, pages 666–671, Cambridge, Massachusetts, USA, 20–23 June 1995. Institute of Electrical and Electronics Engineers (IEEE). ISBN 0-8186-7042-8. doi:10.1109/ICCV.1995.466875. URL `https://doi.org/10.1109/ICCV.1995.466875`.

Logan Kugler. The State of Virtual Reality Hardware. *Communications of the ACM*, 64(2):15–16, 25 January 2021. doi:10.1145/3441290. URL `https://doi.org/10.1145/3441290`.

Jérémy Lacoche, Nico Pallamin, Thomas Boggini, and Jérôme Royan. Collaborators Awareness for User Cohabitation in Co-located Collaborative Virtual Environments. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology (VRST 2017)*, pages 15:1–9, Gothenburg, Sweden, 8–10 November 2017. Association for Computing Machinery (ACM). ISBN 978-1-4503-5548-3. doi:10.1145/3139131.3139142. URL `https://doi.org/10.1145/3139131.3139142`.

Chengyuan Lai, Xinyu Hu, Afham Ahmed Aiyaz, Ann Segismundo, Ananya Phadke, and Ryan P. McMahan. The Cognitive Loads and Usability of Target-based and Steering-based Travel Techniques. *IEEE Transactions on Visualization and Computer Graphics*, 27(11):4289–4299, November 2021. doi:10.1109/TVCG.2021.3106507. URL `https://doi.org/10.1109/TVCG.2021.3106507`.

Mikko Laitinen, Jonas Lundberg, Magnus Levin, and Alexander Lakaw. Utilizing Multilingual Language Data in (Nearly) Real Time : The Case of the Nordic Tweet Stream. *Journal of Universal Computer Science (J.UCS)*, 23(11):1038–1056, 28 November 2017. doi:10.3217/jucs-023-11-1038. URL `https://doi.org/10.3217/jucs-023-11-1038`.

Mikko Laitinen, Jonas Lundberg, Magnus Levin, and Rafael Messias Martins. The Nordic Tweet Stream : A Dynamic Real-Time Monitor Corpus of Big and Rich Language Data. In *Proceedings of the Digital Humanities in the Nordic Countries 3rd Conference (DHN 2018)*, CEUR Workshop Proceedings, pages 349–362, Helsinki, Finland, 7-9 March 2018. Sun SITE Central Europe Workshop Proceedings (CEUR-WS). URL `http://urn.kb.se/resolve?urn=urn:nbn:se:lnu:diva-78277`.

Douglas Lanman, Henry Fuchs, Mine Mark, Ian McDowall, and Michael Abrash. Put on Your 3D Glasses Now: The Past, Present, and Future of Virtual and Augmented Reality. In *ACM SIGGRAPH 2014 Courses (SIGGRAPH 2014)*, pages 12:1–173, Vancouver, British Columbia, Canada, 10–14 August 2014. Association for Computing Machinery (ACM). ISBN 978-1-4503-2962-0. doi:10.1145/2614028.2628332. URL `https://doi.org/10.1145/2614028.2628332`.

Steven M. LaValle. *Virtual Reality*. Online, 2020. URL `http://lavalle.pl./vr/`.

Joseph J. LaViola, Jr. MSVT: A Virtual Reality-Based Multimodal Scientific Visualization Tool. In *Proceedings of the Third IASTED International Conference on Computer Graphics and Imaging*, page 7. International Association of Science and Technology for Development (IASTED), November 2000. URL `http://cs.brown.edu/people/jlaviola/pubs/msvt.pdf`.

Joseph J. LaViola, Jr., Ernst Kruijff, Ryan P. McMahan, Doug Bowman, and Ivan P. Poupyrev. *3D User Interfaces: Theory and Practice*. Addison-Wesley Professional, 2nd edition, 2017. ISBN 9780134034324. URL `https://www.pearson.com/us/higher-education/program/La-Viola-3-D-User-Interfaces-Theory-and-Practice-2nd-Edition/PGM101825.html`.

Bongshin Lee, Arjun Srinivasan, Petra Isenberg, and John Stasko. Post-WIMP Interaction for Information Visualization. *Foundations and Trends in Human–Computer Interaction*, 14(1):1–95, 11 June 2021. doi:10.1561/1100000081. URL `https://doi.org/10.1561/1100000081`.

Charlotte P. Lee and Drew Paine. From The Matrix to a Model of Coordinated Action (MoCA): A Conceptual Framework of and for CSCW. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW 2015)*, pages 179–194, Vancouver, British Columbia, Canada, 14–18 March 2015. Association for Computing Machinery (ACM). ISBN 978-1-4503-2922-4. doi:10.1145/2675133.2675161. URL `https://doi.org/10.1145/2675133.2675161`.

Jiwon Lee, Mingyu Kim, and Jinmo Kim. RoleVR: Multi-experience in immersive virtual reality between co-located HMD and non-HMD users. *Multimedia Tools and Applications*, 79:979–1005, 2020. doi:10.1007/s11042-019-08220-w. URL `https://doi.org/10.1007/s11042-019-08220-w`.

Georgios Lepouras. Comparing methods for numerical input in immersive virtual environments. *Virtual Reality*, 22:63–77, March 2018. doi:10.1007/s10055-017-0312-5. URL `https://doi.org/10.1007/s10055-017-0312-5`.

Rensis Likert. A Technique for the Measurement of Attitudes. *Archives of Psychology*, 22(140):1–55, June 1932. URL `https://psycnet.apa.org/record/1933-01885-001`.

Hangxin Liu, Zhenliang Zhang, Xu Xie, Yixin Zhu, Yue Liu, Yongtian Wang, and Song-Chun Zhu. High-Fidelity Grasping in Virtual Reality using a Glove-based System. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 5180–5186, Montreal, Quebec, Canada, 20–24 May 2019. Institute of Electrical and Electronics Engineers (IEEE). doi:10.1109/ICRA.2019.8794230. URL `https://doi.org/10.1109/ICRA.2019.8794230`.

Jean-Luc Lugrin, Fred Charles, Marc Cavazza, Marc Le Renard, Jonathan Freeman, and Jane Lessiter. CaveUDK: A VR Game Engine Middleware. In *Proceedings of the 18th ACM Symposium on Virtual Reality Software and Technology (VRST 2012)*, pages 137–144, Toronto, Ontario, Canada, 10–12 December 2012. Association for Computing Machinery (ACM). doi:10.1145/2407336.2407363. URL `https://doi.org/10.1145/2407336.2407363`.

Patrik Lundblad, Jonas Thoursie, and Mikael Jern. Swedish Road Weather Visualization. In *2010 14th International Conference Information Visualisation (IV)*, pages 313–321, London, UK, 26–29 July 2010. Institute of Electrical and Electronics Engineers (IEEE). doi:10.1109/IV.2010.51. URL `https://doi.org/10.1109/IV.2010.51`.

