Measuring User Experience for Virtual Reality

Dissertation
zur Erlangung des Grades
des Doktors der Ingenieurwissenschaften (Dr.-Ing.)
der Fakultät für Mathematik und Informatik
der Universität des Saarlandes

vorgelegt von
Marco Speicher, M.Sc.
Saarbrücken, Juni 2019

UNIVERSITÄT
DES
SAARLANDES

Dean:
Prof. Dr. Sebastian Hack

Head of Committee:
Prof. Dr. Verena Wolf

Reviewers:
Prof. Dr. Antonio Krüger
Prof. Dr. Adalberto Simeone

Committee member:
Dr. Florian Daiber

Day of defense:
10th July, 2019

Acknowledgements

I am thankful to all people whom I had the pleasure to collaborate and work with over the last few years. I apologize to anyone I have forgot.

First and foremost, I want to thank **Antonio Krüger** for the opportunity to work in his group and on this thesis. Thank you for the helpful support and guidance over the last years. Moreover, I thank **Adalberto Simeone** for all invaluable discussions and providing me with helpful advice, and for reviewing my thesis. Special thanks for the joint work with **Anna Feit**, **Vito Gentile** and **Salvatore Sorce**. In addition, big thanks to **Markus Langer** and **Rudolf Siegel** for the teaching hours in statistics and leading me out of dead-ends. A very special gratitude goes out to all group members at UMTL who supported and helped my in different phases (in alphabetic order): **Maximilian Altmeyer**, **Nico Herbig**, **Frederic Kerber**, **Pascal Lessel**, **Frederic Raber**, and **André Zenner**. With a special mention to my office mates **Donald Degraen** and **Frederik Wiehr**, and, of course, **Felix Kosmalla** and **Christian Lander** for the wonderful time and bright moments I was able to spend with you. Nevertheless, I want to thank **Iris Lambrecht** for addressing any legitimate concerns that have raised. Especially, I would thank the students whose work I was allowed to supervise and who helped me to realize my ideas: **Sebastian Cucerca**, **Jan Ehrlich**, **Philip Hell**, **Gian-Luca Kiefer**, **Nadja Rutsch** and **Pascal Ziegler**.

Finally, I want to express my great gratitude to my mentors, my fantastic four: **Florian Daiber**, **Sven Gehring**, **Ralf Jung** and **Markus Löchtefeld**. Thank you for your support and helpful advices in many different circumstances. I would also like to thank my family and friends for the tolerance and understanding over the last years. Yet, I would never have come this far without the most important person in my life, my wonderful wife, **Marie Speicher**. Thank you for the permanent support, compassion, and that you always believed in me!

Abstract

In recent years, Virtual Reality (VR) and 3D User Interfaces (3DUI) have seen a drastic increase in popularity, especially in terms of consumer-ready hardware and software. These technologies have the potential to create new experiences that combine the advantages of reality and virtuality. While the technology for input as well as output devices is market ready, only a few solutions for everyday VR - online shopping, games, or movies - exist, and empirical knowledge about performance and user preferences is lacking. All this makes the development and design of human-centered user interfaces for VR a great challenge.

This thesis investigates the evaluation and design of interactive VR experiences. We introduce the Virtual Reality User Experience (VRUX) model based on VR-specific external factors and evaluation metrics such as task performance and user preference. Based on our novel UX evaluation approach, we contribute by exploring the following directions: shopping in virtual environments, as well as text entry and menu control in the context of everyday VR. Along with this, we summarize our findings by design spaces and guidelines for choosing optimal interfaces and controls in VR.

Zusammenfassung

In den letzten Jahren haben Virtual Reality (VR) und 3D User Interfaces (3DUI) stark an Popularität gewonnen, insbesondere bei Hard- und Software im Konsumerbereich. Diese Technologien haben das Potenzial, neue Erfahrungen zu schaffen, die die Vorteile von Realität und Virtualität kombinieren. Während die Technologie sowohl für Eingabe- als auch für Ausgabegeräte marktreif ist, existieren nur wenige Lösungen für den Alltag in VR - wie Online-Shopping, Spiele oder Filme - und es fehlt an empirischem Wissen über Leistung und Benutzerpräferenzen. Dies macht die Entwicklung und Gestaltung von benutzerzentrierten Benutzeroberflächen für VR zu einer großen Herausforderung.

Diese Arbeit beschäftigt sich mit der Evaluation und Gestaltung von interaktiven VR-Erfahrungen. Es wird das Virtual Reality User Experience (VRUX)-Modell eingeführt, das auf VR-spezifischen externen Faktoren und Bewertungskennzahlen wie Leistung und Benutzerpräferenz basiert. Basierend auf unserem neuartigen UX-Evaluierungsansatz leisten wir einen Beitrag, indem wir folgende interaktive Anwendungsbereiche untersuchen: Einkaufen in virtuellen Umgebungen sowie Texteingabe und Menüsteuerung im Kontext des täglichen VR. Die Ergebnisse werden außerdem mittels Richtlinien zur Auswahl optimaler Schnittstellen in VR zusammengefasst.

Relevant Publications

The work presented in this thesis, including figures, tables and text fragments have been published in the following publications. The chapters of this dissertation are partly based on these publications.

Full conference and journal papers

[286] M. Speicher, A. M. Feit, P. Ziegler and A. Krüger. **Selection-based Text Entry in Virtual Reality.** In *Proc. of the 2018 CHI Conf. on Human Factors in Computing Systems*, CHI '18, ACM, pp. 647:1–647:13. (appears in Section 3.3.3)

[287] M. Speicher, P. Hell, F. Daiber, A. Simeone and A. Krüger. **A Virtual Reality Shopping Experience Using the Apartment Metaphor.** In *Proc. of the 2018 International Conf. on Advanced Visual Interfaces*, AVI '18, ACM, pp. 17:1–17:9. (appears in Section 5.4)

[289] M. Speicher, N. Rutsch, A. Krüger and M. Löchtefeld. **Exploring Spatial Menu Representation and Apartment-based Categorization for Online Shopping.** In *Human-Computer Interaction – INTERACT 2019*, Springer International Publishing. (appears in Section 5.3)

[282] M. Speicher, S. Cucerca and A. Krüger. **VRShop: A Mobile Interactive VR Shopping Environment Combining the Benefits of On- and Offline Shopping.** In *Interact. Mob. Wearable Ubiquitous Technol.* 1, 3 (Sept. 2017), pp. 102:1–102:31. (appears in Section 5.2)

[288] M. Speicher, C. Rosenberg, D. Donald, F. Daiber and A. Krüger. **Exploring Visual Guidance in 360-degree Videos.** In *Proc. of the Int. Conf. on Interactive Experiences for TV and Online Video*, TVX '19, pp. 131–133.

[283] M. Speicher, F. Daiber, S. Gehring and A. Krüger. **Exploring 3D Manipulation on Large Stereoscopic Displays.** In *Proc. of the 5th ACM Int. Symp. on Pervasive Displays*, PerDis '16, pp. 59–66. (appears in Section 2.4.2)

[61] F. Daiber, M. Speicher, S. Gehring, M. Löchtefeld and A. Krüger. **Interacting with 3D Content on Stereoscopic Displays.** In *Proc. of the 4th ACM Int. Symp. on Pervasive Displays*, PerDis '14, pp. 32:32–32:37. (appears in Section 2.4.2)

Poster papers

[285] M. Speicher, J. Ehrlich, V. Gentile, D. Degraen, S. Sorce and A. Krüger. **Pseudo-haptic Controls for Mid-air Finger-based Menu Interaction.** In *Proc. of the 2019 CHI Conf. Extended Abstracts on Human Factors in Computing Systems*, CHI EA '19, ACM, pp. 1787–1793. (appears in Section 4.2)

[281] M. Speicher. **Shopping in Virtual Reality.** In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (March 2018), pp. 1–2. (appears in Section 5.2.3)

[290] M. Speicher, R. Siegel and A. Krüger. **ProductFinder: A Location Aware Product Information Display for Retail Environments.** In *Proc. of the 6th ACM Int. Symp. on Pervasive Displays*, PerDis '17, ACM, pp. 23:1–23:2. (appears in Section 5.2.3)

[284] M. Speicher, F. Daiber, G.-L. Kiefer and A. Krüger. **Exploring Task Performance and User's Preference of Mid-air Hand Interaction in a 3D Docking Task Experiment.** In *Proc. of the 5th Symp. on Spatial User Interaction*, SUI '17, ACM, pp. 160–160. (appears in Section 5.2.2)

Contents

Chapter 1

Introduction

Within the last decade and after arriving on the consumer market, Virtual Reality (VR) is now evolving into an everyday technology [130]. Unlike the VR hardware of the 1990s, current VR hardware is able to render complex high-quality 3D scenes in real-time, provide much higher resolutions and became affordable for the average PC or game console user. Even smartphones can be used for watching movies or riding a roller coaster in VR. Furthermore, precise head and positional tracking of state-of-the-art consumer systems (e.g. Oculus Rift[1] or HTC Vive[2]) prepare the way for more interaction and naturalness. In a recent survey with people who would not buy a VR headset, only 10% stated technical reasons[3] . Instead, the main reasons are: (1) a lack of interest in video gaming, (2) the experience not being appealing to the user, and (3) the fear of physical or social harm. Hence, the rapidly growing interest and availability of VR raises questions considering the everyday use of VR, if we disregard its usefulness for gaming, military or therapy purposes [130]. As examples for everyday VR, this thesis evaluates the field of system control (see Chapter 4) and shopping in VR (see Chapter 5).

More and more everyday VR applications simulating immersive virtual environments (VEs), such as VR shops (e.g. IKEA[4]), gaming experiences (e.g.

[1] https://www.oculus.com/rift/

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

[3] https://bit.ly/2GN83vx

[4] https://demodern.de/projekte/ikea-vr-showroom

climbing [150]) or operating system (OS) simulators (e.g. Virtual Desktop[5]), are appearing in the common VR app stores (e.g. the Oculus or Steam store). But here, the user puts on a head-mounted display (HMD) and starts exploring the VE, instead of sitting in front of the screen like in the PC setting. So what will the VE look like and how will we interact with it? Typical examples of common UI patterns from the non-VR world[6] are drag-and-drop (using a 2D mouse cursor), hyperlinks (used on webpages on the internet), or the pinch-to-zoom gesture (for touch interaction). These patterns have become standard because of their good user experience (UX) and intuitiveness, but usually only for the particular technologies they were designed for. Consequently, the design and evaluation of UIs and interactions have to be revised for VR, particularly regarding VR-specific factors like physical demand, motion sickness and UX [282].

Current research in the fields of VR and Human-Computer Interaction (HCI) is also concerned with the question of how the real experience should be represented in virtual form [79], or whether a new type of experience should be designed that builds on the possibilities of the virtual or physical medium [56]. But one has to differentiate between virtual and physical familiarity when designing for VR, which leads to the assumption that just transferring a given system from one medium to another might not be the optimal solution according to performance and the user's preference. However, until it is possible to work in VR as precisely as in the real world, it is essential to come up with new ways of interaction using consumer hardware, which can support the users in achieving their goals. As a first step, one would adapt existing patterns and guidelines from interaction design for PC and mobile devices. People wearing a HMD could use gamepads or mouse/keyboard, and traditional controls might also be efficient for VR, but maybe not regarding VR-specific factors and metrics.

Overall, VR presents UX challenges as well as opportunities. When using the head as a controller, how should one attract the user's attention in a VE, which surrounds the user completely, unlike the situation of a user sitting in front of, for example, a 22-inch screen? As the user is able to fly

[5] https://www.vrdesktop.net/

[6] http://ui-patterns.com

around in the VE while being stationary in reality, how does motion sickness and workload affect the UX? At the most basic level, UX research is about understanding users and not about software [38]. Therefore, user evaluation is critical for the success of new VR software.

However, when designing novel UIs for VR, questions arise as how to evaluate the designed or implemented prototypes and which VR-specific factors and metrics should be used for the evaluation of VR experiences. We envision future VR experiences to have their origins in UX design. But the current status of VR research does not offer a evaluation framework tailored to VR with a focus on UX. Existing evaluation methods for 3D User Interfaces (3DUI) [29, 76] or characteristics of VR [169] build a positive and promising foundation. But there is a gap in evaluating UX, which has become a fundamental and important metric for UI design [157]. Although isolated methods include external factors such as the system used or physical attributes of the user [28], more VR-specific factors need to be included, such as interactivity or convenience of the VR system, and other non-VR factors might be disregarded.

Therefore, the goal of this thesis is to investigate the interaction in VR focusing on VR-specific factors and metrics to make future VR applications more comparable. Today's VR systems enable a variety of experiences from different areas and UX research may benefit from a VR-based methodology [247]. But to fully understand VR as a valid representation of a VE, it is essential to investigate to what extent not only the performance of the tested method, but also the cognition, behavior and experiences of the users are relevant. Hence, we contribute a new model for UX in VR (see Chapter 3).

1.1 Definitions

In the following, we define the most important terms of this work, thus ensuring that the components of our problem statements, research questions and contributions are clearly understood by the reader. The majority of these terms have controversial definitions, same from different areas of research. We intend solely to provide definitions that should enable the reader to understand how these terms are used in the context of this thesis.

Virtual Reality (VR) describes the experience of being in an interactive computer-generated world [295, 238]. A VR experience is a synonym.

Everyday VR describes activities and interactions that the user experiences in VR regularly in different contexts of daily life [91], like online shopping, games, or movies. The majority of such activities are basic interactions such as object selection and manipulation, navigation, or system control, which includes text input and menu control [29].

VR Systems are user-computer interfaces [42] involving real-time simulation and interaction in a VE through multi-sensory feedback (e.g. visual, auditive and haptics). VR systems require stereoscopic 3D representations in contrast to classic monoscopic desktop settings. Based on Jerald [129], the main components of a VR system are the *virtual environment*, the VR display for *output*, the *tracking system*, one or more *input devices*, and an *external computing node*.

Virtual Environments (VEs) are digital 3D spaces in which the users' movements are tracked in real-time and their environments are provided in the form of a 3D scene. The movements of a user's controller (e.g. gamepad) can also be tracked and the virtual character moves and looks around accordingly.

Head-mounted Displays (HMD) can be described as VR displays where the graphics always appear on screens coupled to the user's head [29]. This creates the illusion that the VE completely surrounds the user.

Commodity Devices are affordable devices that are generally compatible with each other.

Consumer Hardware is intended for everyday use, typically in private homes. It includes commodity devices used for entertainment (e.g. televisions, game consoles), communication (e.g. smartphones) or office activities (e.g. desktop computers). Consumer VR hardware includes affordable input devices (e.g. Leap Motion, Microsoft Kinect) and complete VR systems (e.g. HTC Vive, Oculus Rift, Samsung Gear, Google Daydream). A detailed classification of current consumer VR systems can be found in Section 2.2.5.

Degrees of Freedom (DOF) describe the number of axes that can be controlled by the user. Three DOF are required for controlling the user's head position (width, height, depth) while another three DOF define its orientation (yaw, pitch, roll).

Field of View (FoV) describes a solid angle that determines how much of the VE is visible from the user's point of view. A distinction is made between horizontal and vertical FoV. It should be mentioned that the FoV is independent of eye movements in this definition, but not of head movements. The angle determining how much is visible for the user without eye movements is called *Field of Regard (FoR)*.

User Interfaces (UIs) include hardware and software components which enable the user to interact with the system. The hardware components of a UI include input devices such as keyboards, mice, gamepads or gloves, as well as output devices such as HMDs, monitors, projectors or speakers. Software components of a UI include menu controls and widgets such as buttons or sliders. *Intuitive* UIs are designed based on known metaphors and are distinguished by their familiarity and/or naturalness, e.g. pinching for scaling or grabbing for object selection.

3D User Interfaces (3DUIs) are UIs that involve 3D interaction.

Interaction Techniques are methods allowing the user to accomplish a given task via a UI including both hardware and software components. Current operating systems require the user to interact in only two dimensions, such as by moving a mouse pointer. In VR UIs, though, the user exchanges information with the computer system in 3D space.

Isomorphism characterizes a one-to-one mapping between interactions in the real world and their effects in the VE [29].

Task Performance metrics include quantitative measurements such as *task completion time*, *error rate* and *accuracy* (or precision) [29]. These metrics indicate to what extent users are able to cope with the task and the interaction method.

User Preference metrics usually consist of subjective feedback, such as *usability*, *user experience*, *motion sickness* and *task workload* [29].

Usability is described as the extent to which a system or method can be used by certain users in a certain context to achieve certain goals effectively, efficiently and satisfactorily, see DIN EN ISO 9241.

User Experience (UX) expands on the aspects of usability with aesthetic and emotional factors, such as an appealing, novel and desirable design, aspects of confidence-building and stimulation, or enjoyment during use [157].

Motion Sickness describes the amount of nausea and dizziness a user feels during or after a VR experience. Riding in a roller coaster, sailing aboard a ship, or flying on a plane can cause motion sickness and lead to discomfort [163]. While motion sickness can influence the task performance and user preference, it is not a permanent condition and can be reduced (see Section 2.3.3).

Immersion refers to what technology achieves from an objective point of view [277]. The more feedback a VR system provides in all sensory modalities, i.e. visual, auditory or haptic, and the more precise its tracking (e.g. head tracking), the more *immersive* it is. *Partial* immersion supports the feeling of "observing" a VE (e.g. in a flight simulator), while *full* immersion supports the feeling of "being" [270] in that environment (e.g. in a HMD).

Presence describes human reaction to immersion, i.e. to what extent the user feels present within the VE. In the same immersive system, different people can experience other levels of presence, and different immersive systems can lead to a similar presence for other people [277].

Outside Factors are external factors that potentially affect the performance or preference of a technique, UI or tested device. The experiences in VR can vary between users, or due to the system used, the environmental design, and the tasks to be performed [28, 134, 130].

Design Spaces are standard approaches to represent design rationale and use a semi-formal notation, like the *QOC (Questions, Options, and Criteria)* approach by MacLean et al. [188].

1.2 History of Virtual Reality Experiences

For many decades, researchers around the world have been doing research in VR to improve the technology and explore its possibilities. While much progress has been made in recent years and the technology may finally gain a foothold in the consumer market, VR research is not new and many fundamentals and findings that are still valid today were already developed in the 1980s and 1990s. But at that time, VR hardware was rarely available, expensive, and in many aspects like video resolution, convenience and computing power, less powerful and usable than today. This led to VR being developed and used primarily in military and industrial areas.

The history of VR systems is often considered to begin with the American physicist Edwin Land, who constructed the first polarization glasses in 1932 (see Figure 1.2), from which the Polaroid Company was born in 1937. Even today, various polarization filters are used in stereoscopic films or 3D cinemas to display two images taken from two different points, i.e. one for each eye, in order to create spatial experiences and better depth perception. Six years later, the Cinerama [318] was a further development in the field of more immersive representations of virtual environments. Using three synchronously running cameras and corresponding projectors, a frame rate of initially 26 frames per second (fps, see Section 2.2.2) could be achieved, followed by 24 fps. The curved screen completely filled the peripheral field of view (FoV, see Section 2.2.2) of the viewers and thus achieved a higher immersion. Even today a wide FoV is an important component of Head Mounted Displays (HMD: see Section 2.2.2) regarding immersion and the sense of presence [174].

Another ten years later, Morton Heilig built the first prototype of his

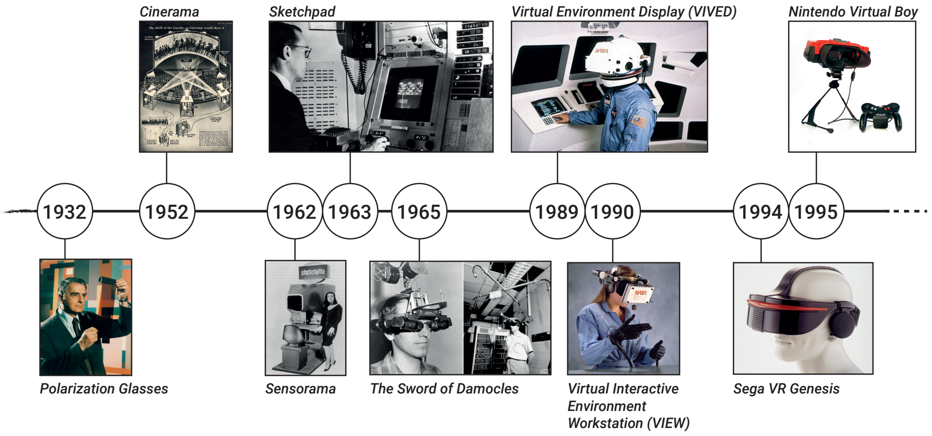

Figure 1.1: This figure shows the virtual reality timeline from its origin in 1932 to the end of the first widespread publicity of VR in the 1990s.

Sensorama [111], the first multi-sensory immersive VR system. With built-in vibrating mechanism, stereoscopic images, and an odor and wind system, it exploits the possibilities of all three feedback dimensions, namely visual, auditory, and haptic [29]. In addition to Sensorama, the Headsight Television System [58] also attracted attention in the early 1960s. This system includes a magnetic motion tracking system to determine head orientation. While the development of Headsight was the first step in the development of HMDs, it lacked the integration of computer and image generation. At the same time, Sketchpad [300] was developed, which can be regarded as one of the first interactive graphics applications. In addition, it made fundamental contributions to HCI research and was one of the first GUIs. It used the Light-Pen, the predecessor of the mouse, so that the user could point to and interact with objects displayed on the screen.

Soon after, the first HMD was developed. The Sword of Damocles [301] was so heavy that it could not be carried by a human and had to be installed on the ceiling using a rack. A wireframe cube floating in space with an edge length of about 5cm is considered to be the first tracked object represented in VR and whose position changes were directly transmitted. Furthermore, in his famous essay "The Ultimate Display" [301], Sutherland describes his vision of a futuristic display that allows the user to dive into virtual

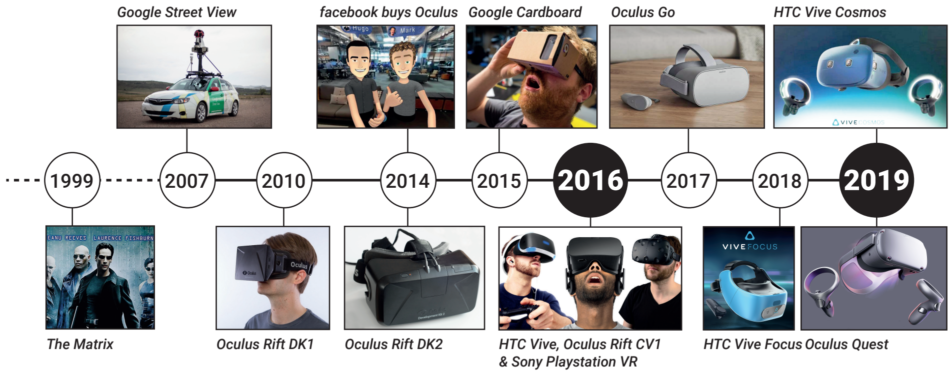

Figure 1.2: This figure shows the virtual reality timeline from the 2000s to today, with its second phase of popularity in 2016 and another possibly to come in 2019.

computer-generated environments via novel multi-modal input and output devices. In Section 2.1.2, we will further discuss the different definitions of VR and the evolution of the term. With the advancement of technology and the discovery of the potential of VR, the 1980s were marked by various NASA VR research projects. With the Virtual Environment Display (VED) [83] and the Virtual Interactive Environment Workstation (VIEW) [84], astronauts were supported in the complex control of a new spacecraft by means of tele-robotic systems. In the Virtual Planetary Exploration (VPE) project, NASA tried to visualize the huge amounts of data transmitted by the Viking probes from Mars in a suitable way. Besides the mainly military and medical interest in VR, there were first attempts in the 1990s to enter the consumer market with the Nintendo Virtual Boy[7], Sega VR or CyberMaxx. The Virtual Boy is considered as one of the biggest flops for Nintendo[8].