Jock Mackinlay, Stuart K. Card, and George G. Robertson. A Semantic Analysis of the Design Space of Input Devices. *Human-Computer Interaction*, 5(2–3):145–190, 1990. doi:10.1080/07370024.1990.9667153. URL `https://doi.org/10.1080/07370024.1990.9667153`.

G. Elisabeta Marai, Angus G. Forbes, and Andrew Johnson. Interdisciplinary Immersive Analytics at the Electronic Visualization Laboratory: Lessons Learned and Upcoming Challenges. In *2016 Workshop on Immersive Analytics (IA)*, pages 54–59, Greenville, South Carolina, USA, 20 March 2016. Institute of Electrical and Electronics Engineers (IEEE). doi:10.1109/IMMERSIVE.2016.7932384. URL `https://doi.org/10.1109/IMMERSIVE.2016.7932384`.

Kim Marriott, Jian Chen, Marcel Hlawatsch, Takayuki Itoh, Miguel A. Nacenta, Guido Reina, and Wolfgang Stuerzlinger. Immersive Analytics: Time to Reconsider the Value of 3D for Information Visualisation. In Kim Marriott, Falk Schreiber, Tim Dwyer, Karsten Klein, Nathalie Henry Riche, Takayuki Itoh, Wolfgang Stuerzlinger, and Bruce H. Thomas, editors, *Immersive Analytics*, Lecture Notes in Computer Science (LNCS, volume 11190), pages 25–55. Springer, Cham, first online edition, 16 October 2018. ISBN 978-3-030-01388-2. doi:10.1007/978-3-030-01388-2_2. URL `https://doi.org/10.1007/978-3-030-01388-2_2`.

Daniel Medeiros, Eduardo Cordeiro, Daniel Mendes, Maurício Sousa, Alberto Raposo, Alfredo Ferreira, and Joaquim Jorge. Effects of Speed and Transitions on Target-based Travel Techniques. In *Proceedings of the 22nd ACM Symposium on Virtual Reality Software and Technology (VRST 2016)*, pages 327—–328, Munich, Germany, 2–4 November 2016. Association for Computing Machinery (ACM). ISBN 978-1-4503-4491-3. doi:10.1145/2993369.2996348. URL `https://doi.org/10.1145/2993369.2996348`.

Paul Milgram and Fumio Kishino. A Taxonomy of Mixed Reality Visual Displays. *IEICE Transactions on Information Systems*, E77-D(12):1321–1329, 25 December 1994. ISSN 0916-8532. URL `https://search.ieice.org/bin/summary.php?id=e77-d_12_1321`.

Paul Milgram, Haruo Takemura, Akira Utsumi, and Fumio Kishino. Augmented Reality: A class of displays on the reality-virtuality continuum. In *Proceedings of Telemanipulator and Telepresence Technologies (SPIE Vol. 2351)*, pages 282–292, Boston, Massachusetts, USA, 21 December 1995. Society of Photo-Optical Instrumentation Engineers (SPIE). doi:10.1117/12.197321. URL `https://doi.org/10.1117/12.197321`.

Patrick Millais, Simon L. Jones, and Ryan Kelly. Exploring Data in Virtual Reality: Comparisons with 2D Data Visualizations. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems (CHI EA 2018)*, pages LBW007:1–6, Montreal, Quebec, Canada, 21–26 April 2018. Association for Computing Machinery (ACM). ISBN 978-1-4503-5621-3. doi:10.1145/3170427.3188537. URL `https://doi.org/10.1145/3170427.3188537`.

Andrew Moran, Vijay Gadepally, Matthew Hubbell, and Jeremy Kepner. Improving Big Data Visual Analytics with Interactive Virtual Reality. In *Proceedings of the 2015 IEEE High Performance Extreme Computing Conference (HPEC 2015)*, page 5, Waltham, Massachusetts, USA, 15–17 September 2015. Institute of Electrical and Electronics Engineers (IEEE). doi:10.1109/HPEC.2015.7322473. URL `https://doi.org/10.1109/HPEC.2015.7322473`.

Tamara Munzner. *Visualization Analysis & Design*. A K Peters Visualization Series. A K Peters/CRC Press, 1st edition, 26 November 2014. ISBN 9781466508910. URL `https://www.routledge.com/Visualization-Analysis-and-Design/Munzner/p/book/9781466508910`.

Carlo Navarra, Katerina Vrotsou, Tomasz Opach, Almar Joling, Julie Wilk, and Tina-Simone Neset. A progressive development of a visual analysis interface of climate-related VGI. *Environmental Earth Sciences*, 80:684:1–15, 3 October 2021. doi:10.1007/s12665-021-09948-1. URL `https://doi.org/10.1007/s12665-021-09948-1`.

Dennis C. Neale, John M. Carroll, and Mary Beth Rosson. Evaluating Computer-Supported Cooperative Work: Models and Frameworks. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work (CSCW 2004)*, pages 112–121, Chicago, Illinois, USA, 6–10 November 2004. Association for Computing Machinery (ACM). ISBN 978-1-58113-810-8. doi:10.1145/1031607.1031626. URL `https://doi.org/10.1145/1031607.1031626`.

Chrystopher L. Nehaniv, Kerstin Dautenhahn, Jens Kubacki, Martin Haegele, Christopher Parlitz, and Rachid Alami. A Methodological Approach relating the Classification of Gesture to Identification of Human Intent in the Context of Human-Robot Interaction. In *Proceedings of the 2005 IEEE International Workshop on Robot and Human Interactive Communication (ROMAN 2005)*, pages 371–377, Nashville, Tennessee, USA, 13–15 March 2005. Institute of Electrical and Electronics Engineers (IEEE). ISBN 0-7803-9274-4. doi:10.1109/ROMAN.2005.1513807. URL `https://doi.org/10.1109/ROMAN.2005.1513807`.

Tina-Simone Neset, Lotten, Wiréhn, Tomasz Opach, Erik Glaas, and Björn-Ola Linnér. Evaluation of indicators for agricultural vulnerability to climate change: The case of Swedish agriculture. *Ecological Indicators*, 105:571–580, October 2019. doi:10.1016/j.ecolind.2018.05.042. URL `https://doi.org/10.1016/j.ecolind.2018.05.042`.

Tina-Simone Neset, Julie Wilk, Sara Cruz, Marisa Graça, Jan Ketil Rød, Maaike J. Maarse, P. Wallin, and Lotta Andersson. Co-designing a citizen science climate service. *Climate Services*, 24:100273:1–10, December 2021. doi:10.1016/j.cliser.2021.100273. URL `https://doi.org/10.1016/j.cliser.2021.100273`.