The film "The Matrix" by Wachowski was released in 1999. The film shows characters who live in a simulated reality, i.e. they cannot distinguish between the virtuality and the "true reality", and as in today's VR systems, physical laws can be overridden. Non-isomorphic (or magical) interactions such as tractor beams for selection [229] or teleportation for navigation [33]

[7] https://nintendo.fandom.com/wiki/Virtual_Boy

[8] https://www.telegraph.co.uk/technology/0/biggest-technology-flops-history/nintendo-virtual-boy/

replace isomorphic (or natural) interfaces. Although some earlier films, such as "Tron" (1982) and "Lawnmower Man" (1992), dealt with the representation of virtual reality, "The Matrix" had a great cultural influence and brought the theme of simulated and virtual reality into the mainstream[9].

Today, about 20 years later, public interest in VR technology is greater than ever before. Driven by the interest of the entertainment industry and other industries in VR, the development of VR-HMDs has made great progress in recent years. As a result, after the successful launch of second-wave consumer systems in 2016 (Oculus Rift, HTC Vive, Sony Playstation VR), the third wave of VR HMDs using inside-out tracking (e.g. Oculus Quest, HTC Vive Cosmos) is now on the verge of reaching mass markets around the world at affordable prices for consumers (see Figure 1.2). The current VR technology benefits from highly available, cost-effective and powerful hardware, which was not available for personal use or simply not affordable in the 80s and 90s. In today's huge and constantly growing gaming industry, most households have systems for simulating complex VEs. With the availability of high-performance gaming laptops, VR systems today are more flexible. In Section 2.2.5, we provide an overview of the VR systems currently available on the market and classify them according to interactivity, comfort and graphic quality.

1.3 Problem Statement

In order to design enjoyable and efficient interaction techniques, UIs and applications for VR systems, the key challenge is to build an understanding of how users perceive the content and which external and internal factors play a key role for UX. This requires understanding how each factor is defined and how it can be measured and improved. To gather the needed insights, we address the following problems:

How can we categorize different types of commodity VR systems? In the last decade, there was no commodity VR hardware available. VR researchers were forced to resort to homebrewed or expensive devices and

[9] https://www.revolvy.com/page/History-of-science-fiction-films

computer systems, which could not be connected and integrated easily. With the constantly growing development and availability of technologies in various areas of VR, novel and affordable input and output media for VR is now available for everyone. Until 2017, VR HMDs could simply be categorized in two ways: portable (or mobile and wireless) or not. Moreover, with the advent of the first VR HMDs without external sensors and computing nodes (e.g. Oculus Go) and mobile solutions for high-end hardware (e.g. the HTC Vive wireless module), a distinction has to be made here. Therefore, we want to categorize and illustrate the differences and commonalities of different types of consumer VR systems regarding comfort, graphics, and interactivity aspects.

How can we classify different types of VR-specific factors? Certain VR-specific factors and influences restrict the use of VR UIs. We want to provide an example set of outside factors, which are clarified and determined prior to the evaluation phase. Outside factors like health and safety issues related to motion sickness might have the potential to significantly delay, or at worst prevent, the further development of immersive VR interfaces. But task- and environment-specific aspects like difficulty of the task and appearance of the environment could also have an influence on the UX in VR. We also want to classify VR-relevant metrics, e.g. how motion sickness or workload is defined and caused, which should provide conclusions about the effectiveness, efficiency and user preference of the evaluated artifacts.

How can we measure and improve UX in VR across varying contexts? We want to develop principles for the design and evaluation of effective and usable VR environments, interaction techniques and user interfaces. The main purposes of usability evaluation refer to metrics, such as objective performance and subjective reactions [161], to capture everything about an artifact or a user who influences the use of the artifact.

How can applications for VR leverage performance and preferences? Traditional input might not work in full immersive VR UIs, because VR users are not stationary in general. To overcome limitations in the tracking space or anatomical constraints, non-isomorphic techniques allow users to interact using "supernatural" metaphors. We want to investigate if natural (or isomorphic) interaction has an impact on task performance. And do isomorphic concepts provide higher user preference due to their familiarity, or can the user adapt to the non-isomorphic methods? How this can be accomplished has so far not been fully explored.

VR-specific affecting factors, like environmental or human-related influencing aspects (e.g. appearance of the VE and physical attributes of the user) and metrics (e.g. motion sickness, presence and workload) need to be integrated and considered within the evaluation. In order to form a basis using the full capabilities of UX in VR, this thesis addresses the problems mentioned above in the following research questions.

1.4 Research Questions

As mentioned before, the work in this thesis explores isomorphic and non-isomorphic interaction in VR, i.e. the mappings between interactions in the real world and their effects in the virtual environment. Simply adapting 2D UIs from non-VR is not always the best solution [29]. Thus, different sets of interaction techniques and UIs were designed, which rely on well established metaphors. These artifacts were evaluated and investigated in two distinguishable application areas for everyday VR, namely system control in VR like text input or menu control (see Chapter 4) and shopping concepts (see Chapter 5), with respect to our novel approach for evaluating VR experiences (see Chapter 3). The research objectives of this thesis are split into four directions: first, a theoretical analysis of interaction in VR and its evaluation characteristics; second, identifying outside and inside factors for VR UX evaluation; third, shopping in VR; fourth, VR system control; those suit the unique properties of everyday VR. This thesis thereby addresses the following research questions:

What are the key aspects of the evaluation of VR experiences? The work in this thesis includes a theoretical analysis of existing VR systems in order to identify the outside factors (e.g. environment, system, task or human), which could influence the VR-specific measurements like UX, motion sickness or workload. Moreover, we derive a novel evaluation approach based on VR-specific characteristics to provide means for making interactive VR experiences comparable with each other to inform the design and development process. In particular, we address the missing link between objective performance and subjective users' preferences concerning the metrics.

What are suitable interaction techniques for different contexts? There is a lack of user-friendly and intuitive UIs and interaction techniques, as well as a connection to previous findings from basic research on VR and 3DUI. While military, education and gaming topics are well investigated, basic topics such as shopping, text entry or menu control are rarely studied in VR research regarding performance [96, 194] and user preferences (e.g. UX, workload, and motion sickness). This work therefore focuses on the development and evaluation of novel and immersive VR experiences, aiming to include the main advantages and limitations of isomorphism, e.g. grabbing vs. tractor beam for selection and manipulation in VR [287, 241].

How should interaction be designed for VR? One purpose of this research is to develop guidelines and principles for the design and evaluation of effective and usable VR environments. Design guidelines for VR UIs are directly adapted from non-VR systems without any further investigation, like head or controller pointing on a QWERTY keyboard for text input in VR. But what if tracked hand-held controllers are not available, or low physical demand and motion sickness are of particular importance? Overall, does isomorphic interaction using the user's hand or fingers for input have any impact on task performance or the user's preference? To answer these questions, we present example design principles for shopping in VR (see Section 5.2.2), as well as a design space for text entry in VR (see Section 4.1.2).

1.5 Methods and Approach

The methodology can be divided into several steps. First, we conduct a thorough literature review and present a classification of current VR systems.

Second, this work provides a multi-factorial approach to the evaluation of different artifacts in VR (see Chapter 3.4), which recommends not only external factors but also different VR-specific objective and subjective evaluation metrics. Our approach for VR UX evaluation combine the benefits of state-of-the-art usability evaluation methods and approaches, like the user-centered design (UCD) approach (see Section 3.2.1), testbed evaluation approach (see Section 3.2.2) or sequential evaluation approach (see Section 3.2.3), which form the basis for most user evaluations in 3DUI research [29, 88]. We address the challenges and risks of user-centered evaluation in VR, analyzing data using VR-specific metrics, and providing heuristics and guidelines for future VR developers and designers.

In the further course of the work, different studies examined the correspondence of performance and preferences in different environments for everyday VR applications and situations, e.g. text entry in VR (see Section 4.1), finger-based menu control in VR (see Section 4.2) or shopping in VR (see Chapter 5). In this context, we develop and evaluate interactive high- and low-fidelity prototypes involving users to ensure that an efficient, user-friendly and pleasant result is achieved. We consistently involve the users in all phases of development and place their needs at the center of our decisions. We uncover the requirements and expectations users have of the system, the goals they pursue when using it and the context in which the system is used. This knowledge is built in a user-centric approach through a series of experiments and studies.

In the context of reporting on the experimental measurements and evaluations of users, the potentials, prerequisites and possibilities for the implementation and design of VEs and UIs as empirical research instruments in the field of HCI are discussed. Based on quantitative data, statistically significant differences between the evaluations of the tested interaction techniques and UIs were found; analysis based on experimental results showed differences, similarities and limitations. Overall, this thesis provides knowledge to researchers and professionals engaged in the design of technological interfaces about the usefulness of VR in the evaluation of UX.

Now, we will give a brief overview of the aforementioned key contributions of this thesis, which are discussed in detail in Chapter 3, 4 & 5.

1.6 Contributions

The goal of this thesis is to investigate the interaction in VR with a focus on VR-specific factors and metrics to make future VR applications and their interfaces and devices more comparable. Today's VR systems enable a variety of applications from different areas. In order to improve existing interactive VR experiences and to support the development of new applications, this work contributes in the following three areas:

1.6.1 Theoretical Contributions

We introduce the Virtual Reality User Experience (VRUX) evaluation approach with focus on 3DUI- and VR-specific external factors, and combine the metrics of evaluation of 3DUIs, and the characteristics of VR (see Section 3.4). This allows a more refined and differentiated classification of interaction techniques and UIs for VR. Furthermore, we present a classification of common affordable and commodity VR systems based on their interactivity, comfort and graphics quality. The influence of marketable consumer input and output devices has been examined in more detail with regard to task performance and user preferences.

1.6.2 Technical Contributions

While military, education and gaming topics are well investigated, basic topics such as shopping, text entry or menu control are rarely studied in VR research regarding performance [96, 194] and user preferences (e.g. UX, workload, and motion sickness). To fill in the missing link between on- and offline shopping, we designed and implemented a VR online shopping framework based on real data from a local retailer, which forms a basis for our two immersive VR online shopping environments. These prototypes aimed to maintain the benefits of online shops, such as search functionality and availability, while simultaneously focusing on shopping experience, clarity and immersion. Additionally, we present a finger-based pseudo-haptic UI for menu control in VR based on physical metaphors. Furthermore, we present a selection of different text input methods for VR.

1.6.3 Design Contributions

Previous research has taught us not to assume that transferring conventional UIs into VR environments would be accepted by the users [161]. Instead, interaction should be tailored specifically to the immersive VEs, resulting in a "better" version of the reality [296], e.g. bypassing physical limitations like gravity or arm reach. In order to support and help upcoming VR designers and developers of VR shops to create experiences that do not frustrate or cause motion sickness, we describe design principles for shopping in VR (see Section 5.2.2) and a design space for selection-based text entry in VR (see Section 4.1.2), including questions, options and criteria. There is little research in comparing text entry methods in VR [32, 96, 335], so we believe that there are still open questions, and VR designers in particular are still provided with little guidance for text entry in VR.

Furthermore, we formulate guidelines to guide the work, including actionable insights on how to optimize performance, usability, satisfaction, and experience for the users of VR. In this context, we provide the main characteristics of online and offline shopping, as well as a list of potential guidelines and lessons learned to inform the design of VR shop and text entry interfaces. The design spaces and the evaluated methods provide a solid baseline for comparison of future VR applications. Particularly in VR, UX and workload turned out to be essential factors for task performance.

1.7 Outline

The remainder of this thesis is structured as follows. Chapter 2 describes the theoretical background and related work for this thesis. It includes a definition of VR, as well as a categorization of human factors in VR and current VR systems. In Chapter 3, we present a framework for prototyping and evaluation of VR experiences with 3D content. This framework is developed with respect to sustainability and future work. The intention is to use commodity VR systems, which can be easily understood by novices and configured by a minimal effort of instruction and they should be natural to use. Based on the evaluation framework and inspired by related work, a set of various everyday VR scenarios are designed and implemented, like system control (see Chapter 4) or shopping in VR (see Chapter 5).

In Chapter 4, we present and compare six methods for selection-based text entry in VR (see Section 4.1), as well as two visual approaches to mid-air finger-based menu control in VR environments (see Section 4.2).

Chapter 5 starts with a theoretical background focused on VR in retail, alternative shopping concepts, and commercial VR shop applications. A customer survey (see Section 5.2.2) has been conducted to explore the main characteristics between on- and offline shops (see Section 5.2.2), followed by principles for VR shops (see Section 5.2.2). In Section 5.2, we present a user study with a product search task based in a WebVR online shop using speech input in combination with VR output. In a subsequent study (see Section 5.3), we compared traditional linear store representation and categorization of online shops with a novel approach using the apartment metaphor in a spatial grid. As a concluding system in a third study (see Section 5.4), we adapted the apartment metaphor to the representation a shop and explored selection and manipulation in VR. Finally, the work concludes by summarizing the main contributions of this thesis and potential future work, followed by concluding remarks.

The majority of the work that is presented within this thesis was carried out in collaboration with researchers and students from different institutions. Therefore, the scientific plural "we" is used throughout the thesis. All URLs are treated as references and were last visited on the 30th of April 2019.

Chapter 2

Background and Related Work

This chapter presents the theoretical background for this work, which refers to the evolution and definition of the Virtual Reality (VR) term, the interaction in VR, as well as user interfaces (UIs) in full immersive virtual environments (VEs) and the human factors that play a main role in the design and development of VR applications. After the definition of VR, we describe the main components of a state-of-the-art VR system. In the following, we address the confounding terminologies of the main characteristics and human factors in VR, namely usability and user experience (UX), task workload, motion sickness, as well as immersion and presence. VR is a very promising technology because it is the combination of and missing link between known reality and virtuality. This requires applications and concepts of interaction and UIs, which are the most important contributions of this work, combining the advantages of both worlds.

2.1 Virtual Reality

Virtual reality (VR) enables people to experience immersive virtual worlds in such a way as if they themselves are part of these worlds. During such an experience, the virtual world stimulates many channels of the user's perceptual system, such as the visual, acoustic or haptic channel. These stimuli are calculated by computer systems that simulate the *virtual environment (VE)*. This powerful technology promises to change our lives like no other. By artificially stimulating our senses, our bodies are tempted to accept another version of reality. In virtuality, laws of nature can apply, but do not have to.

But VR should not be confused with *simulated reality*, which assumes that the simulation cannot be distinguished from the "true reality" to a certain degree. An example scenario for a simulated reality is the disembodied brain or the brain-in-a-jar problem [244]. The brain could be placed in a jar of "life-sustaining" liquid and connected to a supercomputer to receive and give impulses from and to the simulated reality. But here, it is – from the philosophical point of view – not possible for the brain to know or distinguish if it is located in a human head or in the actual jar, as the brain receives the same impulses in both situations (see Figure 2.1). Another famous example of a simulated reality is the Matrix film series, where people are no longer aware that they are living in a simulation. This kind of simulation differs clearly from the current, technologically feasible concept of VR, which can still be easily distinguished from reality. There have been and continue to be discussions on this topic, ranging from philosophical discourse [244, 64, 57] to practical applications in computer science [86].

2.1.1 Reality versus Virtuality

Virtuality is a philosophical concept based on the French thinker Gilles Deleuze [69], who used the term "virtual" to refer to an aspect of an ideal reality that is nevertheless real. However, before we proceed, it is important to clarify the terminology of "real and virtual environments". For this purpose,

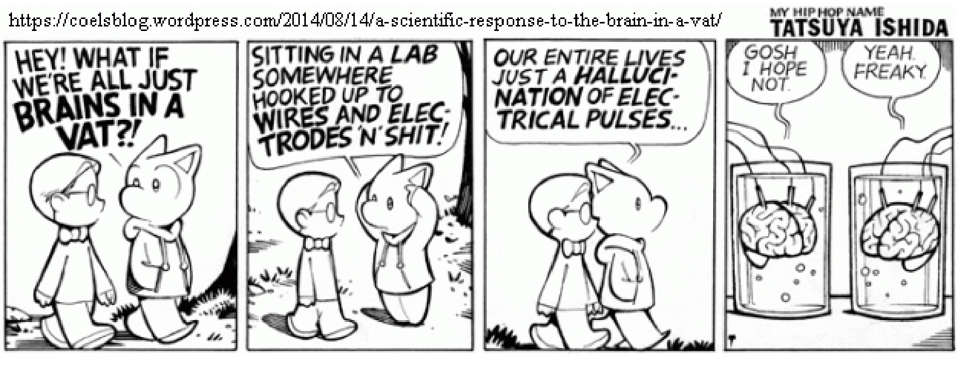

Figure 2.1: Cartoon by Tatsuya Ishida. Brain-in-a-jar thought experiment.

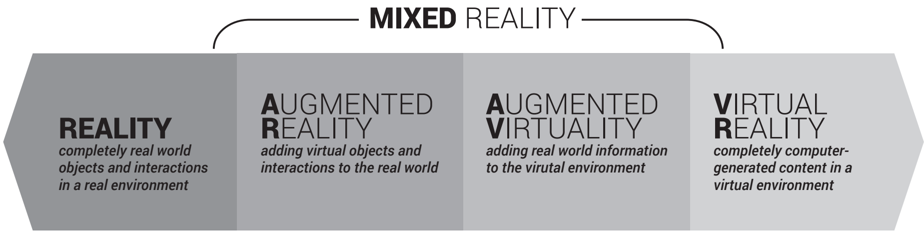

Figure 2.2: Reality-Virtuality Continuum by Milgram et al. [203].

the *Reality-Virtuality Continuum* [203] offers a simple and clear overview by means of a continuous scale, which moves between a completely virtual environment and the real environment (see Figure 2.2). It includes all possible variations and compositions of real and virtual objects. The area between the two extremes, in which both the real and the virtual are mixed, is called *mixed reality*. This in turn consists of both *augmented reality (AR)*, where the virtual complements the real, and *augmented virtuality (AV)*, where the real complements the virtual. In a VR environment, the user is thus completely in an unreal world, where the laws of nature (e.g. gravity or time) can apply but do not have to. Looking at the general differences between AR and VR, AR is also described as “a limited form of VR with see-through HMD” [203].

2.1.2 Evolution and Definition of the Virtual Reality Term

There is no concrete definition of VR, so opinions differ depending on the field in question and the mode used to achieve VR. VR has different definitions based on different views, like the one emerging from human-machine interfaces (HMIs), which not only enable the user to control the machine, but also allow observation and intervention. In this context, prior works [9, 338] define VR as a means of natural interaction technology, i.e. the user can be immersed into a computerized or simulated environment, and operate and interact naturally with the environment. VR is also defined based on its devices, more precisely as the provision of a 3D reality that is realized by a series of sensor devices like head-mounted displays (HMD), data gloves, and others [301]. Here, VR is interpreted as a software and hardware environment, which simulates a real world, in which the user

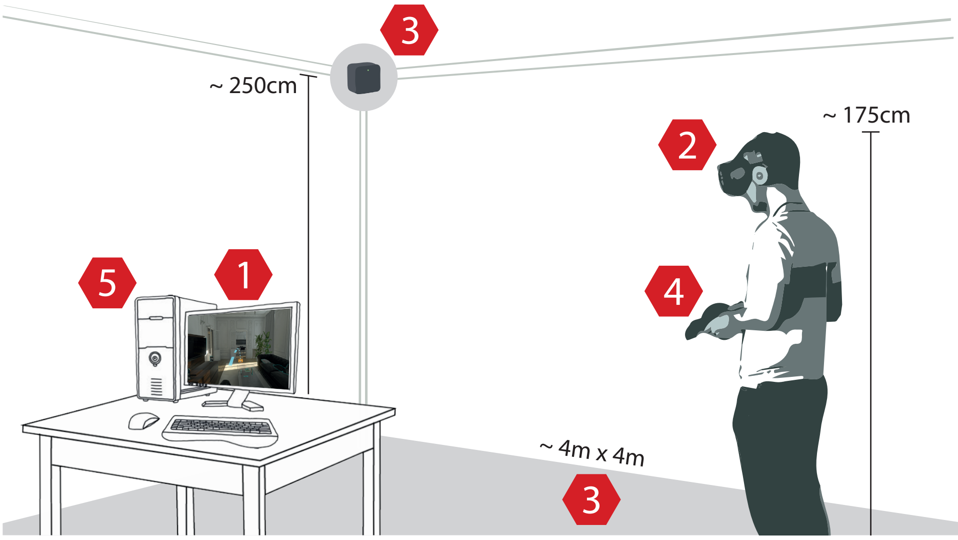

Figure 2.3: Basic components of a VR system. (1) Virtual environment. (2) Virtual reality display, here HTC Vive. (3) Tracking sensors and area, here HTC Vive optical trackers (at $2.5m$) and $4 \times 4m^2$ tracking space. (4) Tracked hand-held controllers. (5) External computing node, here a PC for experiment control and filling out questionnaires.

can operate and control the VE by interacting with devices. So from a technical perspective, VR can be defined as a medium in terms of a collection of technical hardware. Similarly, Coates [53] defines VR as an electronic stimulation wearing HMD and "wired clothing", which enables the user to interact with the 3D environment. Furthermore, VR has been interpreted by Greenbaum [295] as an "alternate world filled with computer-generated images" that respond to user interaction, as well as by Pimental [238] as an "interactive, immersive experience generated by a computer". Finally, Burdea and Coiffet [42] extended the former definitions of VR as a "high-end user-computer interface" by involving real-time simulation and interaction in a VE through multi-sensory feedback (e.g. visual, auditive and haptics).

2.2 Devices and Application Examples

Figure 2.3 shows the main components of a HMI that are fundamental to any immersive VR system. Based on Jerald [129], the main components of a VR system consist of a *virtual environment (VE)* as follows: the VR display for *output*, a *tracking system*, *hand* or *hands-free input devices*, and finally an *external computing node*. Here, the function of the *input devices* is to interact with the *VE* and the *output devices* help in feeling the presence [226]. An optional *tracking system* can be used to track the head and body movements, as well as potential input devices used. Software running on an *external computing node* (e.g. PC, smartphone or internal hardware like the Oculus Go) is implemented with a game engine like Unity3D or UnrealEngine and is used to control and synchronize the whole environment. In the following, we describe each component in more detail.

2.2.1 Virtual Environment

A VE is a digital space in which the users' movements are tracked and their environment provided in the form of a 3D scene. An example of a VE is when in a computer game, the movements of a user's controller (e.g. gamepad or mouse/keyboard) can be tracked and the character moves and looks around accordingly. Movement and viewing direction are therefore generally controlled by the user and not by the system. There are different forms of VEs that are determined by the capabilities of the platform or hardware used to create the environment.

A VE is said to be immersive when the computer-generated environment appears to enclose the user and when those parts of the physical world that are not integral system components are not visible. In a head-mounted display (HMD), the graphics always appear on screens coupled to the user's head, but this creates the illusion that the VE completely surrounds the user. In a driving simulator (see Figure 2.2.2), the environment appears behind the window, outside the vehicle, and is updated as the vehicle moves so that the VE seems to surround the user. The physical "cockpit" of the simulator does not block the view, but is part of the simulation. For all common VR systems, accurate and precise head tracking is required to perceive the VE to be as immersive as possible [322].

2.2.2 Virtual Reality Displays

Besides the classic monoscopic output of the virtual world, e.g. 2D objects on conventional PC monitors, VR systems use stereoscopic 3D representations, i.e. one image per eye. These computer-generated images are then displayed jointly to the user either by using AR/VR HMDs [85], common desktop monitors or in entire rooms equipped with large projection surfaces. The literature on VR shows the use of a variety of different display formats, as seen in Figure 2.4, such as fish tank VR (stereoscopic 3D on ordinary monitors), HMDs, large panoramic screens or Cave Automatic Virtual Environments (CAVEs) [270, 75, 298, 254].