Thomas Neumayr, Mirjam Augstein, and Hans-Christia Jetter. Phases and Success Criteria of Collaboration in the Ubiquity Era. In *Proceedings of the 15th International Conference on Advances in Mobile Computing & Multimedia (MoMM 2017)*, pages 201–205, Salzburg, Austria, 4–6 December 2017. Association for Computing Machinery (ACM). ISBN 978-1-4503-5300-7. doi:10.1145/3151848.3151874. URL `https://doi.org/10.1145/3151848.3151874`.

Thomas Neumayr, Hans-Christian Jetter, Mirjam Augstein, Judith Friedl, and Thomas Luger. Domino: A Descriptive Framework for Hybrid Collaboration and Coupling Styles in Partially Distributed Teams. *Proceedings of the ACM on Human-Computer Interaction (CSCW)*, 2(CSCW):128:1–24, November 2018. doi:10.1145/3274397. URL `https://doi.org/10.1145/3274397`.

Huyen Nguyen, Benjamin Ward, Ulrich Engelke, Bruce Thomas, and Tomasz Bednarz. Collaborative Data Analytics Using Virtual Reality. In *Proceedings of the 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (IEEE VR 2019)*, pages 1098–1099, Osaka, Japan, 23–27 March 2019. Institute of Electrical and Electronics Engineers (IEEE). doi:10.1109/VR.2019.8797845. URL `https://doi.org/10.1109/VR.2019.8797845`.

Thi Thuong Huyen Nguyen and Thierry Duval. A Survey of Communication and Awareness in Collaborative Virtual Environments. In *Proceedings of the International Workshop on Collaborative Virtual Environments (3DCVE 2014)*, pages 1–8, Minneapolis, Minnesota, USA, 30 March 2014. Institute of Electrical and Electronics Engineers (IEEE). ISBN 978-1-4799-5217-5. doi:10.1109/3DCVE.2014.7160928. URL `https://doi.org/10.1109/3DCVE.2014.7160928`.

Donald A. Norman. Natural User Interfaces Are Not Natural. *Interactions*, 17(3):6–10, May–June 2010. doi:10.1145/1744161.1744163. URL `https://doi.org/10.1145/1744161.1744163`.

Norwegian National Committee For Research Ethics in Science and Technology. *Guidelines For Research Ethics in Science and Technology*. The Norwegian National Research Ethics Committees, 2nd edition, June 2016. ISBN 9788276820751. URL https://www.forskningsetikk.no/en/guidelines/science-and-technology/guidelines-for-research-ethics-in-science-and-technology/.

Heather L. O'Brien, Paul Cairns, and Mark Hall. A practical approach to measuring user engagement with the refined User Engagement Scale (UES) and new UES short form. *International Journal of Human-Computer Studies*, 112:28–39, April 2018. doi:10.1016/j.ijhcs.2018.01.004. URL https://doi.org/10.1016/j.ijhcs.2018.01.004.

Manuel Olbrich, Holger Graf, Jens Keil, Rüdiger Gad, Steffen Bamfaste, and Frank Nicolini. Virtual Reality Based Space Operations – A Study of ESA's Potential for VR Based Training and Simulation. In Jessie Y. C. Chen and Gino Fragomeni, editors, *Virtual, Augmented and Mixed Reality: Interaction, Navigation, Visualization, Embodiment, and Simulation (VAMR 2018)*, Lecture Notes in Computer Science (LNCS, volume 10909), pages 438–451. Springer, Cham, 2 June 2018. ISBN 978-3-319-91581-4. doi:10.1007/978-3-319-91581-4_33. URL https://doi.org/10.1007/978-3-319-91581-4_33.

Noritaka Osawa, Kikuo Asai, and Yuji Y. Sugimoto. Immersive Graph Navigation Using Direct Manipulation and Gestures. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST 2000)*, pages 147–152, Seoul, Korea, 22–25 October 2000. Association for Computing Machinery (ACM). ISBN 978-1-58113-316-5. doi:10.1145/502390.502418. URL https://doi.org/10.1145/502390.502418.

Kaitlyn Michelle Ouverson, Carsten Scherber, Emily Oldham, and Stephen B. Gilbert. What Does "Asymmetric VR" Mean? A Directed Content Analysis of Co-Located Use of VR by Users on Reddit. *Frontiers in Virtual Reality*, 2:765881:1–18, 10 December 2021. doi:10.3389/frvir.2021.765881. URL https://doi.org/10.3389/frvir.2021.765881.

Solomon Sunday Oyelere, Nacir Bouali, Rogers Kaliisa, George Obaido, Abdullahi Abubakar Yunusa, and Ebunayo R. Jimoh. Exploring the trends of educational virtual reality games: a systematic review of empirical studies. *Smart Learning Environments*, 7:31:1–22, 19 October 2020. doi:10.1186/s40561-020-00142-7. URL https://doi.org/10.1186/s40561-020-00142-7.

Minna Pakanen, Paula Alavesa, Niels van Berkel, Timo Koskela, and Timo Ojala. "Nice to see you virtually": Thoughtful design and evaluation of virtual avatar of the other user in AR and VR based telexistence systems. *Entertainment Computing*, 40:100457:1–16, January 2022. doi:10.1016/j.entcom.2021.100457. URL https://doi.org/10.1016/j.entcom.2021.100457.

Tapan Parikh, Sara Perl Egendorf, Isaiah Murray, Ali Jamali, Bryan Yee, Sammi Lin, Kendra Cooper-Smith, Brendan Parker, Koron Smiley, and Jenny Kao-Kniffin. Greening the Virtual Smart City: Accelerating Peer-to-Peer Learning in Urban Agriculture With Virtual Reality Environments. *Frontiers in Sustainable Cities*, 3:815937:1–7, 3 February 2022. doi:10.3389/frsc.2021.815937. URL https://doi.org/10.3389/frsc.2021.815937.

Byungho Park. Psychophysiology as a Tool for HCI Research: Promises and Pitfalls. In Julie A. Jacko, editor, *Human-Computer Interaction. New Trends. 13th International Conference, HCI International 2009, San Diego, CA, USA, July 19-24, 2009, Proceedings, Part I*, Lecture Notes in Computer Science (LNCS, volume 5610), pages 141–148. Springer, Berlin, Heidelberg, 1st edition, 2009. ISBN 978-3-642-02574-7. doi:10.1007/978-3-642-02574-7_16. URL https://doi.org/10.1007/978-3-642-02574-7_16.

Vladimir I. Pavlovic, Rajeev Sharma, and Thomas S. Huang. Visual Interpretation of Hand Gestures for Human-Computer Interaction: A Review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):677–695, July 1997. doi:10.1109/34.598226. URL https://doi.org/10.1109/34.598226.

Tekla S. Perry. Virtual Reality Goes Social. *IEEE Spectrum*, 53(1):56–57, January 2016. doi:10.1109/MSPEC.2016.7367470. URL `https://doi.org/10.1109/MSPEC.2016.7367470`.

Mark Peter, Robin Horst, and Ralf Dörner. VR-Guide: A Specific User Role for Aysmmetric Virtual Reality Setups in Distributed Virtual Reality Applications. In *Tagungsband 15. Workshop der GI-Fachgruppe VR/AR*, pages 83–94, Düsseldorf, Germany, 2018. Gesellschaft für Informatik (GI).

Stefan Peters. *Dynamics of spatially extended phenomena - Visual analytical approach to movements of lightning clusters*. PhD thesis, Technische Universität München, München, Germany, September 2014. URL `http://publica.fraunhofer.de/documents/N-419044.html`.

David Pinelle, Carl Gutwin, and Saul Greenberg. Task Analysis for Groupware Usability Evaluation: Modeling Shared-Workspace Tasks with the Mechanics of Collaboration. *ACM Transactions on Computer-Human Interaction*, 10(4):281–311, December 2003. doi:10.1145/966930.966932. URL `https://doi.org/10.1145/966930.966932`.

Márcio S. Pinho, Doug A. Bowman, and Carla M.D.S. Freitas. Cooperative Object Manipulation in Immersive Virtual Environments: Framework and Techniques. In *Proceedings of the ACM symposium on Virtual reality software and technology (VRST 2002)*, pages 171–178, Hong Kong, China, 11–13 November 2002. Association for Computing Machinery (ACM). ISBN 1-58113-530-0. doi:10.1145/585740.585769. URL `https://doi.org/10.1145/585740.585769`.

Sebastian Pirch, Felix Müller, Eugenia Iofinova, Julia Pazmandi, Christiane V. R. Hütter, Martin Chiettini, Celine Sin, Kaan Boztug, Iana Podkosova, Hannes Kaufmann, and Jörg Menche. The VRNetzer platform enables interactive network analysis in Virtual Reality. *Nature Communications*, 12:2432:1–14, 23 April 2021. doi:10.1038/s41467-021-22570-w. URL `https://doi.org/10.1038/s41467-021-22570-w`.

Karolin Poels, Yvonne A. W. de Kort, and Wijnand A. IJsselsteijn. D3.3 : Game Experience Questionnaire: development of a self-report measure to assess the psychological impact of digital games. Technical report, Technische Universiteit Eindhoven, 1 January 2007. URL `https://research.tue.nl/en/publications/d33-game-experience-questionnaire-development-of-a-self-report-me`.

Matti Pouke, Johanna Ylipulli, Ilya Minyaev, Minna Pakanen, Paula Alavesa, Toni Alatalo, and Timo Ojala. Virtual Library – Blending Mirror and Fantasy Layers into a VR Interface for a Public Library. In *Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia (MUM 2018)*, pages 227–231, Cairo, Egypt, 25–28 November 2018. Association for Computing Machinery (ACM). ISBN 978-1-4503-6594-9. doi:10.1145/3282894.3282927. URL `https://doi.org/10.1145/3282894.3282927`.

Jaziar Radianti, Tim A. Majchrzak, Jennifer Fromm, and Isabell Wohlgenannt. A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education*, 147:103778:1–29, 9 December 2020. doi:10.1016/j.compedu.2019.103778. URL `https://doi.org/10.1016/j.compedu.2019.103778`.

Eric D. Ragan, Andrew Wood, Ryan P. McMahan, and Doug A. Bowman. Trade-Offs Related to Travel Techniques and Level of Display Fidelity in Virtual Data-Analysis Environments. In *Proceedings of the Joint Virtual Reality Conference of ICAT – EGVE – EuroVR (JVRC 2012)*, pages 81–84, Madrid, Spain, 16–19 October 2012. The Eurographics Association. ISBN 978-3-905674-40-8. doi:10.2312/EGVE/JVRC12/081-084. URL `https://doi.org/10.2312/EGVE/JVRC12/081-084`.

Lisa Rebenitsch and Charles Owen. Review on cybersickness in applications and visual displays. *Virtual Reality*, 20(1):101–125, 2016. doi:10.1007/s10055-016-0285-9. URL `https://doi.org/10.1007/s10055-016-0285-9`.

Khairi Reda, Alessandro Febretti, Aaron Knoll, Jillian Aurisano, Jason Leigh, Andrew Johnson, Michael E. Papka, and Mark Hereld. Visualizing Large, Heterogeneous Data in Hybrid-Reality Environments. *IEEE Computer Graphics and Applications*, 33(4):38–48, July–August 2013. doi:10.1109/MCG.2013.37. URL `https://doi.org/10.1109/MCG.2013.37`.

David Rempel, Matt J. Camilleri, and David L. Lee. The design of hand gestures for Human–Computer Interaction: Lessons from sign language interpreters. *International Journal of Human-Computer Studies*, 72(10–11):728–735, October–November 2014. doi:10.1016/j.ijhcs.2014.05.003. URL `https://doi.org/10.1016/j.ijhcs.2014.05.003`.

Nico Reski and Aris Alissandrakis. Open data exploration in virtual reality: a comparative study of input technology. *Virtual Reality*, 24:1–22, March 2020. ISSN 1434-9957. doi:10.1007/s10055-019-00378-w. URL `https://doi.org/10.1007/s10055-019-00378-w`.

Nico Reski, Aris Alissadrakis, and Jukka Tyrkkö. Collaborative exploration of rich corpus data using immersive virtual reality and non-immersive technologies. In *2nd International Conference: Approaches to Digital Discourse Analysis (ADDA 2)*, ADDA 2 Book of Abstracts, pages 7–9, Turku, Finland, 23–25 May 2019. University of Turku. URL `http://urn.kb.se/resolve?urn=urn:nbn:se:lnu:diva-83858`.

Nico Reski, Aris Alissandrakis, and Andreas Kerren. Exploration of Time-Oriented Data in Immersive Virtual Reality Using a 3D Radar Chart Approach. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society (NordiCHI 2020)*, pages 33:1–11, Tallinn, Estonia, 25–29 October 2020a. Association for Computing Machinery (ACM). ISBN 9781450375795. doi:10.1145/3419249.3420171. URL `https://doi.org/10.1145/3419249.3420171`.

Nico Reski, Aris Alissandrakis, Jukka Tyrkkö, and Andreas Kerren. "Oh, that's where you are!" – Towards a Hybrid Asymmetric Collaborative Immersive Analytics System. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society (NordiCHI 2020)*, pages 5:1–12, Tallinn, Estonia, 25–29 October 2020b. Association for Computing Machinery (ACM). ISBN 9781450375795. doi:10.1145/3419249.3420102. URL `https://doi.org/10.1145/3419249.3420102`.