The purpose of the fish tank technology is to simulate a display that behaves like a window to the VE. A fish tank VR is thus characterized by a stereo image of a 3D scene viewed on one or more monitors with a perspective projection coupled to the viewer's head position [68, 195]. However, this simulation requires the knowledge of the user's position, or at least the coupling of position and orientation of the head in real time with the VE [81]. The results of a study by Ware et al. [322] suggest that precise head tracking could be more important than stereoscopy for the perception of three-dimensionality on a screen. On the other hand, the goal of the first CAVE was to create a VR environment suitable for scientific research and to provide a UI for controlling high-performance computer applications running on remote supercomputers. It provides a novel way of visualizing the VE and creates a sense of presence by surrounding the user with multiple projections, i.e. displaying stereoscopic images on the walls and floor [191]. An advanced variant of a CAVE is the CABANA [62] system, which has movable walls so that the display can be configured in different display forms such as a wall or L-shape.

All these different display formats have basically different properties and are better or worse suited for different VR applications. For example, they have significantly different potentials for single users, distributed and non-distributed collaborative applications [43]. Demiralp et al. [70] suggest that fish tank VR displays are more effective than CAVEs for applications where the task is outside the user's reference frame: more specifically, when

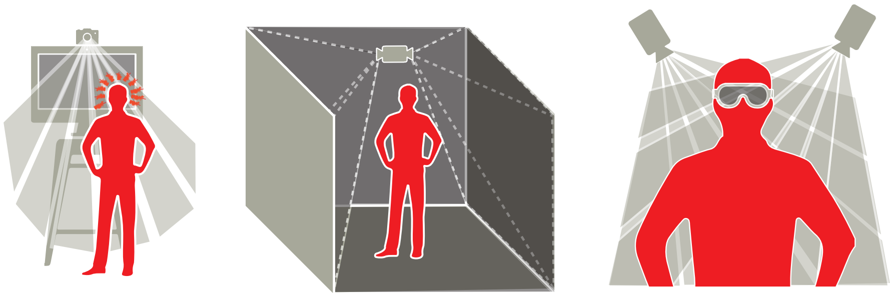

Figure 2.4: Common formats for virtual reality displays. From left to right: fish tank VR, CAVE, HMD.

the user views and manipulates the virtual world from the outside, and the interactive virtual object is smaller than the user's body and can be viewed in the fish tank display without clipping. Furthermore, it should be noted that when using projection screens, i.e. world-fixed displays (e.g. CAVEs), physical objects can stand in the way of virtual objects. This is not the case with HMDs, as their displays are placed close to the user's eyes. Moreover, the representation and material of physical objects can be substituted on purpose [274]. VR HMDs exclude the interaction between the VE and physical objects and do not support peripheral vision [163], which can lead to higher immersion. But this can also lead to a loss of orientation and increase the risk of motion sickness. An HMD-based VR system remains attractive, as it is potentially one of the most cost-effective VR systems that allows users to experience a rich and perceptible VE.

Finally, ensuring this in the best possible way presents many challenges, such as precise tracking of head position and orientation, low latency and careful calibration [130]. Ultimately, we see the future for VR applications in fully immersive VEs and therefore focus on VR HMDs in the further course of this thesis. We will now give a short overview of the display technology of VR HMDs, including more technical aspects.

Display Technology

Head-mounted displays (HMDs) were defined by Bowman et al. [29] as "a computer graphics display worn on the user's head so that the graphics

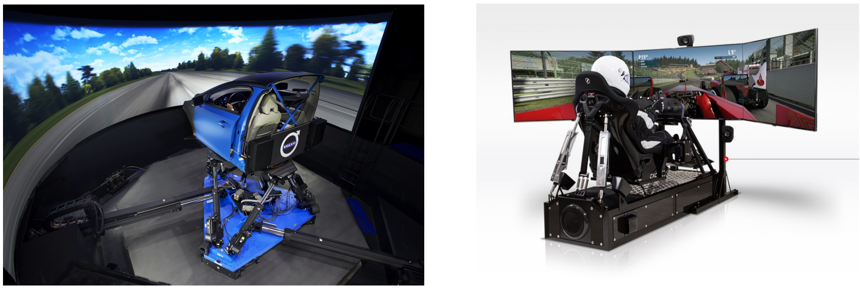

Figure 2.5: Left: VI-grade's DiM Driving Simulator at Volvo Car Group[12]. Right: Motion Pro 2 Racing Simulator[13].

displayed are continuously in front of the user's eyes". Nowadays, the most commonly used display technology in HMDs is Organic Light Emitting Diode (OLED), which is also used in smartphones, TVs and monitors.

When comparing the different display technologies, the basic rule is: the more pixels per inch (ppi) of the display, the clearer the image. But how much resolution (per eye) is enough? Is there a threshold above which pixel density brings little or no added value for us? The current industry standard, or good quality of photos, usually requires 300 ppi, and magazines are typically printed at 300 dots per inch. This led Apple to claim that the 326 ppi of their Retina display[14] is considered a soft threshold at which the human eye can no longer see the difference. But since the HMD is a few centimeters away from the eye, this high pixel density could make the difference between blurriness and clarity, i.e. the lines separating the pixels become visible. This issue is called the "screen-door effect" [71] and can be observed by looking at a projection or computer screen close up.

Field of View

According to Badcock et al. [12], the human horizontal field of view (FoV) for one eye is about 140°, and both eyes see the same in a range of 120°, when looking straight ahead. The monocular FoV for both eyes is about 160°, with the help of eye rotation up to 200°, by head rotation up to 250° and finally by body rotation full 360° (see Figure 2.6). The FoV including the eye or

[12] `https://bit.ly/2JXKqSp`

[13] `https://www.cxcsimulations.com/products/motion-pro/`

[14] `https://prometheus.med.utah.edu/~bwjones/2010/06/apple-retina-display/`

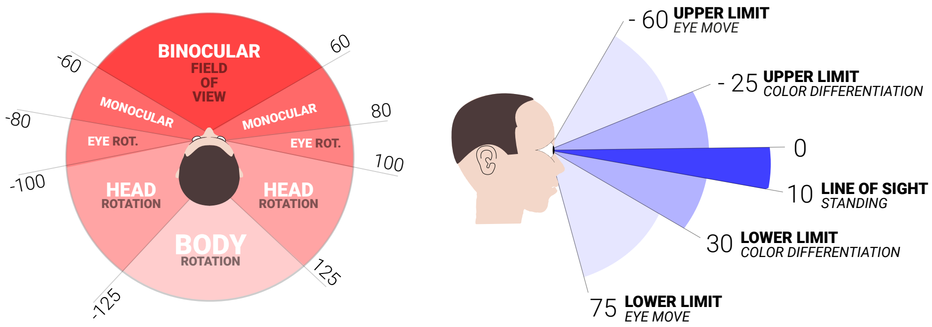

Figure 2.6: Left: horizontal field of view. Right: vertical field of view.

head rotation is called the field of regard. A wide FoV without eye or head rotation was an important missing component of consumer HMDs in the 90s. Current VR displays offer a FoV between 100° (Google Daydream, Sony Playstation VR) and 110° (HTC Vive, Oculus Rift); the Starbreeze StarVR prototype offers up to 210°. The vertical FoV is limited physically by the forehead when looking up to 60° and down to 75°, amounting to about 135° for the vertical FoV [280].

Depending on the *field of view (FoV)* available, VR displays can be classified as fully or partially immersive. Full and partial immersion in VR are fundamentally different user experiences. Partial immersion supports the feeling of viewing a VE, while full immersion supports the feeling of being [270] in that environment. The user's potential for immersion in a VE is often measured from the FoV, which describes how much can be covered from the user's point of view without head movement. More precisely, a wide FoV contributes to the user's sense of presence and reduces motion sickness in a VE [175], while limited FoV can have a negative influence on task performance [231]. Nonetheless, distance perception in VEs is not caused by a limited FoV of the HMD [147].

Refresh and Frame Rates

The term *screen refresh rate* is widely used among casual gamers and is usually given in *Hz*. It refers to how quickly a display can change its content over a period of time. The idea is that the more times the screen is refreshed, the smoother the video and the less flicker. The most common refresh rate for today's TVs is 60 Hz for NTSC-based systems and 50 Hz for PAL-based systems. Modern TVs operate at more than 100 Hz, which can be classified as flicker-free. This is not to be confused with the *video frame rate*, which measures how often (per second) a video source can transfer an entire frame of new data to a display. For comparison: Movies generally run at a uniform speed of 24 frames per second (fps). To put it simply, the more frames used in a given second, the smoother and clearer the motion will be. Since VR is supposed to achieve a certain degree of immersion and realism, a clear and 'lifelike' motion can be crucial for the overall experience. So far it seems that 60 fps per eye might be the minimum [130]. In general, "real-time" means interaction at a frame rate that ensures that images move smoothly as the view direction changes. The minimum that is considered to be real-time might be as low as 10 fps, or as high as 30 fps per screen [130].

Latency

In the context of VR, latency is the time it takes to update the display in response to a change in head position or orientation [129]. If the user's head moves in one direction while the eyes fixate a stationary object in the VE, the image has to be shifted immediately to the opposite direction. Otherwise, the actually stationary object seems to be moving, if the eyes remain fixed on it. With the early VR systems, latency was identified as a big issue and contributor to motion sickness [130].

In the last decades, many computer graphics researchers developed different strategies to both reduce the latency and to minimize the side effects of any remaining latency. With the development of multi-GPU technology, rasterization for the left and right eye can be done in parallel, using one GPU for each [307]. Moreover, mesh simplification algorithms (e.g. level-of-detail or vertex clustering) can be used to reduce the 3D model complexity in the VE while retaining the most important structures [44]. In the vertex clustering algorithm, clusters are generated by creating a uniform 3D grid wherein the vertices are mapped into. Further approaches to improve rendering performance are stencil buffer or multi-resolution shading by exploiting the shape and distortion due to the lens in a HMD [125].

Fortunately ,and thanks to the latest generation of sensing, GPU and display technology, latency is no longer the major issue in most VR systems [250]. It may be around 15 to $25ms$ or down to about $7ms$ [4] for optimally calibrated systems using high-end hardware, which can even be compensated for by predictive methods in the tracking system.

2.2.3 Tracking Systems

As already mentioned, immersive VR requires a computer-generated VE to respond to user movements. In its simplest form, this requires the ability to track a user's head for orientation and possibly positioning in the VE [13]. The tracking system is therefore another main component of VR systems. There are different types of tracking systems used in VR systems, but they all have a few things in common. The majority have a *transmitter* device that generates a *signal*, a *sensor* that detects the signal, and a *controller* that processes the signal and sends information to the *computer*.

Sensor

Most common systems require the user to carry the emitters (or trackers) themselves and to be surrounded by sensors connected to the environment, such as the Oculus sensors [158]. Another approach is the Lighthouse principle used by the HTC Vive [220], where two base stations (i.e. the sensors) emit time-controlled infrared pulses with 60 pulses/second, which are then recorded by the HMD and the controllers. This wireless synchronization reduces the number of cables between sensors, computer and HMD, making the base stations more practical, compared to the USB Oculus sensors.

Signal

The signals sent to the sensors are divided into different signal types [201]: electromagnetic (mainly used in medicine), acoustic (more commonly used in underwater research), mechanical (e.g. Binocular Omni-Oriented Monitor (BOOM) [40]) and optical signals [250]. Electromagnetic signals are disadvantaged if nearby magnetic fields exist, which interfere with them, like

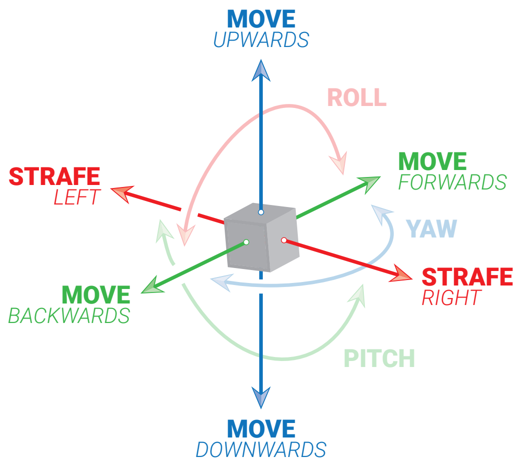

Figure 2.7: Degrees of Freedom (DoF)

human bodies or metals. Acoustic signals have many disadvantages, like the slowly moving sound, which can be strongly influenced by temperature and humidity, leading to a slower update rate. Although the update rate for mechanical signals is rather high, the user's freedom of movement is restricted, as with for example the BOOM device [40], which is fixed to the ceiling or mounted on a trolley.

Because of the others' disadvantages, it is mainly optical tracking systems that are used in current VR systems. Here, light is used to measure the position and orientation of an object or the user. The signal transmitter in an optical device typically consists of a set of infrared LEDs or other reflectors. The sensors are cameras that can detect the emitted infrared light and send information to the system's processing unit. However, the main benefit over other signals is the fast update rate of the optical signal, which means that latency problems are minimized and real-time interaction is possible. In contrast, the line of sight between sensors and transmitter can be obscured, which interferes with the tracking process. In addition, ambient light or infrared radiation can make a system less effective.

Finally, the speed of the program with which the user's position and viewing direction is converted into visual data is decisive for immersing oneself in the virtual world. Users should at least be able to look around, even in menus, during pauses in the game or in interim sequences [29]. In order to recognize the user's position and convert it into data, the VR system needs tracking technology that is as accurate as possible. Mobile VR systems cannot use external camera tracking because they cannot be deployed at a fixed location – but some new standalone portable VR technologies, such as the Oculus Quest or HTC Vive Focus, use inside-out tracking without external sensors. In addition to the typical accelerometers and gyroscopes, this method uses several ultra-wide-angle cameras that are embedded in the front side of the VR glasses to scan the room. According to Oculus, this "arena-scale" method allows a tracking area of more than 100 square meters, which is enormous compared to the areas of the current high-end VR systems like HTC Vive ($\sim 20m^2$) or Oculus Rift ($\sim 15m^2$).

2.2.4 Input Devices for Virtual Reality

Input devices are used to select and manipulate objects or to navigate in VEs, more precisely to send control instructions to an operating system. They enable the user to control a variety of degrees of freedom (DOF) and can be divided into standard input devices, trackers, and gesture interfaces [29]. Standard input devices include mice and keyboards, joysticks or gamepads. These are mainly 2-DOF input devices for controlling a cursor on a screen. Tracked devices require at least three DOF for their orientation, whereas the Microsoft Kinect Skeleton Tracking even provides six DOF per body joint. For most interactive VR experiences, one or more 6-DOF hand-held controllers are often the best choice. For simpler tasks that require only navigation or viewing direction control, i.e. no direct interaction with the VE, non-hand input such as pure head pointing is often used.

Non-Hand Input Devices

This section describes VR input by hands-free method, such as head and eye tracking, treadmills and speech (see Figure 2.8).

Head and Positional Tracking The HMD itself serves as input as well as an output device for the general VR systems. In portable VR systems,

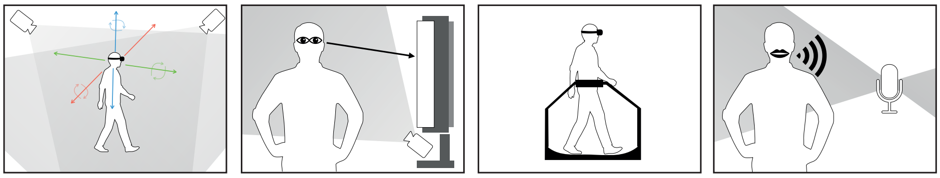

Figure 2.8: Non-hand input devices, from left to right: head and positional tracking, eye tracking, treadmill, and speech input.

internal gyroscope sensors provide sufficient information for the alignment of the head and allow the user to at least look around in the virtual world. The common form of head tracking interaction is pointing by looking or head pointing [130], which is often used for text input [335, 286]. The selection is then confirmed by a user-controlled trigger (button or dwelling time), as described in Section 2.2.4. More powerful VR systems, such as HTC Vive or Oculus Rift, also support positional tracking, which allows body movement. Full-body tracking extends head and positional tracking, either absolute and optical with Microsoft Kinect or a OptiTrack body-suit, but also purely relative with several IMU sensors per body joint.

Eye Tracking Another emerging tracking technology that is becoming more and more interesting for VR systems is eye tracking, e.g. for sports training research [77]. However, eye tracking for VR plays a decisive role in rendering and is already used to achieve better image quality and performance with the help of foveated rendering [230], e.g. by improving camera movements and blurring effects. Another use case is the patent for “using eye tracking to detect expressions in VR” [113].

Treadmills Immersive VR-specific devices that do not require hand input are on their way to the end user. An early example of a world-grounded non-hand form of input specifically designed for VR is Disney’s Aladdin Magic Carpet, which provides an intuitive physical interface for navigating in the VE [233]. In recent years, further VR prototypes have been used for balance training in psychology [135]. Another device is the omnidirectional treadmill [63], which consists of a giant treadmill with many smaller, vertical treadmills attached to its surface. This allows the user to walk virtually unlimited distances in all directions. Medina et al [198] have introduced another useful treadmill device. This human-sized hamster ball suffers from the fact that it cannot reproduce exactly the inertial feedback people experience when walking in the real world.

Speech Another powerful sensor for hands-free interaction is the microphone, which can be used to interact with the VE by speech. Speech input is a very comfortable and natural input modality, and there has been much research in the last decades in the field of speech recognition, with major companies like Microsoft and Google[15] bringing services to market. With the increasing amount of available data and computational resources, it becomes more and more accurate [263, 260].

In general, speech input is a nice addition to other input devices. This makes it a natural way to combine different input modes to form a more cohesive and natural interface. Speech input, when working properly, can be a valuable tool in applications in VEs, especially when hands-free interaction is required, e.g. when both hands of the user are occupied. Speech input is associated with the question of what methodology to use or where the microphone should or can be placed. Ideally, a large-capacity microphone would be the best solution so that the user does not have to wear a headset or orient himself to the position of the microphone. However, noises from other people or machines in the room may be recorded unintentionally, or the system may not know when to listen to the user's voice and when not to. Often a user talks to another person without intending to issue voice commands (see the Alexa issue), but the application "thinks" that the user is talking to it. Push-to-talk would be a solution here, but could influence the naturalness of speech and ultimately the cognitive load of the user [160].

Hand Input Devices

The human hand is an ideal direct manipulation input device for VR [127]. In this section, we discuss tracked and non-tracked hand-held controllers, hand-worn controllers, hand and finger tracking, as well as smartphones.

[15] https://cloud.google.com/speech/

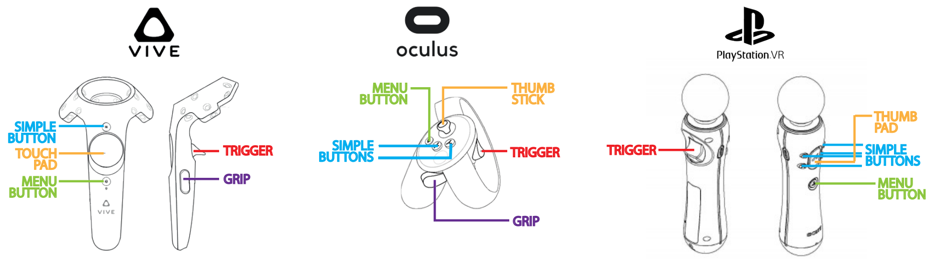

Figure 2.9: Tracked hand-held companion controllers, from left to right: HTC Vive[19], Oculus Touch[20], and Playstation Move[21].

Hand-held Controllers Nowadays, the most common form of interaction in virtual worlds is with non-tracked hand-held controllers such as gamepads or the traditional keyboard and mouse. Whether for game consoles, PCs or smartphones, familiarity and comfort are the clear advantages over novel VR-specific controllers. Here, the majority of game consoles use wireless non-tracked hand-held controllers with directional pads (d-pad) or thumb sticks. Companion controllers like the Oculus Touch or Vive controllers are mainly based on the common gamepad designs, but can be used bi-manually. The Oculus Touch even looks and feels like they have just cut through an XBOX controller (see Figure 2.9). Hence, companion controllers are usually equipped with simple buttons, d-pads, and trigger buttons. But the Vive controllers also have grip buttons, which are triggered when the controller is squeezed. Another difference and advantage of the VR companion controllers is that they are usually tracked and thus offer at least three DOF for pointing (Oculus Go or Google Daydream) or six DOF (Oculus Touch or HTC Vive) per controller, which already offers more complex spatial interactions, such as object manipulation or spatial gestures.

Haptics in Virtual Reality In the last decades, researchers have found that haptic feedback can enhance the UX and immersion in VR significantly [305]. The field of haptic technology for VR can be divided into active and passive haptics [25]. Participants in a low-cost VR training environment for military

[19] `https://www.vive.com/us/accessory/controller/`

[20] `https://www.oculus.com/rift/accessories/`

[21] `https://www.playstation.com/en-us/explore/accessories/`

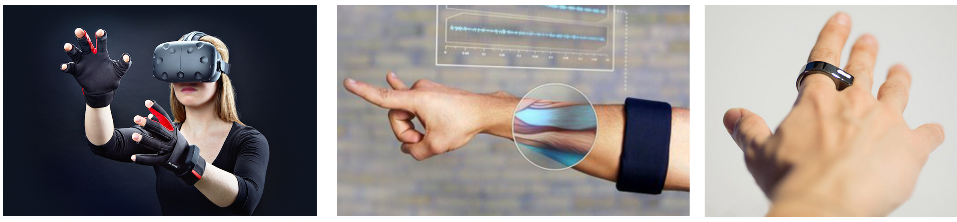

Figure 2.10: Hand-worn controllers, from left to right: Manus VR Glove[25], Myo Band[26], and Scroll Ring[27].

and emergency personnel made fewer procedural errors and completed tasks more rapidly with the addition of active haptic feedback [133]. Proxies with passive haptics provide tactile and kinesthetic feedback, whereas active haptics (e.g. gloves or gamepads) have to produce these effects using sophisticated algorithms and actuators. Training and simulation applications have shown that actually holding something during the interaction is more comfortable than holding nothing and just feeling vibrotactile feedback [51].

Although not directly related to haptics in VEs, Swindells et al. [302] introduced the TorqueBar, as a device for dynamic ungrounded kinesthetic feedback. The TorqueBar can dynamically change its physical properties to modify its center-of-mass location during runtime, controlled by a computer. In this context, Zenner et al. [339] introduced *Shifty* based on their concept of Dynamic Passive Haptic Feedback (DPHF) for VR. Their weight-shifting prototype is able to automatically adapt passive haptic feedback in order to simulate different virtual weights and lengths of an object by changing its internal weight distribution. Another example of a tracked hand-held haptic controller for VR is the so-called Haptic Revolver [325]. Here, customizable wheels are used to simulate different textures or shapes under the user's fingertip, so that the user can feel different virtual surfaces.

Hand-worn Controllers Among the hand-worn controllers are gloves, muscle tension sensors (e.g. Myo Band), and even rings. Many still believe that gloves are the ultimate VR interface, as they are theoretically superior

[22] `https://bit.ly/2HmE3Zv`

[23] `https://bit.ly/2EQTVl3`

[24] `https://bit.ly/2TsNEEv`

to all other controllers and especially optically tracked devices, because no free FoV or optimal lighting conditions are necessary. Just like bare hands, gloves also have the advantage that hands and fingers can interact fully with other physical objects in the real world. This increases the comfort of the interaction and reduces the danger of the gorilla-arm effect [160].