Nico Reski, Aris Alissandrakis, and Andreas Kerren. An Empirical Evaluation of Asymmetric Synchronous Collaboration Combining Immersive and Non-Immersive Interfaces Within the Context of Immersive Analytics. *Frontiers in Virtual Reality*, 2:743445:1–29, 17 January 2022. ISSN 2673-4192. doi:10.3389/frvir.2021.743445. URL `https://doi.org/10.3389/frvir.2021.743445`.

Falko Rheinberg. Intrinsische Motivation und Flow-Erleben [Intrinsic motivation and flow experience]. In Jutta Heckhausen and Heinz Heckhausen, editors, *Motivation und Handeln*, Springer Lehrbuch (SLB), pages 365–387. Springer, Berlin, Heidelberg, 4th edition, 2010. ISBN 978-3-642-12693-2. doi:10.1007/978-3-642-12693-2_14. URL `https://doi.org/10.1007/978-3-642-12693-2_14`.

Falko Rheinberg, Regina Vollmeyer, and Stefan Engeser. Die Erfassung des Flow-Erlebens [The assessment of flow experience]. In Joachim Stiensmeier-Pelster and Falko Rheinberg, editors, *Diagnostik von Motivation und Selbstkonzept [Diagnostics of motivation and self-concept]*, pages 261–279. Hogrefe, Göttingen, Germany, 1st edition, 2003. ISBN 9783801716745. URL `https://www.hogrefe.de/shop/diagnostik-von-motivation-und-selbstkonzept-65344.html`.

Jonathan C. Roberts, Panagiotis D. Ritsos, Sriram Karthik Badam, Dominique Brodbeck, Jessie Kennedy, and Niklas Elmqvist. Visualization beyond the Desktop–the Next Big Thing. *IEEE Computer Graphics and Applications*, 34(6):26–34, November–December 2014. doi:10.1109/MCG.2014.82. URL `https://doi.org/10.1109/MCG.2014.82`.

Katja Rogers, Sukran Karaosmanoglu, Dennis Wolf, Frank Steinicke, and Lennart E. Nacke. A Best-Fit Framework and Systematic Review of Asymmetric Gameplay in Multiplayer Virtual Reality Games. *Frontiers in Virtual Reality*, 2:694660:1–18, 27 July 2021. doi:10.3389/frvir.2021.694660. URL `https://doi.org/10.3389/frvir.2021.694660`.

Jose Luis Rubio-Tamayo, Manuel Gertrudix Barrio, and Francisco García García. Immersive Environments and Virtual Reality: Systematic Review and Advances in Communication, Interaction and Simulation. *Multimodal Technologies and Interaction*, 1(4):21:1–20, 27 September 2017. doi:10.3390/mti1040021. URL `https://doi.org/10.3390/mti1040021`.

Franca A. Rupprecht, Taimur K. Khan, Gerrit van der Veer, and Achim Ebert. Criteria Catalogue for Collaborative Environments. In *Proceedings of the 31st International BCS Human Computer Interaction Conference (HCI 2017)*, pages 165–171, Sunderland, UK, 3–6 July 2017. BCS Learning and Development. doi:10.14236/ewic/HCI2017.4. URL `https://doi.org/10.14236/ewic/HCI2017.4`.

Jan Ketil Rød and Maaike J. Maarse. Using Citizen Sensing to Identify Heat-Exposed Neighbourhoods. *Urban Science*, 5(1):14:1–14, 2021. doi:10.3390/urbansci5010014. URL `https://doi.org/10.3390/urbansci5010014`.

W. J. Sarmiento, Anderson Maciel, Luciana Nedel, and Cesar A. Collazos. Measuring the Collaboration Degree in Immersive 3D Collaborative Virtual Environments. In *Proceedings of the International Workshop on Collaborative Virtual Environments (3DCVE 2014)*, pages 23–28, Minneapolis, Minnesota, USA, 30 March 2014. Institute of Electrical and Electronics Engineers (IEEE). ISBN 978-1-4799-5217-5. doi:10.1109/3DCVE.2014.7160931. URL `https://doi.org/10.1109/3DCVE.2014.7160931`.

Jeff Sauro and James R. Lewis. Standardized Usability Questionnaires. In *Quantifying the User Experience: Practical Statistics for User Research*, pages 185–240. Morgan Kaufmann, first online edition, 29 February 2012. ISBN 978-0-12-384968-7. doi:10.1016/B978-0-12-384968-7.00008-4. URL `https://doi.org/10.1016/B978-0-12-384968-7.00008-4`.

Kjeld Schmidt. The Problem with "Awareness": Introductory Remarks on "Awareness in CSCW". *Computer Supported Cooperative Work (CSCW)*, 11:285–298, September 2002. doi:10.1023/A:1021272909573. URL `https://doi.org/10.1023/A:1021272909573`.

Ralf Schmälzle and Clare Grall. Psychophysiological Methods: Options, Uses, and Validity. In Jan Van den Bulck, David R. Ewoldsen, Marie-Louise Mares, and Erica Scharrer, editors, *The International Encyclopedia of Media Psychology*, page 8. John Wiley & Sons, Inc., 2020. ISBN 978-1-4471-0685-2. doi:10.1002/9781119011071.iemp0013. URL `https://doi.org/10.1002/9781119011071.iemp0013`.

Ben Shneiderman. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In *Proceedings 1996 IEEE Symposium on Visual Languages*, pages 336–343, Boulder, Colorado, USA, 3–6 September 1996. Institute of Electrical and Electronics Engineers (IEEE). doi:10.1109/VL.1996.545307. URL `https://doi.org/10.1109/VL.1996.545307`.

Ben Shneiderman, Catherine Plaisant, Maxine Cohen, Steven Jacobs, Niklas Elmqvist, and Nicholas Diakopoulos. *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. Human-Computer Interaction Series (HCIS). Pearson, 6th edition, 20 June 2017. ISBN 9781292153919.

Ronell Sicat, Jiabao Li, JunYoung Choi, Maxime Cordeil, Won-Ki Jeong, Benjamin Bach, and Hanspeter Pfister. DXR: A Toolkit for Building Immersive Data Visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):715–725, January 2019. doi:10.1109/TVCG.2018.2865152. URL `https://doi.org/10.1109/TVCG.2018.2865152`.

Richard Skarbez, Nicholas F. Polys, J. Todd Oggle, Chris North, and Doug A. Bowman. Immersive Analytics: Theory and Research Agenda. *Frontiers in Robotics and AI*, 6:82:1–15, 10 September 2019. doi:10.3389/frobt.2019.00082. URL `https://doi.org/10.3389/frobt.2019.00082`.

Richard Skarbez, Missie Smith, and Mary C. Whitton. Revisiting Milgram and Kishino's Reality-Virtuality Continnum. *Frontiers in Virtual Reality*, 2:647997:1–8, 24 March 2021. doi:10.3389/frvir.2021.647997. URL https://doi.org/10.3389/frvir.2021.647997.

Mel Slater. A Note on Presence Terminology. *Presence Connect*, 3(3):5, 10 September 2003. URL http://publicationslist.org/melslater.

Mel Slater and Martin Usoh. Presence in Immersive Virtual Environments. In *Proceedings of IEEE Virtual Reality Annual International Symposium*, pages 90–96, Seattle, Washington, USA, 18–22 September 1993. Institute of Electrical and Electronics Engineers (IEEE). doi:10.1109/VRAIS.1993.380793. URL https://doi.org/10.1109/VRAIS.1993.380793.

Mel Slater and Sylvia Wilbur. A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, 6 (6):603–616, 1997. doi:10.1162/pres.1997.6.6.603. URL https://doi.org/10.1162/pres.1997.6.6.603.

Dave Snowdon, Elisabeth F. Churchill, and Alan J. Munro. Collaborative Virtual Environments: Digital Spaces and Places for CSCW: An Introduction. In Elisabeth F. Churchill, David N. Snowdon, and Alan J. Munro, editors, *Collaborative Virtual Environments: Digital Places and Spaces for Interaction*, Computer Supported Cooperative Work (CSCW), pages 3–17. Springer, London, 2001. ISBN 978-1-4471-0685-2. doi:10.1007/978-1-4471-0685-2_1. URL https://doi.org/10.1007/978-1-4471-0685-2_1.

Björn Sommer, David G. Barnes, Sarah Boyd, Thomas Chandler, Maxime Cordeil, Tobias Czauderna, Mathias Klapperstück, Karsten Klein, Toan D. Nguyen, Hieu Nim, Kingsley Stephens, Dany Vohl, Stephen Wang, Elliott Wilson, Yan Zhu, Jian Li, Jon McCormack, Kim Mariott, and Falk Schreiber. 3D-Stereoscopic Immersive Analytics Projects at Monash University and University of Konstanz. *Electronic Imaging*, 2017(5):179–187, 29 January 2017. doi:10.2352/ISSN.2470-1173.2017.5.SDA-109. URL https://doi.org/10.2352/ISSN.2470-1173.2017.5.SDA-109.

Misha Sra, Aske Mottelson, and Pattie Maes. Your Place and Mine: Designing a Shared VR Experience for Remotely Located Users. In *Proceedings of the 2018 Designing Interactive Systems Conference (DIS 2018)*, pages 85–97, Hong Kong, China, 9–13 June 2018. Association for Computing Machinery (ACM). doi:10.1145/3196709.3196788. URL https://doi.org/10.1145/3196709.3196788.

Aaron Stafford, Wayne Piekarski, and Bruce H. Thomas. Implementation of God-like Interaction Techniques for Supporting Collaboration Between Outdoor AR and Indoor Tabletop Users. In *Proceedings of The Fifth IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR 2006)*, pages 165–172, Santa Barbara, California, USA, 22–25 October 2006. Institute of Electrical and Electronics Engineers (IEEE). ISBN 1-4244-0650-1. doi:10.1109/ISMAR.2006.297809. URL https://doi.org/10.1109/ISMAR.2006.297809.

Kay M. Stanney, Ronald R. Mourant, and Robert S. Kennedy. Human Factors Issues in Virtual Environments: A Review of the Literature. *Presence: Teleoperators and Virtual Environments*, 7(4): 327–351, 1998. doi:10.1162/105474698565767. URL https://doi.org/10.1162/105474698565767.

Jan-Philipp Stauffert, Florian Niebling, and Marc Erich Latoschik. Latency and Cybersickness: Impact, Causes, and Measures. A Review. *Frontiers in Virtual Reality*, 1:582204:1–10, 26 November 2020. doi:10.3389/frvir.2020.582204. URL https://doi.org/10.3389/frvir.2020.582204.

Anthony Steed and Ralph Schroeder. Collaboration in Immersive and Non-immersive Virtual Environments. In Matthew Lombard, Frank Biocca, Jonathan Freeman, Wijnand IJsselsteijn, and Rachel J. Schaevitz, editors, *Immersed in Media: Telepresence Theory, Measurement & Technology*, pages 263–282. Springer, Cham, 1st edition, 2015. ISBN 978-3-319-10190-3. doi:10.1007/978-3-319-10190-3_11. URL https://doi.org/10.1007/978-3-319-10190-3_11.

Frank Steinicke. *Being Really Virtual: Immersive Natives and the Future of Virtual Reality*. Springer International Publishing Switzerland, 1st edition, 2016. ISBN 978-3-319-43078-2. doi:10.1007/978-3-319-43078-2. URL https://doi.org/10.1007/978-3-319-43078-2.

Frank Steinicke, Gerd Bruder, and Scott Kuhl. Realistic Perspective Projections for Virtual Objects and Environments. *ACM Transactions on Graphics*, 30(5):112:1–10, October 2011. doi:10.1145/2019627.2019631. URL https://doi.org/10.1145/2019627.2019631.

Sophie Stellmach and Raimund Dachselt. Look & Touch: Gaze-supported Target Acquisition. In *Proceedings of the 2012 SIGCHI Conference on Human Factors in Computing Systems (CHI 2012)*, pages 2981–2990, Austin, Texas, USA, 5–10 May 2012. Association for Computing Machinery (ACM). doi:10.1145/2207676.2208709. URL https://doi.org/10.1145/2207676.2208709.

Neal Stephenson. *Snow Crash*. Bantam Books, 1st edition, June 1992. ISBN 9780553088533.

Dirk Streeb, Mennatallah El-Assady, Daniel A. Keim, and Min Chen. Why Visualize? Arguments for Visual Support in Decision Making. *IEEE Computer Graphics and Applications*, 41(2):17–22, March–April 2021. doi:10.1109/MCG.2021.3055971. URL https://doi.org/10.1109/MCG.2021.3055971.

Barbara Streppel, Dorothea Pantförder, and Birgit Vogel-Heuser. Interaction in Virtual Environments – How to Control the Environment by Using VR-Glasses in the Most Immersive Way. In Jessie Y. C. Chen and Gino Fragomeni, editors, *Virtual, Augmented and Mixed Reality: Interaction, Navigation, Visualization, Embodiment, and Simulation (VAMR 2018)*, Lecture Notes in Computer Science (LNCS, volume 10909), pages 183–201. Springer, Cham, 2 June 2018. ISBN 978-3-319-91580-7. doi:10.1007/978-3-319-91581-4_14. URL https://doi.org/10.1007/978-3-319-91581-4_14.