But why have gloves still not made it to market maturity despite many advantages and extensive research over the last decades? One reason could be the missing and inconsistent recognition of even a few gestures. It is still a challenge to achieve a constant accuracy of finger guidance and not require constant recalibration when the glove is moving on the hand. Gloves must also be put on, taken off and worn, which can lead to discomfort and sweaty hands. Another phenomenon is the danger of social resistance to wearing extra gloves [130], which was also observed after the introduction of Google Glass and is still a 'deal-breaker' for head-mounted eye tracking. In summary, although hand-worn devices are in theory the safer and better choice for VR applications, optical tracking of the user's hands and fingers is preferred, at the present time.

Hand and Finger Tracking The optical tracking of hands and fingers enables real-time interpretation of hand and finger movements to process them into concrete user input. Although hand- or finger-based techniques do not require tracked hand-held or hand-worn controllers, solid tracking is still a prerequisite [248]. When the Microsoft Kinect[28] was introduced as another spatial input device, the interaction space was extended by a mass-market 3D tracking solution. The Microsoft Kinect is a depth sensing input device and was part of the Microsoft XBOX game console. In contrast to mobile devices and Nintendo Wiimote, the Kinect is peripheral-free. Further, it provides a joint skeleton of the user, which is the result of processing depth sensor, infrared sensor and microphone input data. This joint skeleton consists of the 3D positions and orientations of the user's hands, shoulders, head, etc., and was mainly used to enable head pointing and hand interaction in 3DUI research [66]. For finger tracking, we would recommend

[28] https://developer.microsoft.com/de-de/windows/kinect

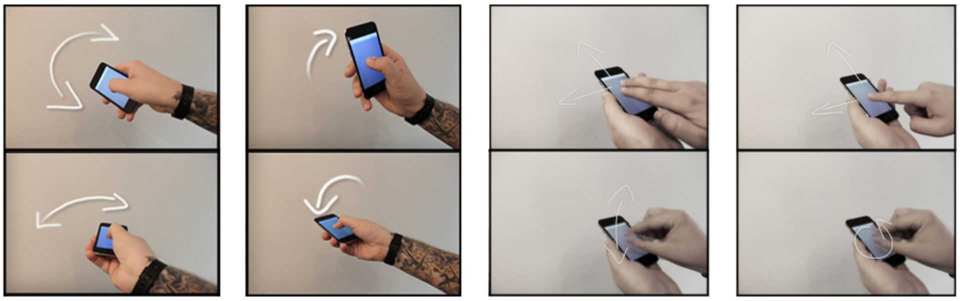

Figure 2.11: Our 3D navigation techniques using smartphone sensors [61], from left to right: tilt, two-finger pan and pinch, and one-finger pan and two-finger rotate.

to use the Leap Motion[29] mounted on the HMD as it is the currently most affordable and available method for finger input and visualization, and because it is much less invasive compared to gloves. Although the Leap Motion is a low-cost and high-efficiency device, it is worth mentioning that it has technical limitations [323] and could therefore have a negative impact on performance and preference compared to other devices. To overcome this, the Leap Motion should be used for all test cases, so that it can be disregarded as an issue.

Smartphone and Tablets The use of mobile devices such as smartphones and tablets for input is becoming increasingly popular as most users are already familiar with their use and metaphors. Bauer and Ebert [16] evaluated their applicability in the field of VR and suggest that for text input keystrokes should be combined with the DOF of a mobile hand-held device. Despite the tracking issues of optical tracking, direct hand manipulation remains the most natural and efficient input method for humans in 3DUIs. By equipping the dominant hand with a 3-DOF sensor, e.g. the accelerometer or gyroscope of a mobile device, its position or orientation can be mapped onto the position or orientation of a virtual object in the VE. We introduced and evaluated 3D interaction techniques using a mobile device (see Figure 2.11) and Kinect separately on a large-scale stereoscopic display [283, 61]. We created the illusion that the user is able to move or rotate this object using

[29] `https://www.leapmotion.com/`

his own hand, which leads to a natural and intuitive interaction within immersive worlds. Furthermore, the use of a physical mobile device serves as a passive haptic prop to support the user's spatial orientation and control (cf. passive real-world interface properties [115]).

2.2.5 State-of-the-art Virtual Reality Hardware

VR consumer hardware has been available for several years and is now affordable for the general public, so the user can enter VR even at home with a PC (Oculus Rift, HTC Vive), game console (Playstation VR), all-in-one system (Oculus Go) or even with a smartphone (Samsung Gear VR, Google Daydream or Cardboard). However, the wide variety of devices often makes us despair about these questions: Which are currently the best VR glasses on the market and what are the differences and commonalities? And how can they be classified? Overall, at the current state of technology "affordable" does not directly mean that HMDs are cheap, and a purchase must be well considered, because PC systems need VR-ready hardware and even casual VR systems, like Google Cardboard or Samsung Gear VR, require a VR-enabled smartphone.

Until 2017, VR HMDs could simply be categorized in two ways: portable (or mobile and wireless) or not. With the advent of the first standalone (or all-in-one) VR HMDs, such as the Oculus Go, a distinction has to be made here, because these glasses no longer require an external computing node and work on a completely self-contained basis. These all-in-one VR systems are typically portable and battery-powered, and work with reused smartphones. So they are even more affordable and convenient than the smartphone VR systems. Finally, they allow the users to try their first VR experiences without a large investment in hardware. But, the low prices are associated with limitations, and come at the expense of quality. In Table 2.1, we illustrate the differences between the state-of-the-art VR HMDs regarding comfort, display, and interactivity aspects.

Besides obviously worse graphics compared to high-end VR-ready PCs, the major drawback of portable VR is that only the head orientation is captured by internal sensors without positional tracking. Today's smartphones

	Google Cardboard	Google Daydream	Samsung Gear	Playstation VR	HTC Vive	HTC Vive Pro	HTC Vive Focus	Oculus Go	Oculus Rift	Oculus Quest
pictures										
release	Jun-14	Nov-16	Nov-15	Oct-16	Apr-16	Apr-18	?	May-18	Mar-16	?
price ($)	19	109	99	299	599	1089	599	199	399	399

Figure 2.12: State-of-the-art VR systems[30], order by manufacturer.

have several built-in sensors that deliver multiple-DOF data, like acceleration relative to free fall or absolute orientation measured by gyroscopes, but with a certain risk of jitter [177]. Here, the calculated device orientation serves as a head rotation, so that the user can look around in the VE.

A step further towards mobile and portable high-end devices are the yet unpublished Oculus Quest[31], which will support inside-out tracking, i.e. tracking of the head and two controllers. These are only behind the previous non-mobile VR systems, such as Rift or Vive, in terms of maximum graphics and computing power, due to external graphics power. But even those *room-scale* high-end VR systems will soon no longer be called "non-mobile" due to wireless modules docked to the HMD, but still not portable. In summary, current VR systems can be classified as:

- **Portable** VR
  - **Smartphone** VR (e.g. Samsung Gear VR or Google Daydream)
  - **All-in-One** VR (e.g. Oculus Go or Quest, HTC Vive Focus)
- **Stationary** VR
  - **Basic** VR (e.g. Oculus Rift, HTC Vive)
  - **Advanced** VR (e.g. HTC Vive Pro)

Overall, the current trend in VR hardware is moving towards mobile consumer hardware with a focus on entertainment (movies & games), educational activities (virtual classroom) and virtual shopping environments. But contrary to this, new flagship devices, like the HTC Vive Pro, are suited

[30] Sources from `https://versus.com/` (26.02.2019)
[31] `https://www.oculus.com/quest/`

Comfort / HMD	HTC Vive Focus	Google Cardboard	Google Daydream	Oculus Go	Samsung Gear	HTC Vive Pro	Oculus Rift	Oculus Quest	HTC Vive	Playstation VR
weight (grams)	?	<150	261	470	345	555	470	570	555	610
wireless	yes	yes	yes	yes	yes	(yes)	no	yes	(yes)	no
adjustable lenses	no	no	no	no	no	yes	yes	?	yes	yes
integrated audio	yes	no	no	yes	no	yes	yes	yes	no	no
external computing node	no	no	no	no	no	yes*	yes*	no	yes*	yes**

Display / HMD	Oculus Quest	HTC Vive Pro	HTC Vive Focus	Oculus Rift	HTC Vive	Oculus Go	Playstation VR	Google Daydream	Samsung Gear	Google Cardboard
resolution (w)	3200	2880	2880	2160	2160	2560	1920	?	?	?
resolution (h)	1440	1600	1600	1200	1200	1440	1080	?	?	?
refresh rate (Hz)	72	90	75	90	90	60	120	?	?	?
field of view ($^\circ$)	?	110	110	110	110	100	100	100	101	90

Interactivity / HMD	Oculus Quest	HTC Vive Pro	Oculus Rift	HTC Vive	Playstation VR	HTC Vive Focus	Google Daydream	Oculus Go	Samsung Gear	Google Cardboard
orientational tracking	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
positional tracking	yes	yes	yes	yes	yes	no	no	no	no	no
tracking area (m^2)	>100	∼20	∼8	∼20	∼5	-	-	-	-	-
controller (DOF)	2x 6	2x 6	2x 6	2x 6	2x 6	1x 3	1x 3	1x 3	-	-
trigger button	2	1	2	1	1	1	-	1	-	-
grip button	-	2	-	2	-	-	-	-	-	-
simple button	2	1	2	1	4	1	1	1	-	1
directional pad	yes	yes	yes	yes	no	yes	yes	yes	yes	no
bi-manual input	yes	yes	yes	yes	yes	no	no	no	no	no

Table 2.1: HMD comparison charts with regard to the categories: *comfort*, *display*, and *interactivity*. The order of the HMDs per table reflects a ranking for the corresponding category, i.e. descending from left to right.

to the industry sector and more professional use. For example, the Vive Pro helmet is equipped with overhead headphones, two front cameras and two microphones. With the higher resolution of 2880×1600 compared to the usual 2160×1200, as well as the availability of a wireless adapter, HTC is taking a further step towards high-end advanced VR technology.

2.3 Human Factors in Virtual Reality

The main issues of VR are caused by bad design or ergonomics affecting people in different ways [130], and the immersion and presence of VR experiences [277]. This section deals with human factors related to VR, how individual terms are defined, and how these factors can be influenced.

2.3.1 Usability and User Experience

Since the beginning of the 1980s, the field of human-computer interaction (HCI) has been concerned with the development of conventions and rules for the design of interactive user interfaces (UIs) [185]. The terms *usability* and *user experience (UX)* are standardized and can be found in DIN EN ISO 9241. Here, usability is described as the extent to which a system or method can be used by certain users in a certain context to achieve certain goals effectively, efficiently and satisfactorily. Thus, the term usability stands for ease of use and user friendliness. Good usability is usually not perceived explicitly, but bad usability is. Therefore, it is important to ensure good usability for all products with a human-machine interface (HMI): whether software, websites, interaction techniques or complex systems for machine control in the workplace, they all benefit from good usability.

According to ISO 9241-210, UX includes the aspects of usability and expands it by aesthetic and emotional factors, such as an appealing, novel and desirable design, aspects of confidence building and stimulation, or fun during use [157]. This approach encompasses the entire UX when using a system. The users should not only get to the goal as efficiently as possible, i.e. quickly and smoothly, but also experience positive and stimulating feelings when using the system. In our studies and prototypes, we pursue a holistic approach, whether in the evaluation or optimization of existing or novel interactive VR systems.

VR contains a multitude of design elements that differ from traditional methods for designing positive UX. While many facets of digital design can be transferred to VR, it has a completely different set of requirements for interaction and interface between the user, the system and the VE than traditional designers and developers are used to from desktop or mobile devices [161]. Traditional digital products are designed to interact through the touch screen of a mobile device or the mouse and keyboard of a desktop or laptop. In contrast, VR applications use a combination of the user's body and senses, as well as proprietary or customized control devices for interaction [56]. There are, of course, analogies to traditional UIs and patterns, but the scope of this thesis is the application of traditional heuristics compared to previously uncommon guidelines. In the context of a VR system, reality is simulated through the use of technology enabling people to experience things that do not exist in the real world. Therefore, it is essential to understand how to adapt our senses and create a real experience.

Today, the creators of new VR experiences often try to avoid the main problems of VR by relying on stationary users to be able to offer comfort and safety, i.e. users either sitting or standing. Especially in VE where the user

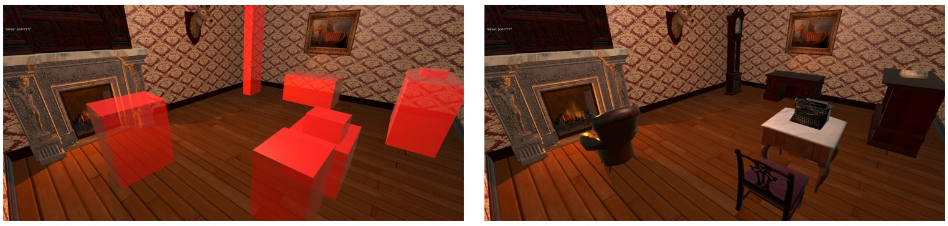

Figure 2.13: These figures show the concept by Garcia et al. [89] of substituting physical objects such as chairs, tables or other furniture with virtual replacements, which fit better into the narrative.

can move freely, UX designers must ensure that users do not risk physical damage to themselves, their environment or nearby furniture. Although there are approaches to substitute physical obstacles with virtual objects (see Figure 2.13), this includes the risks of overuse with VR, which can have a negative impact on a user's physical, health or emotional well-being [130]. In addition, users in VR must feel present in this environment and empowered to understand the new rules it contains. Fully immersive VR systems and the feeling of presence also require that users in VR be given helpful guidance to understand the VE in which they find themselves.

VR applications should therefore allow users not only to passively experience the VE, but also to interact with it. The ability to intuitively pick up, move, shape and create objects is a prerequisite for positive UX in VR [161, 130]. The VR software should be able to precisely track the head and body movements of the user to ensure the highest possible comfort. They should be designed to have a 360° view and allow free movement for a highly interactive experience. Ultimately, the biggest challenge for designers is to provide VR-specific UIs and interactions, such as the issue of navigation (see Section 2.4.3), since the real environment in which the user is in the tracking area is limited, as opposed to the infinite virtual world. Otherwise, if the technical or logistic requirements are not given, e.g. if there is no positional tracking due to inexpensive VR hardware [282], application-specific design guidelines or design spaces are needed. Further discussion on using consumer VR hardware in the context of shopping and system control in VR is described in Chapters 4 and 5.

But what can be fundamentally more important for the UX in VR applications? Is it better to offer the user the freedom to move freely in the VE by natural walking (see Section 2.4.3) and grabbing the object on the one hand, or could teleportation and virtual tractor beams have a more positive effect on the UX within the VE due to the considerably lower physical effort? We investigated these questions in various studies in VR shopping environments (see Chapter 5). In summary, whether a VR experience is developed for a creative design application, an entertainment experience, a gaming platform or something completely different, design and evaluation of UX will play a key role in presenting the VR medium to the user. Finally, attention to detail is crucial when it comes to experience. Basically, designers of VR systems should focus on UX rather on the technology itself [130].

2.3.2 Task Workload

Task workload is the amount of work and effort required for an individual to accomplish a particular task and distinguishes the actual workload and subjective perception of the workload [132]. A typical example of a task in a HMI is text input using a specific input device. Again, a distinction is made between the effort itself, i.e. what the user has to do to successfully complete the task, and the difficulty of the work or task. The methodology basically requires the identification of all aspects of the HMI associated with the tasks of the individual operators, such as mental, physical and temporal demand, as well as frustration, subjective performance and effort [107]. This information can be used by the developers and designers of VR systems and interfaces to develop concepts that are necessary for the interaction of selected system functions or tasks. The associated workload can have a decisive influence on the design and performance of human-computer systems and is therefore listed in this paper as one of the four important human factors in VR. The assessment and evaluation of the workload is especially important when designing new or further developments of existing systems or interaction concepts [107]. Task-specific overload of the user by too-high physical or temporal demand, could be identified and eliminated earlier.

[32] Adapted from `http://www.trimetricsphysio.com/vestibular-rehabilitation/`.

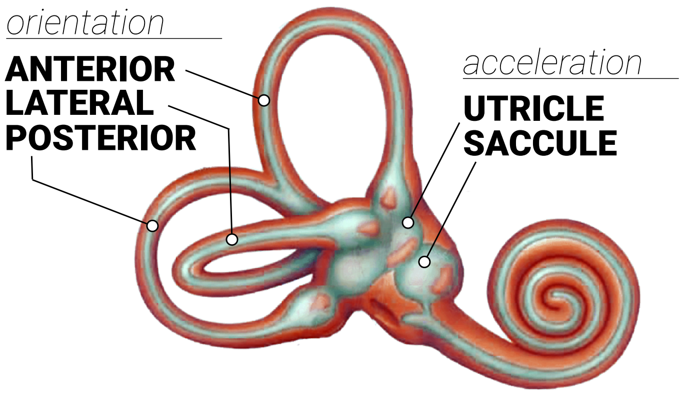

Figure 2.14: The Vestibular System[32]

2.3.3 Motion Sickness

Sickness has been a problem since the emergence and first use of VR systems and can lead to a number of symptoms including nausea, disorientation, headaches, sweating and eye strain [162]. There are several discussions about the term to be used in relation to sickness in VR systems. One approach would be to relate the underlying physiological causes and symptoms, such as motion sickness, simulator sickness, or cyber sickness [140]. Admittedly, the problem is quite complex, since the experiences with sickness in VR can vary greatly from person to person, the technologies used, the design of the environment, and the tasks to be performed [134]. Motion, simulator and cyber sickness are similar in many ways, but differ in their causes. Cyber sickness is considered as a subtype of motion sickness because it does not affect the balance system and is triggered only by visual stimuli. Simulator sickness is also a type of motion sickness that pilots experience after using a flight simulator. Hence, we use the term *motion sickness* in the further course of the work.

Causes and Consequences of Motion Sickness

To develop an overview of motion sickness, an understanding of its origin must first be gained. Motion sickness is related to the vestibular and visual system. The vestibular system is located in the non-acoustic part of the inner ear and consists of three semicircular arcade channels (angular acceleration) and two organs (linear acceleration) [95]. The three semicircular arcades (anterior, lateral, posterior) are adapted to the three possible dimensions or DOF for movement (see Figure 2.14). The other two organs (saccule, utricle) recognize both vertical and horizontal acceleration. This system is therefore comparable with the gyroscope and acceleration sensors in commercially available smartphones. *Vection* is the phenomenon when the vestibular system does not perceive any movement stimuli, while the visual system does. An example is often seen when trains are stationary at the station and a train on the next track starts to accelerate. Although there is no common and officially accepted theory, there are three useful and accepted theories from psychology that can explain the causes of motion sickness [163], as well as some individual factors:

Sensory Conflict Theory The *Sensory Conflict Theory* is the oldest and most accepted one of the theories relating to motion sickness [246]. It is based on two inconsistencies. First, the manipulated stimuli may not be exactly the same as those expected by the central nervous system and the brain. This discrepancy can occur, for example, with aliasing, too-low frame rate, optical distortion, or limited colors. Or second, the sensory systems of the human body that naturally capture the physical outside world (e.g. eyes and ears) transmit the corresponding neural signals, although no real stimuli exist. This appears if a vection is perceived, which cannot be confirmed by inertial forces transmitted through the vestibular system. Basically, the lower the discrepancy between visual and vestibular input, the lower the probability of sickness occurring [297]. Nevertheless, there is no explanation why some users eventually become sick and some do not.

Poison Theory The second theory, called the *Poison Theory*, assumes that motion sickness occurs evolutionarily [308]. Here, sickness is compared to the intake of poison, where physiological effects can affect the visual and vestibular systems. This often results in emptying of the stomach by vomiting. In VEs, misinterpreted stimulation can affect the visual and vestibular senses, such as a strong discrepancy between real and virtual motion [275]. According to this theory, the human body interprets this sickness as a result of poisoning, which can lead to nausea and vomiting. Again, there is no explanation for why only some participants feel nausea.

Postural Instability Theory The *Postural Instability Theory* focuses one of humans' primary behavioral goals, namely to maintain postural stability in the environment [251], i.e. minimizing uncontrolled movements. In general, people walk differently and more safely on a solid surface than on ice, which often leads to slipping and falling. Thus, they change their walking behavior on ice in order to maintain postural stability. However, if the surface on which a person moves abruptly changes, it leads in most cases to postural instability and the danger of falling. Postural instability can therefore cause motion sickness symptoms. The severity of the symptoms (e.g. dizziness) depends directly on the duration and severity of the instability.

Individual Factors There are other approaches and contributing factors to motion sickness in VEs, which cannot be assigned to the above-mentioned theories. First, various display and technology problems can cause motion sickness, such as inaccurate tracking systems [21], interaction lags [232], or flickering [108]. But technology continues to evolve, so many of these problems have already diminished and will almost disappear in the future. However, individual demographic factors can also play a role, such as age, gender and previous illnesses. So children between the ages of 2 and 12 are most susceptible, between 12 and 21 it decreases very fast, and at about 50 years motion sickness should play almost no role anymore [246]. But slight existing illnesses or discomfort can also cause motion sickness [139], like fatigue, insomnia, a hangover, indigestion, menstrual cycles, emotional stress, cold, ear infection or respiratory disease. In summary, persons under the age of 13 or with a medical history of relevant issues should avoid VR.

Reducing Motion Sickness

There are different methods for reducing motion sickness in VEs. Basically, it can be assumed that the fewer DOF, the less motion sickness occurs [292]. But it is not always necessary to control all six DOF regarding position and orientation, e.g. when moving in space with a controller two DOF are usually sufficient. However, for the lowest possible motion sickness rating when moving, it is recommended, if possible, to transfer the real movements in the real world directly to the virtual world [311]. Llorach et al. [178] were able to prove this again with regard to the Sensory Conflict Theory by means of a study in which they compared navigation in a VE with a positional tracking system and a game controller. Further possibilities of navigation in VR, such as teleportation, are described in more detail in Section 2.4.3. Infinite speed techniques (teleportation and animated teleportation box) have proven to cause less discomfort than linear motion [197]. Motion sickness during reading, which already occurs during real-world journeys, could be reduced by the Motion Sickness Prevention System (MSPS) of Miksch et al. [202]. This approach of "reading between the lines" could also be used for VR applications in general.

Another approach besides navigational aspects was presented by Whittinghill et al. [326], who displayed a virtual nose to reduce motion sickness. Lin et al. [173] presented a visual guiding avatar using rotational and translation cues to enhance user experiences in VR by improving motion simulator design. For pointing techniques, with head, eye or controller tracking, at least a cursor should be used as a fixation point, which was shown to reduce motion sickness by Kitazaki et al. [146]. In the TranSection game [170], the combination of head and hand tracking in a 3D virtual environment has led to lower motion sickness ratings. Moreover, recent research in motion sickness indicates that high velocity could amplify the symptoms [211, 304], e.g. driving on a highway causes stronger symptoms than driving in the city. Considering driving simulations, driving the vehicle leads to less motion sickness than being the passenger [267], which contradicts the non-trivial factor of a lower FoV of 90° of the driver compared to the potentially higher 180° FoV of the passenger [174]. Here, researchers presented a trade-off using a dynamically changing FoV and found a negative correlation between motion sickness and enjoyment [80, 168]. On the other hand, high motion sickness can be used, if it is used correctly, to produce pleasing and enjoyable experiences [314], e.g. in a VR rollercaster.

2.3.4 Immersion and Presence

This section addresses the confounding terms *immersion* and *presence*, which are described with the feeling of "being there" experienced in VEs.