Yuta Sugiura, Hikaru Ibayashi, Toby Chong, Daisuke Sakamoto, Natsuki Miyata, Mitsunori Tada, Takashi Okuma, Takeshi Kurata, Takashi Shinmura, Masaaki Mochimaru, and Takeo Igarashi. An Asymmetric Collaborative System for Architectural-scale Space Design. In *Proceedings of the 16th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry (VRCAI 2018)*, pages 1–6, Tokyo, Japan, December 2018. Association for Computing Machinery (ACM). ISBN 978-1-4503-6087-6. doi:10.1145/3284398.3284416. URL https://doi.org/10.1145/3284398.3284416.

Yilu Sun, Omar Shaikh, and Andrea Stevenson Won. Nonverbal synchrony in virtual reality. *PLoS ONE*, 14(9):e0221803:1–28, 16 September 2019. doi:10.1371/journal.pone.0221803. URL https://doi.org/10.1371/journal.pone.0221803.

Ivan E. Sutherland. A head-mounted three dimensional display. In *Proceedings of the December 9-11, 1968, fall joint computer conference, part I (AFIPS '68, Fall, part I)*, pages 757–764, San Francisco, California, USA, 9–11 December 1968. Association for Computing Machinery (ACM). doi:10.1145/1476589.1476686. URL https://doi.org/10.1145/1476589.1476686.

Swedish Research Council. *Good Research Practice*. Swedish Research Council, Vetenskapsrådet, Stockholm, Sweden, 2nd edition, 12 June 2017. ISBN 9789173073547. URL https://www.vr.se/english/analysis/reports/our-reports/2017-08-31-good-research-practice.html.

Anthony Tang, Melanie Tory, Barry Po, Petra Neumann, and Sheelagh Carpendale. Collaborative Coupling over Tabletop Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2006)*, pages 1181–1190, Montreal, Quebec, Canada, 22–27 April 2006. Association for Computing Machinery (ACM). ISBN 1-59593-372-7. doi:10.1145/1124772.1124950. URL https://doi.org/10.1145/1124772.1124950.

Bruce H. Thomas, Gregory F. Welch, Pierre Dragicevic, Niklas Elmqvist, Pourang Irani, Yvonne Jansenand Dieter Schmalstieg, Aurélien Tabard, Neven A. M. ElSayed, Ross T. Smith, and Wesley Willett. Situated Analytics. In Kim Marriott, Falk Schreiber, Tim Dwyer, Karsten Klein, Nathalie Henry Riche, Takayuki Itoh, Wolfgang Stuerzlinger, and Bruce H. Thomas, editors, *Immersive Analytics*, Lecture Notes in Computer Science (LNCS, volume 11190), pages 185–220. Springer, Cham, first online edition, 16 October 2018. ISBN 978-3-030-01388-2. doi:10.1007/978-3-030-01388-2_7. URL https://doi.org/10.1007/978-3-030-01388-2_7.

Lui Albæk Thomsen, Niels Christian Nilsson, Rolf Nordahl, and Boris Lohmann. Asymmetric collaboration in virtual reality: A taxonomy of asymmetric interfaces for collaborative immersive learning. *Tidsskriftet Læring Og Medier (LOM)*, 12(20):28, 2019. doi:10.7146/lom.v12i20.109391. URL https://doi.org/10.7146/lom.v12i20.109391.

Christian Tominski, James Abello, and Heidrun Schumann. Interactive Poster: 3D Axes-Based Visualizations for Time Series Data. In *Poster Compendium of IEEE Symposium on Information Visualization (InfoVis 2005)*, pages 49–50, Minneapolis, Minnesota, USA, 23–25 October 2005. Institute of Electrical and Electronics Engineers (IEEE). URL https://vcg.informatik.uni-rostock.de/~ct/pub_files/Tominski05AxesBasedVis3D.pdf.

Jukka Tyrkkö, Mikko Laitinen, and Magnus Levin. Actually in Nordic tweets. *World Englishes*, 40(4): 631–649, 14 May 2021. doi:10.1111/weng.12545. URL https://doi.org/10.1111/weng.12545.

Amy C. Ulinski, Zachary Wartell, Paula Goolkasian, Evan A. Suma, and Larry F. Hodges. Selection Performance Based on Classes of Bimanual Actions. In *Proceedings of the IEEE Symposium on 3D User Interfaces (3DUI 2009)*, pages 51–58, Lafayette, Louisiana, USA, 14–25 March 2009. Institute of Electrical and Electronics Engineers (IEEE). ISBN 978-1-4244-3965-2. doi:10.1109/3DUI.2009.4811205. URL https://doi.org/10.1109/3DUI.2009.4811205.

D. W. F. van Krevelen and Ronald Poelman. A Survey of Augmented Reality Technologies, Applications and Limitations. *International Journal of Virtual Reality*, 9(2):1–20, 1January 2010. doi:10.20870/IJVR.2010.9.2.2767. URL https://doi.org/10.20870/IJVR.2010.9.2.2767.

Michael A. Vidulich and Pamela S. Tsang. The Confluence of Situation Awareness and Mental Workload for Adaptable Human–Machine Systems. *Journal of Cognitive Engineering and Decision Making*, 9(1): 95–97, 2015. doi:10.1177/1555343414554805. URL https://doi.org/10.1177/1555343414554805.

Milena Vuckovic, Johanna Schmidt, Thomas Ortner, and Daniel Cornel. Combining 2D and 3D Visualization with Visual Analytics in the Environmental Domain. *Information*, 13(1):7:1–21, 2022. doi:10.3390/info13010007. URL https://doi.org/10.3390/info13010007.

Jorge A. Wagner, Wolfgang Stuerzlinger, and Luciana Nedel. Comparing and Combining Virtual Hand and Virtual Ray Pointer Interactions for Data Manipulation in Immersive Analytics. *IEEE Transactions on Visualization and Computer Graphics*, 27(5):2513–2523, May 2021. doi:10.1109/TVCG.2021.3067759. URL https://doi.org/10.1109/TVCG.2021.3067759.

Jorge A. Wagner Filho, Marina F. Rey, Carla M. D. S. Freitas, and Luciana Nedel. Immersive Visualization of Abstract Information: An Evaluation on Dimensionally-Reduced Data Scatterplots. In *Proceedings of the 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 483–490, Reutlingen, Germany, 18–22 March 2018. Institute of Electrical and Electronics Engineers (IEEE). doi:10.1109/VR.2018.8447558. URL https://doi.org/10.1109/VR.2018.8447558.

Jorge A. Wagner Filho, Wolfgang Stuerzlinger, and Luciana Nedel. Evaluating an Immersive Space-Time Cube Geovisualization for Intuitive Trajectory Data Exploration. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):514–524, 1 January 2020. doi:10.1109/TVCG.2019.2934415. URL https://doi.org/10.1109/TVCG.2019.2934415.