The VE users dive in and are *immersed*, if they have the feeling that the virtual world surrounds them and has replaced the physical world as a frame of reference to some extent. In many discussions about the correct use of the two terms, both are often mixed up or confused, as with the sickness terminology. Here, Slater [277] offered a terminology that should clear up the confusion and avoid disputes about the essentials. He stated that the term "immersion" refers to what technology achieves from an objective point of view. The more a system provides feedback in all sensory modalities, i.e. visual, auditory or haptic, and precise tracking, the more it is immersive. Presence, on the other hand, is described as a human reaction to immersion. In the same immersive system, different people can experience different levels of presence, and also different immersive systems can lead to a similar presence for different people [277]. Therefore, presence and immersion are logically separable, but empirically strongly connected [278].

Hence, the goal of immersive VEs should be to let the users experience a computer-generated virtual world as if they were there. In addition to many realistic VR experiences, which also pursue other benefits in addition to pure entertainment, such as training and phobia therapy [262], immersion can potentially offer many other advantages. In this context, observations and comments by participants in a VR experiment show that immersive VR offers a different experience than interaction with 3D applications on desktop PCs or game consoles [199]. Furthermore, novice VR users report a strong feeling of presence when they experience immersive VR for the first time [30, 130]. Stereoscopic VR displays, compared to standard monoscopic desktop screens, selection and manipulation of a virtual object through direct interaction with the real hand, or that head movements changing the view of the virtual world without controlling the camera separately with one hand or controller, offer a unique experience especially for novice VR users [61]. Finally, Gruchalla [101] showed in an empirical study that users of an immersive system perform significantly better than using a

Figure 2.15: These figures show the desktop (left) and CAVE versions of Gruchalla's oil well path- planning application [101].

non-immersive system for a particular task, such as editing a path in an oil well planning application (see Figure 2.3.4). Such results are crucial for immersive VR applications beyond traditional areas for VR such as the military, therapy, training and entertainment.

One of the greatest and most intuitive advantages of immersion is a better spatial understanding in the VE. In the real world, we perceive a stable 3D environment, even if our eyes only process 2D projections. However, the human brain is able to reconstruct useful 3D scenes from these images using only depth cues such as stereopsis, motion parallax, perspective, and occlusion [30]. But immersive VR systems also provide many useful depth cues, such as stereo images, shading, and head tracking, allowing the user to achieve a better spatial understanding through higher levels of immersion compared to non-VR systems. This opens up countless new application areas such as collaboration tasks [216], e.g. in scientific visualization of large amounts of data, or virtual prototyping. Empirical studies investigated the effects of different components of immersive VR, and found that field of regard and stereoscopy did not affect the task performance in a manipulation task in combination [196], but they did individually [61].

Current research tries to identify as many potential benefits of immersive VR as possible. But we want to help other designers and developers to avoid costly or wasteful situations where a highly immersive and application-specific VR system would not be necessary. More precisely, head tracking alone could be sufficient as well, with a simple representation of the VE, in contrast to a room-scale setup with high-resolution 3D models and tracked hand-held controllers [282, 286]. Hence, it would be difficult to justify the effort and development complexity that full immersive VR requires. Therefore, we simplify our definition of immersive VR, so we can demonstrate the benefits of more practical systems and ultimately attract more applications and users for everyday VR. There is already a trend toward lower cost, cf. more consumer VR systems (see Section 2.2.5). We expect this trend to continue as we learn more about the advantages and disadvantages of VR-specific factors (see Chapter 3).

2.4 Interaction Concepts for Virtual Reality

Navigation and object manipulation in immersive VEs are universal interaction tasks. Even though they have been in the focus of research, there are still no universal and suitable solutions for VR-based environments and commodity hardware. In this thesis, we investigate whether VR should portray real-life experience in virtual form, or design a novel type of experience that builds on the affordances of the virtual medium. We thus created different VR concepts, based on *isomorphic* and *non-isomorphic* interaction techniques. In this context, isomorphism characterizes the mappings between actions in the real world and their effect in the virtual environment [243, 29, 340], e.g. moving the virtual character by moving the HMD or walking around in the tracking space. An isomorphic technique uses one-to-one mappings and is considered as the most natural approach. However, this is also accompanied by the risk of unfamiliarity, confusion, and misunderstandings for the user [209]. To overcome limitations in the tracking space or anatomical constraints, a non-isomorphic UI or interaction technique does not imitate the physical reality and allows users to select and manipulate objects using "supernatural" metaphors (e.g. ray-casting [161, 229]). Instead, interaction should be tailored to the immersive VEs, resulting in a "better" version of reality [296], e.g. bypassing physical limitations like gravity or arm reach.

2.4.1 Virtual Reality User Interfaces

User interfaces (UIs) are defined by the hardware and software that mediates the interaction between humans and computers. The UI of a system includes input and output devices such as mice, keyboards, monitors, and speakers, as well as software components such as menus, windows, toolbars, etc. [118]. Concerning intuitive VR UIs, Knöpfle and Voß [148] presented a VR UI to support experts in the automotive industry. They stated that a menu interface is essential for changing object properties and system control, e.g. switching between functions and modes. Their first prototype resembling a standard planar desktop interface resulted in high workload and low UX, which overshadowed the benefits of its simplicity, lower instrumentation and higher familiarity. Subsequently, after using a jog dial instead, the performance and preference could be improved, but not significantly.

3D vs. 2D

Previous research has taught us not to assume that transferring conventional 2D UIs into 3D VR environments would be accepted by the users [161]. When comparing 3D and 2D UIs, familiarity (e.g. from general desktop settings) and affordance (e.g. buttons and switches in a cockpit based on physical metaphors) need to be considered [56], in particular regarding user instrumentation. But one has to differentiate between virtual and physical familiarity, e.g. when designing for VR, which leads to the assumption that just transferring a given system from one medium to another might not be the optimal solution according to the user's preference. However, until it is possible to work in VR as precisely as in the real world, it is essential to come up with new ways of interaction which can support the users in achieving their goals.

Viewport and Interaction Distance

Hezel and Veron [112] stated that the human eyes' accommodation and convergence make it possible to comfortably display objects starting at a distance of about 0.25 meters. Furthermore, every display should have the same distance to the user, because of a performance drop when separating information in the visual field by depth [303]. In addition, Shupp et al. [273] examined the effects of viewport size and distance in the context of geospatial tasks, like searching or route tracing. Furthermore, the findings of Ha et al. [106] regarding workspaces in multi-display environments helped to

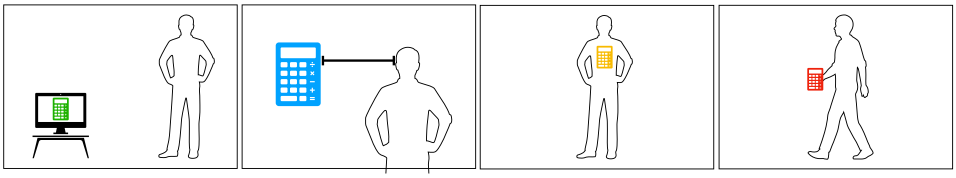

Figure 2.16: Four arrangement options for fixed UIs, from left to right: world-fixed, view-fixed, body-fixed, and hand-fixed [130, 161].

classify the following four arrangement options for UI and control elements in virtual environments [130, 161]: world-, view-, body- and hand-fixed (see Figure 2.16). An example of a world-fixed UI would be a menu interface floating in the air, and the user can look behind it or walk around it. They are mostly used for integrated controls or fixed to objects in the VE [106]. View-fixed UIs are suitable for menu interfaces [106], which should be available at any time or place, such as the main menu. Body- or hand-fixed UIs are often used to simulate displays affixed to some part of the user's body or hand [79], e.g. on a weapon to see the current amount of ammo.

Ens et al. [79] also estimated, based on NASA's Man-System Integration Standards[33], that interactive objects should be between 50 and 60 cm in front of the user and 70 to 80 cm away from his dominant side, depending on the arm's reach. Finally, arm reach has been reported as one of the biggest issues in 3DUIs, and continues to be, especially with selection of objects that are outside the user's area [59].

2.4.2 Selection and Manipulation

As mentioned, 3D input devices, unlike 2D input devices, easily support multiple DOF and offer more natural and intuitive interaction in many task scenarios. But 3D interaction requires more complex techniques, which consequently lead to a higher degree of user instrumentation and workload, such as increased physical and mental demand or frustration [100, 283], thus 3D mid-air selection achieved better results in VR when the participants were in a comfortable pose [181]. In HCI research, novel and improved interaction methods are constantly being developed to facilitate the user's

[33] https://msis.jsc.nasa.gov/

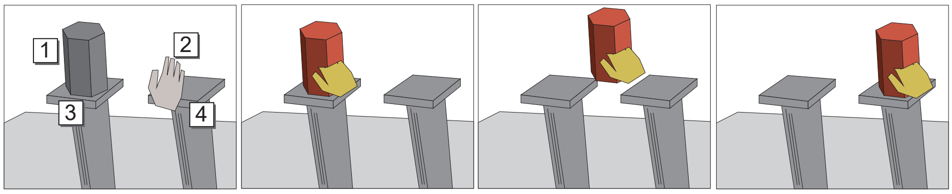

Figure 2.17: Virtual Hand technique, adapted from [242]. The initial state includes (1) the target object to be selected and manipulated, (2) virtual hand controlled by the hand-held motion controller, (3) initial position and (4) desired target position. The user selects the object by moving the virtual hand into it. The manipulation phase is then initiated by pressing a button on the controller. The object is translated and rotated by moving the virtual hand to the target position and finally releasing the button press to confirm.

work with a machine and improve the UX. The most important methods of interaction with the physical environment are touching or picking up objects, i.e. manipulating objects with the hands to influence the environment. This affects the quality of the overall 3DUI: if the user cannot efficiently select and manipulate objects or control the system in the virtual environment, then other challenging tasks cannot be performed. The canonical 3D manipulation tasks are defined by *selection*, *object manipulation* (position, rotation, scale), and manipulation of object or environment attributes through *system control* [29]. Virtual 3D manipulation tasks, like product interaction in a VR shop [287], combine target selection and manipulation. Accordingly, this section concentrates on selection and manipulation of objects in VEs.

Selection

Selection is the most fundamental universal interaction task and includes acquiring or identifying a particular object from the entire set of objects available, which is also called a target acquisition task [2]. The real-world counterpart of the virtual selection task is picking an object with a hand.

Grabbing An isomorphic variant for selecting and manipulating a product in our VR shop prototype is based on the Virtual Hand technique [242]. This concept utilizes a motion controller with a button, which triggers the interaction (see Figure 2.17). In a VR shopping apartment [287], the VR

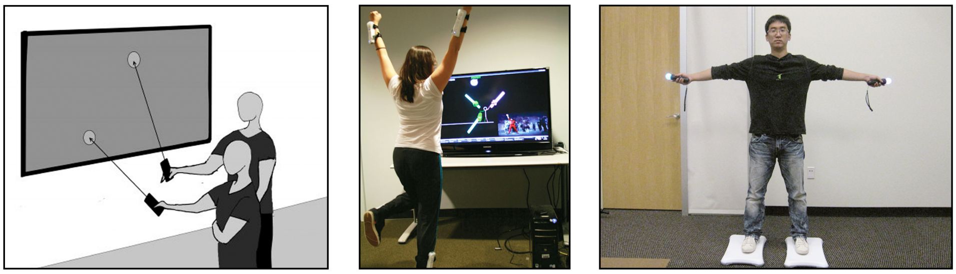

Figure 2.18: Pointing techniques. Left: Smartphone-based [237]. Center: Wiimote-based [331]. Right: Playstation Move-based [105].

controller is held by the user's dominant hand and the grip button triggers the interaction. This is a classical example for isomorphic object selection, although the major drawback is that if objects are out of reach, it requires additional movement by the user and can therefore have a negative impact regarding performance and preference. Here, the Go-Go technique has been developed to extend the virtual hand's reaching distance [241]. If the user's hand or a tracked hand-held controller is close to the body, the mapping between the physical and virtual hands' positions is one-to-one. But if the user moves the physical hand beyond a certain threshold, the mapping becomes non-linear and objects at larger distances can be grabbed.

Pointing The non-isomorphic counterpart to grabbing uses the concept of interaction by pointing in a direction. Direction selection is useful for navigation (e.g. for teleportation, see more in Section 2.4.3), object selection in a direction, or for object manipulation, e.g. to specify a desired object position. The most common form of direction selection is to look or point at the object of interest. This is familiar from real-world interaction and is usually done with the head or a hand-held controller. The basic idea for head pointing is to follow an imaginary ray from the point of view (or object of interest) through the VE [116]. Soojeong et al. [334] explored controller-less selection and claimed that gaze is one of the simplest methods, and has two types: instant and dwelling gaze. Instant gaze triggers object selection directly by looking at it without any confirmation. In contrast, by using dwelling gaze, the selection is triggered if the user looks at an object for

Figure 2.19: This figure illustrates the interaction techniques used for selection and manipulation of the products within our VR shopping apartment [287]. Left: Grabbing the object adapted from [242]. Right: Pointing and tractor beam adapted from [229, 27].

a few seconds. While instant gaze or pointing cause the "Midas touch problem" [126], dwelling gaze prevents it. According to Kim et al. [142], gaze- or pointing-based interaction has several benefits: naturalism, remote controllability, and easy accessibility.

The major drawback of selection by head pointing is that the user is forced to look in the same direction, and the object of interest cannot be occluded [236]. This is unlike controller pointing, which allows bi-manual pointing using both controllers or hands and can increase the task performance, e.g. for text typing in VR [286]. Moreover, it is designed for scenarios where controller tracking is available and extends a ray from each controller. It is actually not novel, because it was used for the Nintendo Wii via optical tracking using infrared. Here, the user controls the cursors' positions on the screen, simply pointing at the object to select it. An object gets selected and highlighted when the cursor (or ray) intersects with it. Pietroszek et al. [237] investigated the use of consumer devices for target selection by pointing in 3D. In particular, they showed that their approach using a smartphone can compete with the Wiimote-based [331] or Playstation Move-based [105] approaches with regard to selection time for smaller objects. Finally, those results indicate that user interaction with appropriate performance can be enabled without requiring specialized hardware or high-precision cameras.

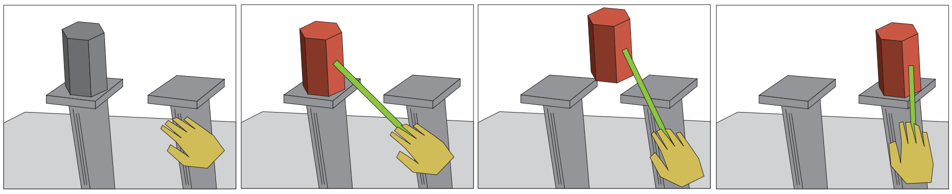

Figure 2.20: Hand-centered Object Manipulation Extending Ray-casting (HOMER) [27].

Manipulation

Manipulating objects in a 3D world involves 6-DOF control: three for positioning (x,y,z) and three for rotation (yaw, pitch, roll). During positioning, the 3D position of an object is changed. The real counterpart of positioning is moving an object from a start point to a destination point. Rotation changes the orientation of an object. The real counterpart of rotation is rotating an object from a start orientation to a destination orientation. Scaling is a task which, due to its simplicity, is often excluded from canonical manipulation tasks, since it is mostly controlled in one dimension.

Typical and common manipulation techniques are virtual hand, Hand-centered Object Manipulation Extending Ray-casting (HOMER) [27], scaled-world grab [206, 236] and world-in-miniature (WIM) [296]. HOMER [27] is a hybrid manipulation technique that combines ray-casting with hand-centered manipulation (see Figure 2.20). Here, the user points her hand at the object to select it and manipulates it with the virtual hand metaphor. In the scaled-world grab technique [206, 236], the whole virtual world is scaled according to the object manipulated by the user (see Figure 2.21). The user selects an object with an image-plane technique, i.e. the zoom level of the entire scene is adjusted in relation to the selected object. With WIM [296], the user does not interact directly with the actual environment (see Figure 2.21), but with a placeholder (here, a miniature model of the VE in the non-dominant hand of the user). Due to its naturalness, direct hand manipulation is the dominant means of interaction in virtual worlds.

One example VR scenario for a 3D object manipulation task could be a furniture arrangement application. As a common task in 3D is docking, we

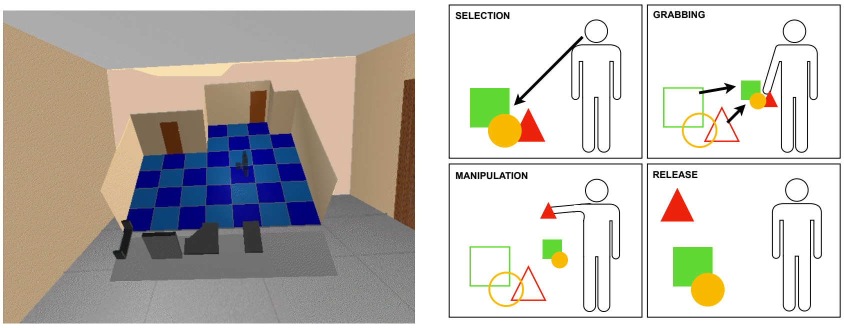

Figure 2.21: World-in-Miniature (WIM) [296] and Scaled-World Grab [206]

explored task performance and users' preference of mid-air hand interaction in a 3D docking task experiment using 3D furniture objects [284]. The experimental results showed that translation and rotation precision benefits from the use of a large projection, while participants preferred a HMD in terms of UX and task workload. Further, Liang et al. [172] investigated how mobile devices can be used for input to distant large 3D displays. In an exploratory study they asked participants to propose interactions for 3D tasks, and applied their findings to a prototypical application for 3D object manipulation. Martinet et al. [193] introduced a 3D manipulation technique based on separation of translation and rotation. It can take from 10–30 seconds to rotate 3D objects using current desktop interfaces and mice [117] – much slower than direct object manipulation, which takes between one and two seconds [321]. We adapted some of these concepts for the input technique on mobile devices [283], e.g. tilting the device to rotate an object.

People often use both hands in the physical world to cooperatively perform many tasks, e.g. text entry, cooking or playing music. In the last decades, researchers have explored the possibility of using both hands simultaneously in a computer interface, which is commonly known as bi-manual interaction. Buxton and Myers [43] showed, in a compound task, that a one-handed interface was inferior to a two-handed interface, which split the compound task into two subtasks that could be performed in parallel by both hands. In a target selection task, other researchers showed that using the non-dominant hand to control a virtual camera while the dominant hand performs the selection was 20% faster than one-handed [14].

2.4.3 Navigation and Travel

Generally, navigation and travel can be divided into four main categories: walking, steering, selection-based and manipulation-based [161]. We limit this discussion to approaches that resemble natural *walking* and *steering*, as well as the *teleportation* technique.

Natural and Virtual Walking

Walking is the traditional isometric mapping found in most immersive VEs. Here, a distinction is made between natural and virtual walking. Slater et al. [279] showed that non-expert users will have a greater sense of being in the VE if they move through the environment via *virtual walking* rather than flying (or sliding) along the floor using buttons on a controller (*push-button-fly*). Based on their findings, Usoh et al. [311] compared these two walking techniques with *natural walking* as a third condition. They found that natural walking is significantly better regarding simplicity, presence and ease of use. However, users often find navigation in VEs more difficult than in the real world. But as natural walking is possible only to a limited extent, due to the limited tracking space of VR systems, there are several approaches to improve virtual walking (see Figure 2.22), like detecting footfalls with a head accelerometer [311].

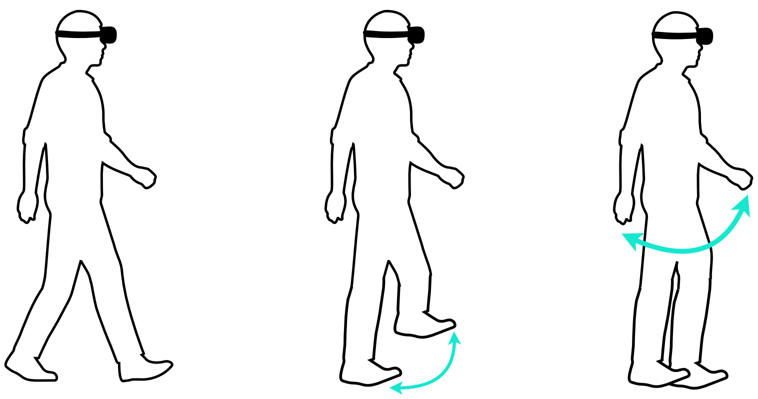

Figure 2.22: Examples of natural walking gestures adapted from [222], from left to right: fully natural, in-place, and arm-swinging.

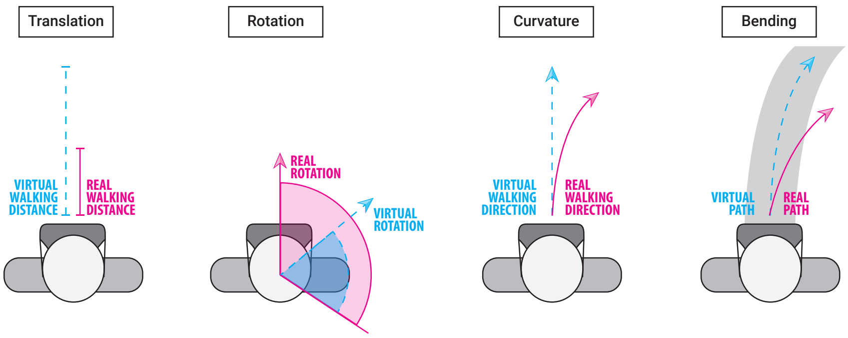

Figure 2.23: Viewpoint manipulation techniques used for redirected walking, adapted from [222]. From left to right: translation [123, 329], rotation [39, 131, 245], curvature [219, 293, 98], and bending [154, 155].

Steering Moreover, there are situations and environments in which real walking is impossible. Here, purely virtual travel techniques come into play. The most prominent and common metaphor is steering, which enables the user to move along a vector in world coordinates (e.g. flying a spaceship) or in a relative direction (e.g. move to the left). Spatial steering can be categorized as gaze-directed [205], hand-directed (pointing) [206], or torso- or lean-directed (e.g. Segway metaphor [61] or PenguFly [313]). Physical steering allows to travel through larger VEs using physical props, like cockpits [37], cycles [35] or specialized devices (e.g. Disney's VR Aladdin [233]).

Redirected Walking A promising approach to deal with limited space, because it includes natural walking, is the so-called redirected walking technique [245]. This approach is designed for small tracking spaces to enable the user to explore a larger VE by natural walking. Redirected walking includes the manipulation of the user's viewpoint (see Figure 2.23) by translation [123, 329] or rotation [39, 131, 245], as well as novel ways like using curvature [219, 293, 98] or bending [154, 155]. Furthermore, Peck et al. [234] explored redirection with distractors, which allowed the user to walk through larger VEs than the tracked space without predefined waypoints. The results indicate that participants would prefer relocation techniques with visual distractors over turning themselves by 180°, either with or without audio instructions.

Figure 2.24: Point & Teleport technique, adapted from Bozgeyikli et al [33].

Teleportation

Besides redirection techniques that manipulate the physical path of walking users, there are techniques that let users travel to a specific place in the VE (*relocation techniques*) [221]. The design space of possible relocation techniques ranges from mimicking real-world isomorphic travel (e.g. elevators or escalators [299]) to magical non-isomorphic travel (e.g. flying using a magic wand [49] or virtual companion [50, 336]).

Bolte et al. [24] combined real walking and teleportation for covering short and long distances, respectively. The target destination is selected based on the user's viewing direction and teleportation is done by physically jumping forward. Other approaches for teleportation in VEs use hand gestures or key clicks [119]. Using key clicks, e.g. on the Vive controller, is less disorienting, requires less cogntitive stress and is more pleasant on average than fist pressing or jumping forward [119].

In contrast to the previously common teleportation in the line of sight, Bozgeyikli et al. [33] describes "Point & Teleport" as a type of teleportation where the target selection is decoupled from the viewing direction and the target is selected by controller pointing instead (see Figure 2.4.3). In a first study, this method was able to clearly outperform the isomorphic and conventional walk-in-place technique as well as the usual joystick locomotion techniques with regard to UX. In a second study, the method was extended with a functionality that allowed the users to determine their desired orientation even before the relocation. Frommel et al. [87] showed in a comparative study that free target selection causes less disorientation and higher presence than teleportation to fixed locations.

In summary, a combination of natural walking and Point & Teleport [33] is suitable for most VR systems, due to its practicability and ease of use, as there is no extra tracking device or controller, as well as no further customization of the VE or tracking space needed.

2.5 Summary

In this chapter, we have presented the theoretical background and related work for this thesis. In the studies and experiments presented in this thesis, the approaches and guidelines mentioned above were taken into account wherever possible. In addition, the literature on human factors, UIs, and interaction in VR research was examined and discussed in this thesis.

Moreover, HCI research on 2D UIs has also been considered when applicable to 3D UIs. We have conducted several user studies and have mainly used participants with little VR experience in our studies. However, we could not completely exclude prior knowledge about 3D games and movies they are exposed to in their everyday lives. Finally, in addition to objective measurements and subjective responses, we collected qualitative data to close the gap between restricted experiments and more open subjective feedback that could inform future applications using the VR UIs and VR interactions studied.

The work reviewed in this chapter shows that there is so far no common basis for the design and development of immersive VEs with regard to VR-specific evaluation metrics. Existing VR systems are often developed as design explorations by game designers with a playful or performance character, but rarely on the basis of a uniform evaluation framework for VR environments that considers user preference and task performance as well as external factors. Therefore, we discuss in more detail the evaluation of

UX in VR applications and present a new model with respect to external influencing factors and potential evaluation metrics (see Chapter 3). Moreover, we provide further related work about current development and evaluation methods and approaches for 3DUI and VR.

But to exploit the full capabilities of VR, we need a general understanding of the characteristics and properties influencing the outcome of fully immersive VEs. Furthermore, there is a need for the investigation and exploration of isomorphism in VR, especially for everyday VR experiences, as well as adapting approved concepts from interacting with VR UIs.

While the technology for input as well as output devices is market ready, only a few solutions for everyday VR exist, and empirical knowledge about performance and user preferences is lacking. Prior research and developers of substantial VR UIs (e.g. menu interfaces or text input) did not take the user's demands and needs into account. For example, VR shopping experiences would require the benefits of both worlds, i.e. efficient product search via speech or text entry and user-friendly UIs for exploration. Moreover, there is no standard method for VR text entry and current commercial systems implement their own techniques [96, 192, 194, 317]. Based on our VR UX evaluation framework (see Chapter 3) and inspired by related work, a set of various everyday VR scenarios are designed and implemented, such as shopping or system control (see Chapters 4 & 5).

Chapter 3

Virtual Reality User Experience (VRUX) Evaluation

> If real is what you can feel, smell, taste and see, then 'real' is simply electrical signals interpreted by your brain.
> – *The Matrix, Movie, 1999.*

A large percentage of the interactions taking place in immersive VEs fall into a small number of general categories: *navigating* from place to place, *selecting* virtual objects within the environment, and *manipulating* the position and/or orientation of virtual objects. Given techniques with good performance characteristics for these three types of interaction in environments, a large number of complex and effective VE applications could be developed. The question is whether isomorphic interaction concepts and representations of objects and environment is the right tool, or whether one should instead use the advantages and power of the virtuality, e.g. the multi-dimensionality of VR or the absence of physical laws. As an example, natural walking is a very natural and realistic method of locomotion compared to teleportation (see Section 2.4.3). Although natural walking offers a higher immersion due to the lower discrepancy between real and virtual, the non-isomorphic Point and Teleport [33] method allows navigation without physical boundaries through the tracking area and a lower physical effort because the user can stay in one place.

This, however, makes the design of UIs and interaction techniques for immersive VEs more complex, since they must first be evaluated in either an appropriate context or generally in comparison to conventional methods. In this chapter, we will first discuss in more detail related work on evaluation of 3DUIs and VR applications and then present a new approach and model for UX evaluation in the context of everyday VR with respect to external influencing factors and potential evaluation metrics.

3.1 Purposes of User Experience Evaluation

One purpose of this research is to develop principles for the design and evaluation of effective and usable VR environments, interaction techniques and UIs. In the remainder of this thesis, we use the term *artifacts* for environments, entire UIs or parts of them such as input or output devices or interaction techniques. LaViola et al. [161] highlighted three main purposes of usability evaluation. Here, the term usability refers to metrics, such as *objective performance* and *subjective responses* (see Section 3.3), to capture everything about an artifact or a user who influences the use of the artifact. First, usability evaluation is described as the assessment, analysis and testing of artifacts. Second, the identification of usability problems and issues leads to changes in the design of the tested artifacts. Third, usability evaluations can lead to design guidelines for future developers and designers to get a better general understanding of the identified usability issues.

However, as described in Section 2.3, there are different interpretations of user experience and usability that lead to different measurement and application areas. ISO definition 9241-210 and a survey at Nokia [19] suggest that UX measurements are similar to usability measurements, but with additional aspects of anticipation and hedonic responses. Thus, the evaluation of UX helps not only to improve human performance and the system, but also to understand how and why people use the artifact. Therefore, in the following we will not only talk about usability, but also about UX evaluation to refer to both approaches. Finally, our concept of Virtual Reality User Experience (VRUX) evaluation encompasses the objective performance of an artifact as well as the measurement and analysis of VR-relevant factors.

3.2 Evaluation Methods and Approaches

Besides the main purposes of usability evaluation, LaViola et al. [161] identified three key characteristics of usability evaluation methods: involvement of representative users (requirement of users or expert-only), evaluation context (generic or application-specific), and whether results produced are qualitative or quantitative. In general, there are several methods for evaluation, such as cognitive walkthrough [240], heuristic or guidelines-based expert evaluation [183], formative evaluation [118], summative or comparative evaluation [118, 265], questionnaires [118], and interviews and demos [118]. In our work, we chose questionnaires, interviews and demos for preliminary tests and focus on summative and comparative evaluation in combination for the main studies, as well as providing guidelines and heuristics for future developers and researchers. The approaches set out in the following sections are the current state of the art of usability evaluation for VR and form the basis for most user studies in VR [161]. Our approach for VR UX evaluation combines their benefits, following a user-centered process, addressing the challenges and risks of user-centered evaluation in VR, analyzing data using VR-specific metrics, and providing heuristics and guidelines for future developers and designers of VR experiences.

3.2.1 User-centered Design Approach

Nowadays, the basis for most software developments is the design and development life-cycle of the user-centered design (UCD) approach [1, 316]. Here, the focus is on gaining a better and more detailed understanding of the person using the artifact to be tested [224], i.e. who will be using the system, device or interaction technique. Although there is an international standard (ISO 13407) for UCD, there are many variations of the process, like the (most prominent) iterative approach [1, 88]. A UCD process involves users throughout the design and development and includes four main iterative phases within a design and development cycle (see Figure 3.1, as follows: specify **context of use**, specify system and user **requirements**, create **design solutions**, and **evaluate design** against requirements.

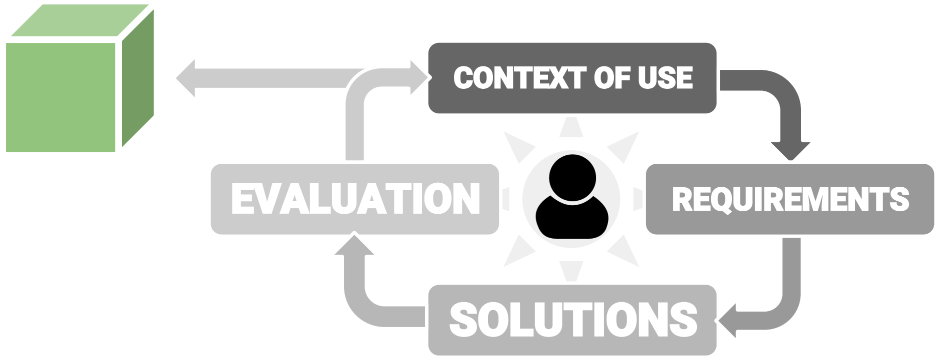

Figure 3.1: User-centered Design (UCD) approach, adapted from ISO 13407.

The differential advantages to the conventional linear approaches lie in the strategy and maintenance, the methodology and its measurable results. The UCD approach is centered on the user, not on development or the client, and is designed to improve the product instead of solving issues. On the other hand, UCD is not a perfect approach, because it can be time-consuming and costly. Designing UIs with user-centric intentions can often lead to non-user-centric results. The designers and developers should therefore take a closer look at the contextual impact and the way people actually use the UI.

3.2.2 Testbed Evaluation Approach

The testbed evaluation approach by Bowman et al. [28] aims at a generic-specific, more rigorous and more systematic approach to formally compare interaction techniques in VEs, regardless of the use case. The process of testbed evaluation comprises five different and consecutive phases (see Figure 3.2): initial evaluation, design phase (definition of taxonomy and design space, outside factors, and performance metrics), user evaluation, data analysis (quantitative performance results, heuristics and guidelines), and integration in the user-centered application. As already mentioned, the process starts with an initial evaluation to gain an intuitive understanding of the generic interaction tasks and techniques. Through this initial experience and gained familiarity with the technique, observations, and assessments of users, this phase provides good initial decisions for the other phases.

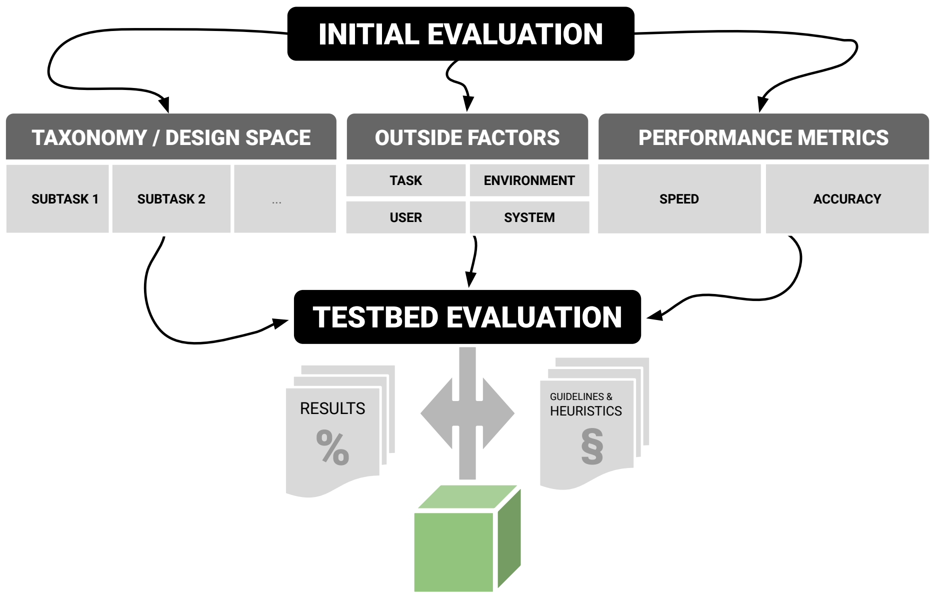

Figure 3.2: Testbed evaluation approach, adapted from Bowman et al. [28].

The results of testbed experiments allow the researchers to formulate more general guidelines and heuristics for the interaction in VEs in order to provide decision support for future developers of optimal VR UIs. This is followed by the design phase, which includes the definition and formulation of a taxonomy and design spaces of the task to be investigated, the metrics to be used and the potentially influencing outside factors. The taxonomy or the design space serves to divide the task into individual subtasks, such as changing the object color into selecting the object, then selecting and confirming the color. The individual subtasks can be assigned to different interactions, e.g. changing RGB and HSV values on sliders or touching in a 3D cylindrical color space. The results of a testbed evaluation depend on task, environment, user and system factors that should be included as secondary independent variables in the evaluation [28].

But the testbeds also have disadvantages compared to conventional evaluation. This systematic evaluation approach is usually more time-consuming, more expensive to implement and requires more experimental subjects. Testbed experiments generate complex data sets that are difficult to analyze and interpret. However, the benefits for future researchers and

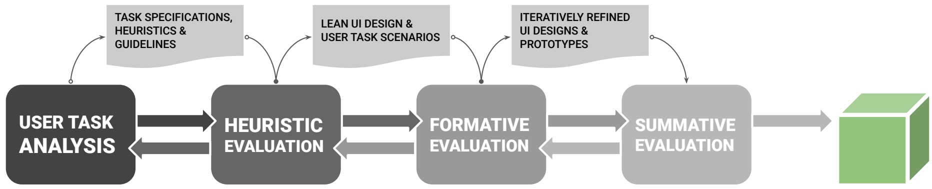

Figure 3.3: Sequential evaluation adapted from Gabbard et al. [88].

developers of building on the experimental results generated outweigh the disadvantages [28]. Since the results of a testbed evaluation must be reproducible in principle, reusability is one of the greatest advantages. Once an environment has been implemented, further artifacts can easily be tested and compared with the artifacts already tested. The growing knowledge base also offers the possibility of prediction within the defined taxonomy.

3.2.3 Sequential Evaluation Approach

Over the last decades, the application- and domain-specific sequential evaluation approach [28] has served as the standard approach for GUI usability evaluation. Gabbard et al. [88] adapt the basic idea and address the design and evaluation of 3D UIs by combining expert- and user-based techniques. As Figure 3.3 shows, this adaption includes four sequential evaluation steps: user task analysis, heuristic or guideline-based evaluation, and formative and summative evaluation. The sequential user task analysis resembles the taxonomy part of the testbed approach, but it provides a more application-specific task description including sequences, dependencies and limitations. The resulting preliminary guidelines and heuristics are evaluated in the following step, which provides a lean UI design and representative user task scenarios, all based on the task analysis. After that, the formative evaluation lead iteratively to refined and more detailed UI designs. Finally, the experimental results of the summative evaluation can be applied to similar applications and domains with similar devices and system. Even then, it should be possible for future designers and developers to make a general use of the derived guidelines and results.

Fully sequential evaluations might reduce study durations to the maximal extent, but require high-quality systems and equipment, and might lead to difficulties in quality management and a reduced clarity of the presentation of results. Furthermore, fully sequential designs should only be used if proof of efficacy is to be provided for a new method or a new breakthrough UI and generalizability and estimation of effect strength are of secondary importance. If, on the other hand, research into the theoretical and scientific background and the validation or classification of a UI or interaction technique are of major importance, generic-specific approaches with a fixed scope or moderately group-sequential approaches should be used.

3.2.4 Frameworks for Prototyping and Evaluating 3DUIs

In recent years, interest in 3D-related technologies such as 3D movies, AR and VR applications and games has increased. Although 3D interaction has a long research tradition, there is still a need for research on how different devices can be efficiently used as input for 3DUIs. Current 3DUIs, such as those provided by VR systems, consist of stereoscopic projection and tracked input devices. But these are often expert systems with complex UIs and require high instrumentation. Moreover, evaluation of VR artifacts is much more complex than 2DUI evaluation due to heterogeneous VR devices and techniques. On the other hand, immersive VEs enable the user to intuitively and naturally perceive 3D data. Interaction with stereoscopically displayed objects is still a challenging task even in VR-based environments [294].

Domingues et al. [76] provided a method for allowing developers and experimenters to quickly evaluate 3DUIs or parts of them (e.g. input devices or interaction techniques) during the design and the development lifecycle. Their approach is based on the V-Model for software development, and is used in the acceptance and system test phases of 3DUIs, as well as for debugging purposes in the integration testing and unit design phases. The MorphableUI framework [60] suggests sensor-based interactions with multiple mobile devices for stereoscopic 3DUIs. This framework allows the seamless addition of various input devices and can be used for rapid prototyping and evaluation of input procedures.

On the one hand, these frameworks are good for generic evaluation such as comparing input devices for navigation in general, but not for application-specific evaluation. Furthermore, they do not address the human factors of VR applications such as important outside factors. A more detailed description of evaluation metrics for VR can be found in the next section.

3.3 Evaluation Metrics for Virtual Reality

In the following, we present a novel approach for VR UX evaluation with respect to all VR-specific factors and metrics.

3.3.1 Evaluation Metrics for 3DUIs

As VR applications necessarily consist of *3D user interfaces (3DUI)*, Bowman et al. [29] differentiate between three evaluation metrics for 3DUI: system performance, task performance, and user preference. Whereas system performance metrics include benchmarking (like average *frame rate*, *latency* or *interaction time*), task performance metrics include quantitative measurements such as *task completion time*, *error rate* and *accuracy* (or precision). These metrics indicate to what extent users are able to cope with the task and interaction method. User preference metrics usually consist of subjective feedback, such as *UX*, *usability*, *motion sickness* and *workload*. Unlike for workload, where NASA TLX [107] is used as the standard questionnaire to measure task workload, there is no standard for UX, usability and motion sickness assessment in VR systems. Therefore, we discuss in the following paragraphs the current state of the art of measuring UX in VR applications and recommend suitable questionnaires, which were also used in our studies.

System Performance

System performance refers to objective measures like frame rate, latency or network delay, and tracking sensitivity. The tracking sensitivity of different technologies like optical tracking, gyroscopic or magnetic sensors may have more or less importance for the methods. Some methods require more accurate tracking (e.g. isomorphic and direct interaction), whereas others can cope with lower tracking precision (e.g. methods using d-pads or buttons for input, with tracking only for visualization of the controllers). As frame rate and delays are of unique concern, a method which does not meet the requirements of 'real-time' interaction would be out of scope anyway and is therefore not worthy of being evaluated or compared.

Task Performance

Task performance measurement is very application-specific and has its own standards for different contexts, e.g. words per minute (WPM) for text entry (see 4.1) or using a stopwatch for product search (see 5.2, 5.3 & 5.4). For a broader discussion of the task performance metrics for different applications in everyday VR, cf. Section 3.3.3.

User Experience and Usability

UX questionnaires are often used to determine the UX of artifacts or entire systems. Some questionnaires have been established which have different degrees of popularity and are often used in practice due to their popularity. Thomaschewski et al. [306] compared the dimensions and factors of the established UX questionnaires (AttrakDiff2, meCUE, SUS, SUMI, UEQ and VisAWI). Their results indicate that UX in VR should be measured by the User Experience Questionnaire (UEQ) [157], an end-user questionnaire consisting of 26 very short questions to measure UX quickly in a simple and immediate way. However, it is not always easy to decide whether subjective results from questionnaires can really show whether a product meets the desired quality aspect. Benchmarks developed especially for the UEQ allow comparison of the results with a large number of other user studies [261].

For the usability assessment, the System Usability Scale (SUS) [36] can be used; it is likely the most popular questionnaire for measuring attitudes toward system usability. It is a reliable and valid measure of perceived usability. Furthermore, it performs as well as or better than commercial and homegrown internal questionnaires.

Motion Sickness

When developing immersive VR applications using HMDs, motion sickness is one of the most crucial problems and a main factor influencing user preference and experience. The main symptoms of motion sickness are nausea, eye strain, headaches, and blurred vision [268], which are greater when using immersive VR. Although they are negligible in a desktop setting, they can be used as a reference for comparing different VR applications based on results in a desktop setting. To compare VR systems and methods, two different but comparative approaches for measuring motion sickness are available [65], namely objective and subjective measurement methods.

Objective Measurements The difficulty of measuring motion sickness objectively (e.g. based on measurements like heart rate or skin conductivity) is that there are no underlying physical symptoms that have been proven to be directly related to motion sickness. Cobb et al. [54] were able to show by measurements taken during and after wearing an HMD that motion sickness can be an individual problem in VR for ease of use. Further research on motion sickness symptoms and physiological changes used heart rate, blink rate, stomach upset and electroencephalography (EEG) for measurement [143]. Basically, objective measurement techniques represent a good individual result, but in fact they are not yet very mature and physically limit the user. Furthermore, no valid connections between measured objective values and the subjective perception of sickness could be found.

Subjective Measurements The Pensacola Motion Sickness Questionnaire (PMSQ) [139] is based on the Pensacola Diagnostic Index (PSI) [97] and is considered as one of the first motion sickness tests. The participants use scales to evaluate potential symptoms such as headache, dizziness, warmth, drowsiness, and nausea. Another widely used instrument for measuring motion sickness is the so-called Nausea Profile (NP) [213], which has been developed for medical use in order to capture complex experiences with patient nausea in more detail. As the use cases were increasingly tested in flight simulators, the PMSQ was modified and the 16-Point Simulator

Sickness Questionnaire (SSQ) [138] was developed, which is one of the most frequently used questionnaires for motion sickness measurement. Somewhat later, the Motion Sickness Assessment Questionnaire (MSAQ) [93] was developed; it enables the identification of multivariate measures. The results correlate both with the PSI and with the Nausea Profile [93]. Important here and a clear advantage over the other methods is that a further dimension of sickness was introduced in connection with the so-called "sopite syndrome" [165], which includes drowsiness and negative effects.

Finally, it can be noted that motion sickness influences both the wider acceptance of the technology itself and the continuous improvement of the technology. For military applications, the cost-benefit factor can far outweigh the ergonomic aspects. For more everyday entertainment applications, however, outside factors like health and safety issues related to motion sickness have the potential to significantly delay, or at worst prevent, the further development of immersive VR interfaces. Therefore, methods to reduce motion sickness in virtual environments are essential for survival and further entry into the consumer market.

3.3.2 Characteristics of Virtual Reality

Lee and Chung [169] formulated three important characteristics of VR: *immersion*, *interactivity*, and *presence*. Immersion describes to what extent the customer's senses are isolated from the real and stimulated by the virtual world. In particular, outside factors like field of view, frame rate, type of VR system and display (HMD or fish tank VR), as well as the degrees of freedom, determine the immersion of a VR experience. Presence reflects the subjective experience of being in one environment, but physically situated in another. Interactivity indicates to what extent users can participate in manipulating virtual content in real time. Furthermore, additional outside factors which could have an influence on the feeling of presence were identified by Mütterlein and Hess [214], namely content quality, initial excitement, isolation, and distraction.

The common questionnaire for VR applications from Slater et al. [278] measures the presence of a VE. Here, it is important to customize this questionnaire by changing the location description from 'office' to the appropriate environment of the test case. The Slater Usoh Steed (SUS) Count shows the mean of the SUS count of 6 or 7 scores amongst the 6 questions. The SUS Mean uses the mean score across the 6 questions instead.

3.3.3 Application-specific Metrics for Everyday VR

There are only a few categories of applications for VEs that are currently used for everyday VR such as gaming, text entry or shopping. The general requirements of these applications for interaction techniques and UIs cover a broad range, such as the already mentioned characteristics for VR, system and task performance as well as usability or motion sickness. In addition, there are many new applications of VEs that are being explored, which may require interactions with the environment with different properties. Therefore, we propose application-specific metrics that include mapping to a set of performance metrics, such as measurable characteristics of the performance (speed, accuracy) of a technique in a particular context. With this indirect mapping, application designers can define the desired levels of different metrics and then select a technique that best fits those requirements and the context. In the following, we present selected application-specific metrics for everyday VR applications that have also been used in experiments later in the work (see Section 5.2 and 4.1).

Virtual Reality Shopping Experiences

Although there has been a lot of research in multimedia and e-commerce [218], we believe that there are still open questions. In particular, designers of VR shopping environments are still faced with little guidance as there is no link between VR and e-commerce. So, we introduce the Virtual Reality Shopping Experience (VRSE) model in order to fill the gap between VR and e-commerce [282]. This model combines the metrics of customer satisfaction, evaluation of 3D UIs, and the characteristics of VR (see Figure 3.4).