Mattias Wallergård, Joakim Eriksson, and Günter Alce. *Virtual Reality — En Introduktion [Virtual Reality – An Introduction]*. Studentlitteratur, 1st edition, 2022. ISBN 9789144122281. URL https://www.studentlitteratur.se/kurslitteratur/teknik-datorer-it-och-bygg/data-allmant/virtual-reality/.

Ray Chen Wang. *Geometric Fault Detection Using 3D Kiviat Plots and their Applications*. PhD thesis, The University of Texas at Austin, Austin, Texas, USA, May 2017. URL http://hdl.handle.net/2152/47439.

Xiyao Wang, Lonni Besançon, Florimond Gueéniat, Mickael Sereno, Mehdi Ammi, and Tobias Isenberg. A Vision of Bringing Immersive Visualization to Scientific Workflows. In *Proceedings of the 2019 ACM Conference on Human Factors in Computing Systems (CHI) - Workshop on Interaction Design & Prototyping for Immersive Analytics*, page 8, Glasgow, Scotland, UK, 4–9 May 2019. Association for Computing Machinery (ACM). URL https://hal.archives-ouvertes.fr/hal-02053969.

Matthew Ward, Amit Barde, Paul N. Russell, Mark Billinghurst, and William S. Helton. Visual Cues to Reorient Attention from Head Mounted Displays. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 60(1):1574–1578, 8 September 2016. doi:10.1177/1541931213601363. URL https://doi.org/10.1177/1541931213601363.

Matthew O. Ward, Georges Grinstein, and Daniel Keim. *Interactive Data Visualization: Foundations, Techniques, and Applications*. A K Peters/CRC Press, 2nd edition, 2015. ISBN 9781482257373. URL https://www.crcpress.com/Interactive-Data-Visualization-Foundations-Techniques-and-Applications/Ward-Grinstein-Keim/p/book/9781482257373.

Finn Welsford-Ackroyd, Andrew Chalmers, Rafael K. dos Anjos, Daniel Medeiros, Hyejin Kim, and Taehyun Rhee. Asymmetric Interaction between HMD Wearers and Spectators with a Large Display. In *Poster session at The 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (IEEE VR 2020)*, page 2, Atlanta, Georgia, USA, 22–25 March 2020. Institute of Electrical and Electronics Engineers (IEEE). doi:10.13140/RG.2.2.34994.56006. URL https://doi.org/10.13140/RG.2.2.34994.56006.

Christopher D. Wickens and C. Melody Carswell. Information Processing. In Gavriel Salvendy and Waldemar Karwowski, editors, *Handbook of Human Factors and Ergonomics*, pages 114–158. John Wiley & Sons, Inc., 5th edition, 13 August 2021. ISBN 9781119636113. doi:10.1002/9781119636113.ch5. URL https://doi.org/10.1002/9781119636113.ch5.

Josef Wideström, Ann-Sofie Axelsson, Ralph Schroeder, Alexander Nilsson, Ilona Heldal, and Åsa Abelin. The Collaborative Cube Puzzle: A Comparison of Virtual and Real Environments. In *Proceedings of the 3rd international conference on Collaborative virtual environments (CVE 2000)*, pages 165–171, San Francisco, California, USA, 10–12 September 2000. Association for Computing Machinery (ACM). ISBN 1-58113-303-0. doi:10.1145/351006.351035. URL https://doi.org/10.1145/351006.351035.

Christopher K. Wikle, Andrew Zammit-Mangion, and Noel Cressie. *Spatio-Temporal Statistics with R*. The R Series. Chapman & Hall/CRC Press, 2019. ISBN 978-1-138-71113-6. URL https://spacetimewithr.org.

Wesley Willett and Petra Isenberg. Understanding Digital Note-Taking Practice for Visualization. *IEEE Computer Graphics and Applications*, 35(4):38–51, July–August 2015. doi:10.1109/MCG.2015.52. URL https://doi.org/10.1109/MCG.2015.52.

Markus Wirth, Stefan Gradl, Jan Sembdner, Soeren Kuhrt, and Bjoern M. Eskofier. Evaluation of Interaction Techniques for a Virtual Reality Reading Room in Diagnostic Radiology. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology (UIST 2018)*, pages 867–876, Berlin, Germany, 14–17 October 2018. Association for Computing Machinery (ACM). ISBN 978-1-4503-5948-1. doi:10.1145/3242587.3242636. URL https://doi.org/10.1145/3242587.3242636.

Isabel Wohlgenannt, Alexander Simons, and Stefan Stieglitz. Virtual Reality. *Business & Information Systems Engineering*, 62:455—-461, 2 July 2020. doi:10.1007/s12599-020-00658-9. URL https://doi.org/10.1007/s12599-020-00658-9.

Katrin Wolf, Markus Funk, Rami Khalil, and Pascal Knierim. Using Virtual Reality for Prototyping Interactive Architecture. In *Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia (MUM 2017)*, pages 457–464, Stuttgart, Germany, 26–29 November 2017. Association for Computing Machinery (ACM). ISBN 978-1-4503-5378-6. doi:10.1145/3152832.3156625. URL `https://doi.org/10.1145/3152832.3156625`.

Yuanjie Wu, Yu Wang, Sungchul Jung, Simon Hoermann, and Robert W. Lindeman. Using a Fully Expressive Avatar to Collaborate in Virtual Reality: Evaluation of Task Performance, Presence, and Attraction. *Frontiers in Virtual Reality*, 2:641296:1–15, 7 April 2021. doi:10.3389/frvir.2021.641296. URL `https://doi.org/10.3389/frvir.2021.641296`.

Haijun Xia, Sebastian Herscher, Ken Perlin, and Daniel Wigdor. Spacetime: Enabling Fluid Individual and Collaborative Editing in Virtual Reality. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology (UIST 2018)*, pages 853–866, Berlin, Germany, 14–17 October 2018. Association for Computing Machinery (ACM). ISBN 978-1-4503-5948-1. doi:10.1145/3242587.3242597. URL `https://doi.org/10.1145/3242587.3242597`.

Ji Soo Yi, Youn ah Kang, John T. Stasko, and Julie A. Jacko. Toward a Deeper Understanding of the Role of Interaction in Information Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1224–1231, November–December 2007. doi:10.1109/TVCG.2007.70515. URL `https://doi.org/10.1109/TVCG.2007.70515`.

Björn Zimmer and Andreas Kerren. OnGraX: A Web-Based System for the Collaborative Visual Analysis of Graphs. *Journal of Graph Algorithms and Applications*, 21(1):5–27, January 2017. doi:10.7155/jgaa.00399. URL `https://doi.org/10.7155/jgaa.00399`.

Lnu.se

ISBN: 978-91-89709-26-3 (print), 978-91-89709-27-0 (pdf)

Linnéuniversitetet