Most users see the interface of online shops as appropriate, mainly because of their search functionality and opening hours [137]. This can lead to the assumption that ordinary online shops with user-friendly and attractive

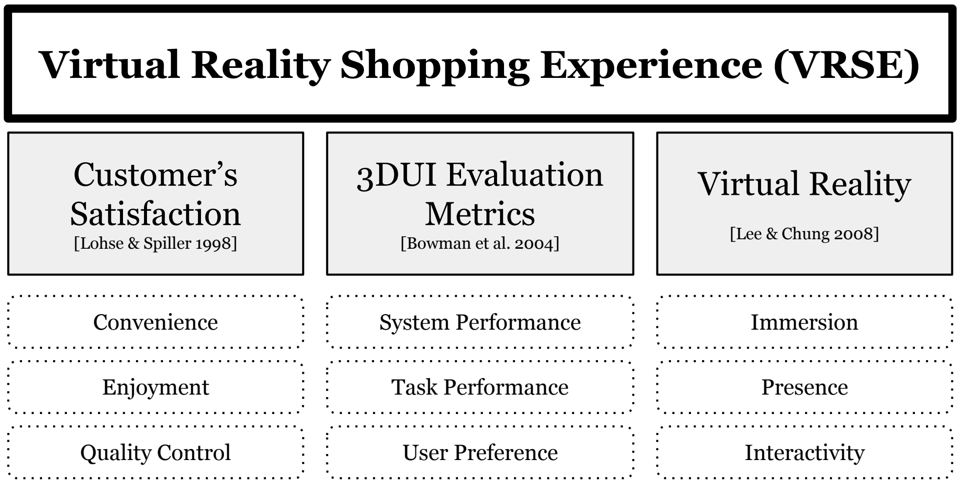

Figure 3.4: The Virtual Reality Shopping Experience (VRSE) model [282].

UIs would provide a higher customer satisfaction, including three main characteristics: *convenience*, *enjoyment* and *quality control* [128, 179]. Satisfied customers could tend to make not only more purchases at once, but also repeated purchases. Therefore, we want to address the main characteristics of customer satisfaction regarding online and offline shopping (see Section 5.2.2). With regard to online shops, convenience and comfort includes store navigation features like search functions, site maps or product indices which are essential for large stores [179]. Enjoyment is important, because people find playful interaction intrinsically interesting, e.g. when they are involved in activities like purchasing something for pleasure and enjoyment [5]. Quality control by customers of online shops is still a challenging task. It requires the ability to test a product before purchase or get an impression of its size and shape, which is not handled well by current online shops. While isolated online stores offer 360° images or videos of products, they cannot actually fill the gap in experience or immersion compared to shopping in brick-and-mortar stores.

In summary, the VRSE model combines the VR characteristics and 3DUI evaluation metrics into a model for evaluating VR shopping experiences with regard to customer satisfaction (see Figure 3.4). This will enable designers to build more usable and effective VR shops and help to move towards a stronger theoretical basis and more principled design guidelines.

Text Entry in Virtual Reality

Text entry in VR-based applications requires validation of findings and answers to new questions and should therefore rely on certain criteria. In particular, novel VR apps and their interaction techniques for text entry are mostly based on guidelines and standards from non-VR areas.

Text entry performance metrics indicate to what extent users are able to cope with the task and interaction method. Text entry evaluations usually focus on the same objective statistics, *speed* and *accuracy*. Nevertheless, repeated trials are necessary to generate great volumes of data, consisting of presented text (the stimulus, i.e. what they were asked to enter) and transcribed text (what they actually entered). When conducting text entry experiments where participants have to enter multiple sentences in a row, the words-per-minute (WPM) for measuring speed should be chosen, which is also the current standard for text entry evaluations [184]. Here, accuracy can be determined by the error rate, i.e. calculating the minimum string distance (MSD) between the presented and transcribed text and divided by the larger number of characters using Levenshtein's algorithm [171].

Gaming in Virtual Reality

In recent decades, video games have become an important form of entertainment, but also useful tools for education and training. Even in the health sector, video games are now used to increase motivation and change behavior, such as interactive fitness or exertion games [252]. Over time, different questionnaires have been developed, which explicitly take up the construct of engagement in the game and the experience of the player. Brockmyer et al. [34] presented the Game Engagement Questionnaire (GEnQ), which identifies the different levels of engagement of the player while playing. One purpose of the GEnQ is, for example, to identify children among the participants who might be threatened by violence in video games.

The Game Experience Questionnaire (GExQ) by IJsselsteijn et al. [122] acknowledges that it is not an easy task to adequately describe and measure the gaming experience. It captures the current experience during and after the game, as well as concerns around gaming with others. The strengths of the GExQ clearly lie in the number of items captured, such as positive affect, competence or challenge. Certain items, however, are difficult for the participant to grasp, since usually in laboratory experiments there is only a limited time available to play a game and to evaluate it accordingly. This questionnaire would therefore be more suitable for longer-term studies, in which participants of a study can adequately deal with the game.

But the biggest criticism of the GExQ is its status, labeled as "manuscript in preparation", which has not prevented prior research from using this questionnaire. However, a recently published validation study ($N = 633$) by Law et al. [164] could not prove the validity of the GExQ, so it is not recommended to use the GExQ; instead, more reliable and valid alternatives like other multi-dimensional measurements such as Player Experience of Need Satisfaction (PENS) [255] should be used. PENS is preferred for predicting fun and enjoyment, but also game ratings, sales, developer loyalty and sustained player interest. This questionnaire provides a detailed model to evaluate games across different genres in terms of in-game autonomy, player competence level, feeling of connection to other players, presence, and intuitive controls.

3.4 Virtual Reality User Experience (VRUX) Model

One goal of this thesis is to provide a model for measuring UX and the evaluation of interaction with VR. The aforementioned methods and metrics for 3DUIs and VR determine the applicability of different approaches and metrics for evaluating different artifacts in general. In the following, we therefore discuss in more detail our approach of UX evaluation in the context of everyday VR applications including a step-by-step evaluation guide. Furthermore, we present a new evaluation model for measuring UX in VR with respect to external factors and potential metrics. Finally, we present an example study procedure showing when and how to apply which measure and factor with respect to our VRUX model.

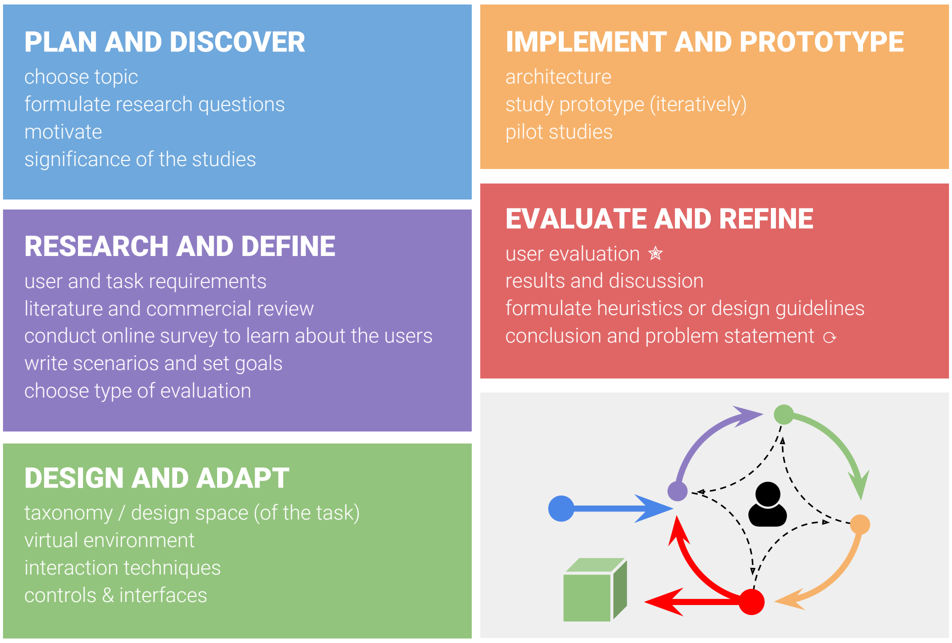

Figure 3.5: Step-by-step VR UX evaluation guide, based on the user-centered iterative model and sequential evaluation approach [88].

3.4.1 VR User Experience Evaluation Approach

Plan and Discover! Ask yourself important questions before starting the evaluation. Although you may be limited by certain guidelines for your location or work, e.g. due to financial or logistical reasons, the choice of topic is the first and most important step. Regardless, it is important to keep a few questions in mind: Are there enough related research results on this topic? Is the topic new and unique enough to give new opinions? Can the topic's complexity be reduced? Is it relevant to the domain or target group? Why is VR a possible modality; does it have to be an immersive 3D environment, or not? Often only the novelty effect or emerging hype is the main reason for choosing VR applications. Briefly state the high-level reason (or reasons) for conducting a VR study, and decide if it should be evaluated in a generic or application-specific context.

Research and Define! The goal of the second stage is to understand the context of use and to define the user and task requirements. In this phase, the prospective users of the product are identified, as well as why, how and for what purpose they will use it, and under what conditions. These specifications (or objectives) are of particular importance as external conditions can affect the performance and usability of an artifact. However, any requirement should be clearly defined and formulated to ensure a successful and usable conclusion of the evaluation. Here, *objectives* are questions to answer what we need to know in the design process and what knowledge gaps we need to fill. *Hypotheses* are assumptions whose validity we are unsure of, i.e. what we assume or think we understand about our users. The hypotheses about an artifact should be expected and reasonable, but not trivial or completely pulled out of thin air. Finally, based on the time, budget and participants available, the methods should be selected to provide as many details as future researchers need in order to make the study reproducible. In practice, the methodology should inform us about what will happen, for how long and where.

Design and Adapt! The third stage is to move from task to design and create design solutions, which may be done in further steps, building from a rough concept to a complete design. First of all, a *taxonomy* has to be formulated, including all subtasks and defining how to achieve or fail at the task and all subtask goals for every tested condition. A *design space* can be based on MacLean et al. [188], which is a standard approach to represent design rationale and uses a semi-formal notation, called QOC (Questions, Options, and Criteria). The questions, as the key parts of this design space approach, identify key design issues; options are specified to provide potential answers, and criteria are defined for assessing and comparing the options. Here, the study design will be created, including the type of study (between- or within-subject), power analysis and specification of the independent and dependent variables. This involves specifying the artifacts' properties (e.g. number of DOF or feedback modalities) and parameters (e.g. mobility, naturalness, or fidelity) to be tested using chosen metrics and measures (e.g. effectiveness, efficiency, or preferences). As a reminder, artifacts can be environments, UI or parts of it, such as input or output devices or interaction techniques. After that, the virtual environment where the study should take place, interaction techniques for navigation and manipulation, as well as the controls and interfaces for system control need to be designed and implemented.

Implement and Prototype! In the fourth stage, the artifacts to be tested and compared are integrated into a prototype. This stage is responsible for bringing all ideas to life with code. After specifying the architecture and experimental setup, a prototype for each artifact is implemented iteratively and undergoes several preliminary tests in the form of pilot studies or expert interviews. The choice of programming language, game engine or VR system (high-end or casual) depends on the domain, time available, and complexity of the task.

Evaluate and Refine! The final stage includes carrying out the experiment and evaluation, ideally through UX testing with actual users (see an example procedure in Figure 3.6). Finally, the experimental results can be applied to similar applications and domains and can be easily used by future VR designers and researchers. When designing for VR experiences, a different set of design considerations comes into play than when designing for 2D UIs. The resulting heuristics and guidelines can help upcoming VR designers and developers to create experiences which might survive the hype, and do not frustrate users or make them feel nauseous. After each evaluation, there is a conclusion phase which, in addition to the positive findings, also raises notable limitations, new questions and potential issues.

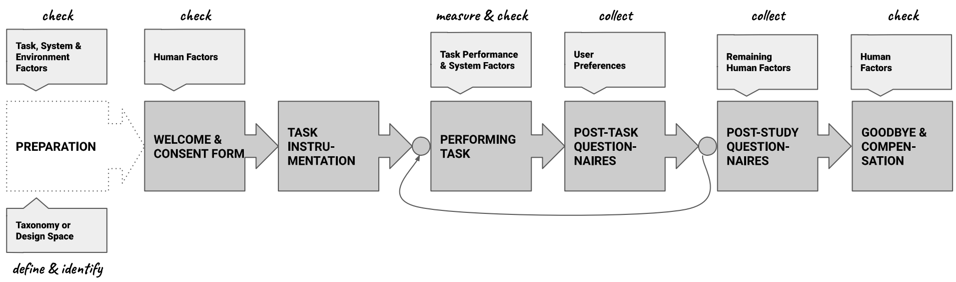

Figure 3.6: Details of an example study procedure in the user evaluation phase. The figure illustrates how and when to apply which measure and factor with respect to the VRUX model.

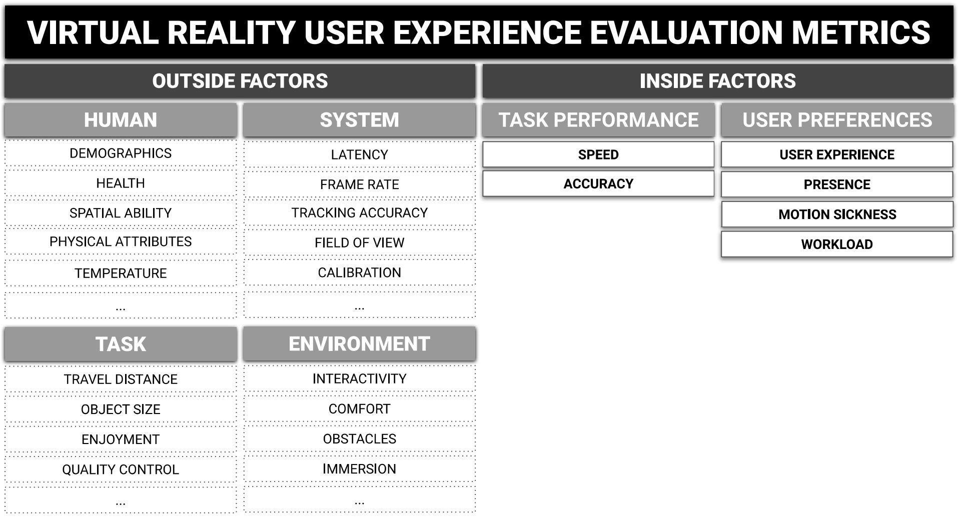

Figure 3.7: Overview of the global metrics of the Virtual Reality User Experience (VRUX) evaluation model.

3.4.2 VR User Experience Evaluation Metrics

We have already learned that there are many ways to interact with objects in VR, including various ways to select and manipulate objects, as well as moving in space or interacting with the system controls. So which certain VR-specific outside factors and influences restrict the use of the artifact under investigation? In the following, we discuss an example set of outside factors which should be clarified and determined prior to an evaluation. Then, we describe VR-relevant metrics and measurements which should be collected during the evaluation and thus provide conclusions about the effectiveness, efficiency and user preference of the evaluated artifact.

Outside Factors

The contextually relevant outside factors are divided into the following categories, adapted from related work [28, 161, 130]: *environment*, *task*, *human* (or *user*), and *system*.

Environment It should always be checked whether the artifact to be tested causes any kind of user discomfort, including motion sickness or fatigue, in the intended virtual as well as physical environment. A minimum of immersion, interactivity and freedom must be ensured in advance, such as adequate selection of the VR system and input devices to avoid potential negative influences. Especially if users can select objects with their hands and heads, and have to keep their heads or arms quiet and possibly in an unfavorable position for a long time, this can easily make them tired and can be very strenuous in the long run, cf. gorilla arm. Performance and pleasure may be promising at first, but later, when the users get tired, the performance can drop and high physical effort can be very frustrating.

Task The way in which the interaction or artifact is designed should really depend on the specific scenario and task. How many objects are there and how overloaded is the environment? The amount of objects and size of the VE could have a strong influence on the difficulty of the task, e.g. when maneuvering through a minefield or searching for a specific object. But also the size of the object or VE that users are supposed to interact with is of high importance, and how far away it is. For manipulating objects in arm reach other interaction techniques are better suited than for objects at larger distances (see Section 2.4.2). Another factor for VR interaction is expressiveness, which refers to the flexibility the user has with a particular technique. Researchers could show that bimanual interaction in VEs outperforms one-handed interaction even for complex tasks [153]. But what if the user can not freely manipulate an object with both hands? Moreover, some VR systems are equipped with only one controller or even none at all, which allows head pointing only. For example, in an interior design app, perhaps the system allows the user to move a piece of furniture to a different position and orientation [284]. But maybe it requires a complex 3D rotation [74], which could make the user feel frustrated and dissatisfied. Furthermore, customer satisfaction [169] including pleasure and quality control are also relevant task-specific factors.

Human The third category includes the human (or user) factors. As already described in Section 2.3.3, demographics like age and gender can have an influence on the application of VR and might cause motion sickness symptoms. Furthermore, there should be a balance between experienced and inexperienced VR users unless the study examines a particular experience group. Overall, it is absolutely necessary to ensure that there are no previous illnesses in the participants that could influence the experimental results: for example, they should not suffer from visual impairment or color blindness, especially in VR applications using a HMD. Participants who complain of occasional arm pain should also be excluded from mid-air 3D interaction studies. As well, spatial ability (e.g. spatial memory or estimating distance) and physical attributes (e.g. arm length or body height) of the participants have to be queried and, if necessary, normalized or excluded. Another human-related factor is the temperature, whereby a distinction is made between the room temperature felt by the participant and the real room temperature.

System One of the main problems of VR is motion sickness, which is caused by bad design of the environment, the UI itself, ergonomics, or parts of the VR system, which affect the UX of a VR application in different ways. Many VR systems use a tracking system that records the person's head movements and adjusts the images accordingly. Here the accuracy of the tracking is essential for usability; it can lead to frustration for the user, but can also have a direct influence on motion sickness. As already described in Section 2.3.3, the field of view has a direct influence on motion sickness. Besides the accuracy, an optimal calibration of the VR system is of special importance for the immersion of the system. An incorrectly calibrated body height or arm length can be unpleasant and quickly lead to the "uncanny valley" effect. Further critical system factors are latency, flicker or too-low frame rates, often caused by badly configured VR systems and insufficient graphics and computing power.

Inside Factors

In this work, we divide the relevant inside factors and recommended measurements for evaluating VR applications into two main categories, adapted from various related work: objective task performance [29] and subjective user preferences [29, 169].

Task Performance The first big category is task performance. How long does it take for users to actually complete the task? And how accurate are they in performing the task? Task performance can be measured in the form of objective data (speed, accuracy), e.g. task completion time for speed and error rate for accuracy.

User Preferences The second category includes usability-related issues and subjective responses acquired by post-task questionnaires. How easy and stimulating is it for the users to learn the interaction techniques, and how easy is it to use the UI? Is it likely to be cognitively or physically demanding? Or is it quite intuitive? Another inside factor, which is specific to VR, concerns how natural and immersive the experience is. Does the tested artifact make the whole experience more or less immersive? Here, the feeling of presence is particularly dependent on the level of immersion of the environment and the VR system. These are just a few examples of how outside factors can have a direct influence on the measurements and thus the evaluation of the artifact. In summary, data describing users' preference about the methods can be collected in the form of subjective feedback (UEQ [157] for UX, NASA TLX [107] for workload, MSAQ [93] for motion sickness, and SUS [276] for presence).

3.5 Summary

Different applications should have different emphases in relation to evaluation metrics. In applications such as psychotherapy and training, interaction may need to be simulated as realistically and immersively as possible so that patients or users can apply their VR skills or fears directly to the real scenarios they are being prepared for. Presence, accuracy (or precision) and naturalness are therefore the most important factors here.

In the gaming sector, if a small game is to be developed without targeting a specific group of players, the interaction does not necessarily have to be realistic, but it should be something intuitive so that everyone can easily use it without high instrumentation or expert knowledge. Therefore, the ease of learning and the ease of use would be more important than pure efficiency or practicality. If, on the other hand, a tool for experts is to be developed in a certain area, e.g. a data visualization tool for big data analysts, task performance is particularly important.

In this chapter, we presented several important results of using our developed methodology for UX evaluation in VR. Our new model is based on related approaches of 3DUI Evaluation [29] and general characteristics of VR [169], as well as classical methods of usability evaluation. In the next two chapters, we present example user evaluations in everyday VR user scenarios, like system control in VR (see Chapter 4) and VR concepts for shopping in VEs (see Chapter 5). Within this framework, we have developed new UIs that perform well in a variety of application scenarios and compared them with traditional concepts with respect to our VRUX model. For this purpose, we have developed general testbeds, as well as application-specific prototypes for VR UX evaluation, which can be reused for future comparisons. Finally, we present a large number of empirical results on the performance of isomorphic and non-isomorphic UIs and interaction techniques. These results lead to general principles and guidelines that can be applied to VR systems to improve performance and user preference. This should be useful and important for those developing VR systems with various levels of complexity of interaction.

Chapter 4

System Control in Virtual Reality

In recent years, VR and 3DUI are more accessible than ever these days and have seen a drastic increase in popularity, especially in terms of consumer-ready hardware and software. While the technology for input as well as output devices is market ready, only a few solutions for system control (e.g. text entry or menu control) and UIs exist, and empirical knowledge about performance and user preferences is lacking. In this chapter, we study text entry in VR by selecting characters on a virtual keyboard (e.g. for product search or entering a password or URL), as well as VR UIs for menu control (e.g. settings or main menu) with respect to our VRUX model (see Section 3.4). First, we discuss the design space for assessing selection-based text entry in VR and evaluate six implemented methods that span different parts of the design space. Our results show that pointing using tracked hand-held controllers outperforms all other tested methods. Then, we present and compare a simple planar UI similar to common 2D desktop interfaces with a pseudo-haptic UI based on physical metaphors. The pseudo-haptic UI performs better in terms of accuracy and user preferences.

The concepts and results of this chapter have been published previously in the following publications: [283, 284, 286, 285]

4.1 Selection-based Text Entry in Virtual Reality

As text-based communication is rarely studied in VR research regarding text entry performance [96, 192, 194, 317], there is a need for evaluating the user preferences (e.g. UX, workload, and motion sickness) and characteristics of VR (e.g. immersion) regarding recent technology and interaction concepts for text entry in VR. Furthermore, design guidelines for text entry are directly adapted from non-VR systems without any further investigation, like head or controller pointing on a QWERTY keyboard. Moreover, traditional text entry might not work in non-desktop 3DUI, because VR users are not fixed in general. Here, non-isomorphic techniques allow users to interact using "supernatural" metaphors to overcome limitations in the tracking space or anatomical constraints.

> *But what if tracked hand-held controllers aren't available, or low physical demand and motion sickness are of particular importance? Does natural (or isomorphic) interaction using fingers or pens for input have any impact on task performance or the user's preference?*

To answer these questions, we present a design space for selection-based text entry in VR, based on MacLean et al. [188]. We further contribute an analysis of this design space using a methodology that forms a basis for the development of VR text entry methods in future VR applications and enables researchers to relate future analysis to ours.

We presented them to participants ($N = 24$) in an empirical study to analyze their text entry performance and user preferences against our design space for selection-based text entry in VR. The results showed that the performance, workload and UX of our implemented pointing methods (head & controller) are above average compared to related work in VR [96, 335] or non-VR [192, 94]. Particularly in VR, UX and task workload turned out to be essential factors for text entry performance. In summary, the design space and the evaluated methods provide a solid baseline for comparison of future selection-based text entry methods in VR.

Figure 4.1: Left: Tap, Dwell or Gesture [335]. Right: FaceTouch [102].

4.1.1 Classification of Selection-based Text Entry in VR

Currently there is no standard method for VR text entry and current commercial systems implement their own techniques. Most of them are based on a selection paradigm, where characters are selected sequentially on a virtual keyboard floating in front of the user. This paradigm is familiar to users from other systems, suitable while standing or walking, and thus easy to adapt for VR. Since there is no taxonomy of text input using virtual keyboards in VR which we can refer to, we chose the term "selection-based" as used by prior work [192]. Nevertheless, performance and preference with such systems could greatly differ depending on the selection method.

Typing on a virtual keyboard providing live feedback has a crucial impact on users' typing performance while wearing an HMD [317, 176]. Here, physical keyboards would be more useful for text-heavy interactions but not for mobile use or while standing or walking, which is normally the case in common VR setups (e.g. HTC Vive). Here, the most common play area settings released by Steam[34] in 2017 indicate that 25% use "standing-only" ($1m^2$), 28% a room-scale setup with 4-$5m^2$, and the rest with larger play areas. Therefore, methods using VR controllers (e.g. pointing or cursor), or no controllers at all (e.g. head pointing [335], FaceTouch [102] or speech [32]), would be more suitable (see Figure 4.1.1). Apart from speech and head pointing, none could approach the performance of a physical keyboard (see Table 4.1), whereas several methods could be more useful regarding the user's preferences. We thus investigated how text entry using a virtual keyboard could be supported in VR with respect to our VRUX model.

[34] https://bit.ly/2HzlvpG

In Table 4.1 we categorize common text entry techniques based on their input method and compare different aspects of their design and achievable performance. Methods differ in the number of hands they require, if they make use of the QWERTY layout, if they require visual attention, if they provide haptic feedback, what input device they require and if it must be tracked in VR. The given performance is based on example prior studies conducted in or outside VR. We observe that performance studied outside VR is generally higher. However, these are estimates from prior work, which are not consistent in their methodology and are thus not directly comparable. Moreover, some methods have not been studied in VR at all.

Besides the lack of comparative performance evaluations in VR, as well as other contexts, little is known about users' preference of these methods, including VR specific factors such as immersion or motion sickness. Therefore, this work compares six common selection-based text entry methods using a virtual keyboard, covering a wide range of design aspects, by the same rigorous methodology for assessing performance and users' preference in VR based on literature review.

Input Method	Qwerty	Eyes free	Hands	Haptic feedb.	Device tracked	WPM in VR	WPM other
Soft button	✓	×	1-2	×	✓	4-7 [96]	33-36 [10]
Mid-air pointing	✓	×	1-2	×	✓	?	13-19 [192, 272]
Head pointing	✓	×	0-1	×	×	10-15 [335]	4.5 [94]
Eye gaze	✓	×	0-1	×	×	?	10-20 [190]
Discrete cursor	✓	×	1-2	✓	×	?	6-7 [330]
Physical keyboard	✓	(✓)	1-2	✓	×	24- 40 [176, 317]	40-60 [184]
Finger gestures	×	✓	1-2	×	✓	6 [96]	22 [291]
Chording/Twiddler	×	✓	1	✓	×	3 [96]	47 [182]
Multi-tap	×	✓	1	✓	×	12 [96]	20 [182]
Handwriting	×	×	1	✓	(✓)	?	15-20 [184]
Speech	×	✓	0	×	×	13 [32]	11 [120]

✓: yes ×: no ?: unknown/not applicable

Table 4.1: Overview of text entry methods evaluated in VR or potentially usable in VR. We compare the following factors: if the method uses QWERTY, if it can be used without visual attention, how many hands are needed for control, if it provides haptic feedback, and if tracked device is shown in VR. We then give performance estimates in WPM for evaluation in and outside VR, with example references.

4.1.2 Design Space

One purpose of this research is to develop principles for the design and evaluation of effective and usable selection-based text entry in immersive VEs. So, we introduce a design space based on MacLean et al. [188], which is a standard approach to represent design rationale and uses a semi-formal notation, called QOC (Questions, Options, and Criteria), for example to fill the gap between VR and selection-based text entry. The questions as the key parts of this design space approach identify key design issues and options providing potential answers to the questions, and criteria for assessing and comparing the options. As we wanted to build on prior work, we decided to use Markussen et al.'s [192] design space for selection-based text entry as a basis and adapted their QOCs for investigating selection-based text entry in VR using consumer hardware.

Questions and Options

There is only a few research in comparing text entry methods in VR [32, 96, 335], so we believe that there are still open questions, and especially designers of VR applications are still provided with little guidance as there is no common link between VR and text entry.

Q: Which keyboard layout? The QWERTY layout has been the standard keyboard layout of the last century [227], which supports the assumption that layout would have a positive impact on preference and performance, although it is not at all superior for expert performance [186, 182]. Even an imaginary "optimal" layout with faster keystrokes would only differ to a small extent [144]. Due to its habitual use in daily life, novel layouts (e.g. Dvorak [78] or OPTI [187]) would require large amount of adaptiveness and instrumentation until the first signs of improvement. However, Bowman et al. [32] suggest using the QWERTY layout for 3DUIs, if symbolic input will be infrequent or if most users will be novices. In summary, we would recommend using the QWERTY layout for VR text entry, if user preference is more important than performance.

Q: 2D or 3D? Bowman et al. [32] state that reducing the number of degrees of freedom (DOF) in 3DUIs is crucial when high-precision interaction is needed, in particular for text entry, a task that has been 2D for decades. So, 2D text entry methods could have a positive impact on intuitiveness, especially when using the QWERTY keyboard layout. But using 2D interfaces in 3D environments can decrease the awareness and feeling of presence (e.g. by taking off the HMD to use a physical keyboard or overlay it in the virtual scene [317, 176, 194]). But as selection-based text entry involves interaction with a virtual keyboard, there is also the question of how to represent the keyboard in the environment. While 2D approaches can mimic typing on a surface and enable users to imagine a keyboard floating in front of them, 3D would suggest more interactivity and increase UX and immersion.

Q: Typing in relation to what? Another question is whether typing should be in relation to an explicit reference point for gestures or selections. Touch-based surfaces (e.g. a Vive controller pad) can implicitly maintain a reference point for the user, whereas this is not the case with mid-air interactions. Mid-air techniques (e.g. keyboard input using fingers or stylus) make use of absolute reference points [215]. Here, the virtual representation of the keyboard is placed at a fixed location in the environment. By using a room-tracking VR system the user can walk around or even through the keyboard, which can increase the feeling of presence. As no empirical study on VR specific text entry method has covered this question yet, it is hard to determine the impact of this question on the criteria. So, in contrast to a fixed absolute position of the keyboard, it could also be positioned relative to the user, e.g. to the head for constant distance while typing or to the non-dominant hand while typing with the dominant [103]. While relative reference points are more flexible, they can cause high instrumentation and complexity of the UI itself.

Q: Position and size of the keyboard? The size of the keyboard representation matters especially for distance-dependent text input methods (e.g. direct input) but also for ergonomic reasons. Current text input methods for VR don't allow one to change position or size of the representation. Bachynskyi et al. [11] identified several clusters in input space for pointing gestures and advise to split the input space for right and left hand (if possible) and make the input representation fit the lower and peripheral input space. Nevertheless, larger input representations require a less precise tracking technique, but may afford more head movements, which consequently could result in higher motion sickness and workload. The immersion is driven by the interactivity, i.e. the absence of customization could reduce the presence when the user cannot manipulate the virtual world as expected [169].

Q: Feedback or not? Although typing on a virtual keyboard lacks real tactile feedback, it can be compensated for by using vibrator feedback from VR controllers, or pseudo-haptic feedback when pressing the virtual keys [167]. Here, the 3D key moves in depth when intersecting with the user's finger or controller, which simulates a physical button press in combination with auditive and visual feedback. Nevertheless, when using indirect keyboard interaction, users need at least visual feedback on tracking of their controller and cursor movements. Pointing-based methods using tracked hand-held controllers usually comes with the drawing of the corresponding ray intersecting the keyboard. However, at least cursor visualization is generally of high importance regardless of the input type, performance and user's preference, because otherwise it can be very confusing for the user. While visual feedback can support the perspicuity of novel methods, it may increase intuitiveness but also influence the feeling of presence. But less feedback can result in higher error rates and 'trial and error' learning.

Criteria

Our design criteria for selection-based text entry in VR will enable designers to build more usable and effective VR systems including text input and help to move towards a stronger theoretical basis and more principled design guidelines. Hence, we focus on the metrics of system and text entry performance (speed, accuracy), as well as the user preferences (UX, task workload) and characteristics of VR (motion sickness, immersion) with respect to our VRUX model (see Section 3.4).

Figure 4.2: This figure shows the VE including stimulus (purple), text input field, and the virtual keyboard.

4.1.3 Concept

Virtual Environment

The VE consisted of a virtual representation of a standard QWERTY keyboard in the participants' interaction zone at $1.3 - 1.7m$ in a comfortable distance for mid-air interaction, a text area for the output at eye sight, and the stimulus above (see Figure 4.2). The position and orientation of the keyboard could be adjusted for each participant. Apart from those three elements, the VE showing a sunset was empty, which made the environment more immersive but not distracting [32]. Visual feedback was similar for all methods. Here, hovering over a virtual key highlights the key in blue and symbolized the virtual cursors. For auditory feedback, the participants wore headphones and got audio feedback when selecting the virtual keys.

Evaluated Text Input Techniques

This section describes how each of the six candidates for VR text entry (see Figure 4.3) is used to select a character from the user's perspective, the important parts of the implementation, what parts of the design space this method covers, and finally the commonalities and differences.

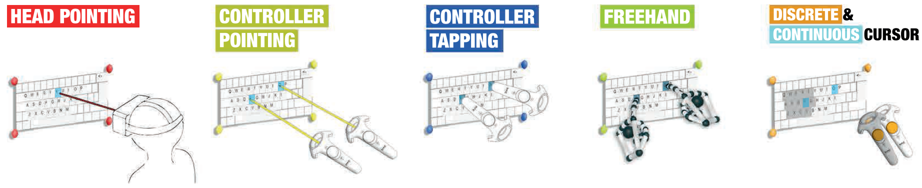

Figure 4.3: This figure illustrates our implemented selection-based text entry candidates for VR. From left to right: Head Pointing (HP, red), Controller Pointing (CP, yellow), Controller Tapping (CT, blue), Freehand (FH, green), Discrete (DC, orange) and Continuous Cursor (CC, light blue).

Head Pointing (HP) Pointing is one of the most fundamental patterns for selection [130] and is usually done with the head or hand-held controller. The basic idea is to follow an imaginary ray from the point of view (or object of interest) through the VE. In our approach of *Head Pointing (HP)*, we extend a ray from the main camera position towards the viewing direction provided by the SteamVR Unity plugin. If the first intersection is a character, it can be selected by a user-controlled trigger (button or dwell time).

Controller Pointing (CP) The implementation of our *Controller Pointing (CP)* method is analogous to HP but allows bi-manual pointing using both controllers. Moreover, it is designed for scenarios where controller tracking is available and extends a ray from each controller. It is actually not novel, because it was used for the Nintendo Wii via optical tracking using infrared. Here, the user controls the cursors' positions on the virtual keyboard by simply pointing at the character to select it, which is further visualized by changing the key's color.

Controller Tapping (CT) This method is implemented within the pen & tablet metaphor [8] and provides a more isomorphic, direct and realistic interaction than pointing. Text entry using a digital pen is less common but advancing fast with the rise of the Apple Pencil and several other tablets which can be operated quickly and accurately with a stylus. In our approach, the HTC Vive controllers are used for *Controller Tapping (CT)*, i.e. the virtual keys by reversing and holding them like digital pens. In contrast to pointing, this method requires physical manipulation.

Freehand Technique (FH) Our most isomorphic and realistic text entry candidate is the *Freehand (FH)* technique, where the user's fingers are tracked to type directly on a virtual keyboard. This technique doesn't require any tracked hand-held controllers, but instead the tracking of the fingers (e.g. using gloves or Leap Motion) [248]. In our approach, we decided to use the Leap Motion, which is mounted on the HMD. It is worth to mention that the Leap Motion provides no accurate tracking due to its technical limitations [323] and could therefore have a negative impact on the performance and preference. The potential gap in tracking precision compared to the other methods was therefore included in our discussion of the experimental results. We chose the Leap Motion as it is the current most affordable and available method for finger input and visualization.

Pad-based Discrete Cursor Control (DC) The majority of gaming consoles use text entry methods using the attached controller, more precisely the directional pad (d-pad) or thumb-sticks. In most instances, the text input is performed by controlling one discrete cursor over a virtual QWERTY keyboard for character selection, and conforming the input by a trigger button. We transferred this to the HTC Vive controller, which is also equipped with a d-pad and is controlled by the user's thumb. Furthermore, we improved the standard method for bi-manual input using both controllers simultaneously. Finally, in our implemented *Discrete Cursor (DC)* method, the user controls a cursor per controller for character selection, whereas each was intended for the corresponding half of the keyboard. In order to separate the two keyboard sections and to make it easier for the user, the left keys were colored darker then the right ones.

Pad-based Continuous Cursor Control (CC) The second method using the controllers' d-pad is the so-called *Continuous Cursor (CC)* method, which extends the functionality of DC. The difference between these two methods is that CC allows continuous control of the cursors. Here, we use the 2D thumb position on the touch pad and set the cursor of the corresponding keyboard half accordingly. Pressing a touch pad triggers the text input. Apart from that, the mode is analogous to DC.

Commonalities and Differences among Methods

In the following, we discuss what the six candidates for text entry in VR using a virtual keyboard have in common, as well as what divides them, according to our design space. First, we use the common QWERTY layout and fixed keyboard position and orientation for all methods according to prior work [227, 32], as our methods should aim at non-experts. The optimal position, orientation and size of the virtual keyboard representation is calibrated in advance for each user and method. In Table 4.2, we list which decision was made for each factor. The main limitation of pointing (HP, CP) is that the user cannot perform eyes-free text entry, which can cause higher mental and also physical demand (e.g. neck pain for HP, hand tremor and gorilla arm for CP). Despite the major advantage of direct mid-air input (FH), i.e. seeing the hands or controllers, there are several challenges including the lack of a touch sense: gorilla arm and the line-of-sight requirement [130]. In addition, FH is the only method where no hand-held controller is used at all, but instead a camera for tracking the user's fingers. Consequently, FH involves stable and solid tracking of the user's fingers, which is still a challenge in current research [323].

Besides that, CP and CT use the position and orientation of the tracked hand-held controllers, while DC and CC get along without tracking. Although techniques using non-tracked hand-held controllers (e.g. gamepad

	HP	CP	CT	FH	CC	DC
Which input device?						
Eyes free?	×	×	×	(✓)	✓	✓
Hands?	0-1*	1-2	1-2	1-2	1-2	1-2
Hand-held controller?	×*	✓	✓	×	✓	✓
Device tracked?	×*	✓	✓	✓**	×	×
Trigger button?	✓	✓	×	×	✓	✓
Feedback or not?						
Colored cursor?	✓	✓	×	×	✓	✓
Pseudo haptics?	×	×	✓	✓	×	×
2D or 3D?						
Character selection?	2D	2D	3D	3D	2D	2D
Key representation?	2D	2D	3D	3D	2D	2D

Table 4.2: Decision made regarding design space questions and options. (*) No additional controller is needed for HP except for a button to trigger the input, as head rotation is given for any VR hardware. (**) Tracking by an external camera, e.g. Leap Motion.

or joystick) aren't physically demanding in principle, there is still a high risk of suffering the so-called 'texting thumb' pain. Apart from CT and FH, all methods use a button to confirm text entry. Furthermore, we use visual and auditive feedback for all methods except for FH and CT. Here, we use a combination of visual, auditive and pseudo-haptic feedback when selecting a character in order to amplify the amount of realism. Finally, while FH and CT involve 3D manipulation of the controllers or hands, all other methods control 2D cursors on a planar virtual keyboard.

4.1.4 Empirical Study

We conducted a controlled laboratory experiment to compare the six text entry methods with respect to the task performance and user preference. Most VR applications require the user to infrequently enter short phrases. Thus, we were interested to compare the methods in a short text entry task, rather than a longitudinal study.

Participants

A total of 24 unpaid participants (5 female) volunteered in this experiment, aged between 22 and 29 years ($M = 25.29$, $SD = 1.89$). Two participants preferred the US QWERTY keyboard layout; the rest preferred the German QWERTZ layout. All participants rated themselves as able to read and copy English sentences ($M = 3.67$, $SD = 0.65$; on a scale from 1 (beginner) to 5 (native speaker)). All participants were right-handed. 28.6% had a visual impairment (glasses or contact lenses), but no participant was color blind. As the HTC Vive allows the user to wear glasses, no further adjustment was needed. On average, participants rated their experience in VR and that with each text entry method on a scale from 1 (novice) to 5 (expert) as follows:

- VR: $M = 2.57$, $SD = 1.50$
- Head Pointing: $M = 1.52$, $SD = 1.01$,
- Controller Pointing: $M = 2.71$, $SD = 1.52$
- Controller Tapping: $M = 2.57$, $SD = 1.34$
- Freehand: $M = 1.90$, $SD = 1.16$
- Gamepad/Joystick Cursor Control: $M = 3.81$, $SD = 1.14$

Apparatus

The VR system used an HTC Vive and ran on a Windows 10 machine with Unity 5.4. A standard desktop computer was used with an i7 CPU, 16 GB RAM and a Nvidia GeForce GTX 980Ti graphics card to fill out the questionnaires and control the experiment. Besides the Freehand technique, where the Leap Motion device was used to track the fingers, the two Vive controllers were used for the other methods, because they are tracked hand-held controllers equipped with d-pad and trigger buttons. The HTC Vive optical trackers (or lighthouses) were installed about $2.5m$ above the ground in two opposite corners to span a maximum Vive tracking area of approximately $4m \times 4m$. The participants were standing in its center while performing the tasks, as shown in Figure 4.4.

Design

The experiment was a within-subjects design, with one independent variable (Input Method) with six levels and six dependent variables related to

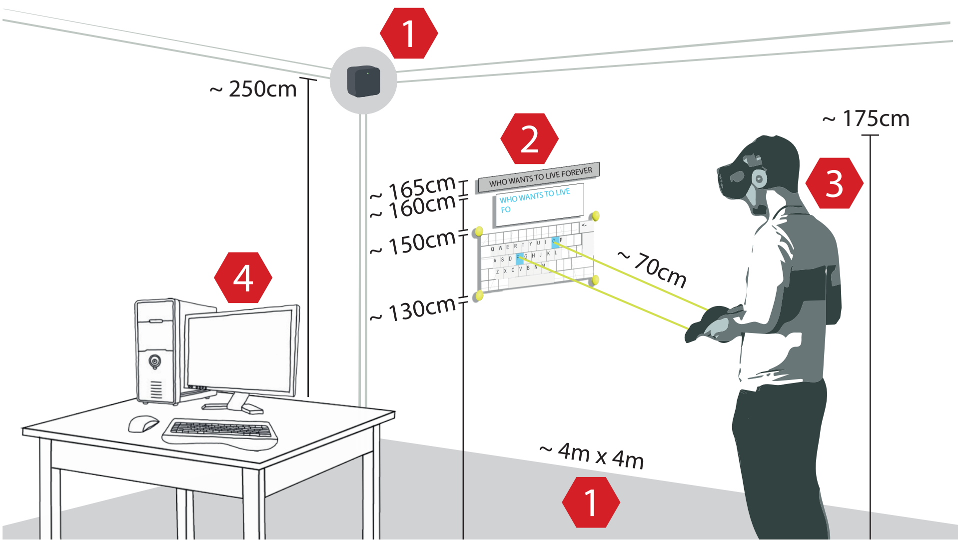

Figure 4.4: This figure illustrates the experimental setup. (1) Vive optical tracker (at $2.5m$) and tracking space with $4 \times 4m^2$. (2) Virtual keyboard, stimulus and text input field. (3) Participant wearing Vive and tracked hand-held controllers. (4) PC for experiment control.

the performance (speed, accuracy) and preference of users (UX, workload, motion sickness, immersion). The input method conditions were counterbalanced using a Latin square. Aside from training, this amounted to: 24 participants $\times$ 6 input methods $\times$ 5 phrases $= 720$ trials.

Task

The task was to transcribe five phrases (trials) with each text entry method as fast and as accurately as possible. Error correction was allowed using backspace. All phrases were randomly chosen from a set of 100 memorable phrases from the Enron corpus [312], with 20–28 characters each. We have chosen the Enron towards Mackenzie's phrase set, because of its higher validity for mobile text entry. Prior work has shown that both are comparable in terms of memorability, performance, and errors [151].

Procedure

The experiment started with a 5-minute SteamVR tutorial to get familiar with the headset and the controllers. It explains the bounding box of the tracking area how to use the controllers and their corresponding buttons. After a short break, the main part of the experiment started, where the participant was to perform all six tasks in Latin-square order, which lasts about 15-30 seconds on average for one of at least five trials per task. Before each condition, the interaction technique was explained and practiced in a warm-up phase of about five minutes. Participants received only minimal instructions about the functionality of the different interaction techniques, so that no explicit conceptual model was assigned to them. The participant was instructed to transcribe phrases as fast and as accurately as possible. Consequently, he or she was allowed, but not forced, to correct errors by using backspace. After each set of trials, the participant was asked to take off the HMD and fill out the post-task questionnaires to gather subjective feedback about the user's preference. After all trials were performed and post-task questionnaires were filled out, the participant was asked to fill out a final questionnaire collecting demographic data (age, gender, experience). Overall, the experiment took about 60 minutes per participant in total.

Evaluation Metrics

We measured task performance in the form of objective data (speed, accuracy) and collected data describing users' preference to the methods, including subjective feedback (UX, workload, motion sickness, immersion).

Task Performance For each participant, we measured text input speed and accuracy by calculating the average words per minute and error rate across the five entered phrases, in accordance with the common standards in text entry research [332], as follows:

- **Words per minute (WPM)** was computed by dividing the number of transcribed words (any 5-character string) by the time it takes to transcribe the text, formally:

$$WPM = \frac{|T| - 1}{S} \times 60 \times \frac{1}{5} \tag{4.1}$$

  where S is the time (in seconds) from the first to the last key press and $|T|$ is the number of characters in the transcribed text.
- **Error rate (%)** was computed by calculating the minimum string distance (MSD) between the presented and transcribed text and dividing it by the larger number of characters, formally:

$$\frac{100 * MSD(P, T)}{max(|P|, |T|)} \tag{4.2}$$

  where P and T denote the presented and transcribed text. MSD is calculated using Levenshtein's algorithm [171].

User Preference We collected a variety of subjective feedback to assess *UX* and *workload*, but also *immersion* and *motion sickness*, important in VR applications. Therefore, we used the following questionnaires:

- **User Experience Questionnaire (UEQ):** consists of 26 short questions to measure UX [157]. The scales cover classical usability (efficiency, perspicuity, dependability) and UX aspects (attractiveness, novelty, stimulation). The higher the score the better the experience.

- **NASA TLX:** assesses task workload based on 6 factors (mental, physical and temporal demand, effort, performance and frustration) [107]. The lower the rating the lower the workload.
- **Motion Sickness Assessment Questionnaire (MSAQ):** assesses the motion sickness based on 16 questions rated on a 9-point scale [93]. The scales cover four dimensions of motion sickness, which were defined as gastrointestinal, central, peripheral, and sopite-related. The lower the score the better.
- **Slater-Usoh-Steed Questionnaire (SUS):** a commonly used questionnaire to measure the user's immersion and presence in a VE. The SUS Count shows the amount of 6 or 7 scores in average amongst the 6 questions, while SUS Mean uses the mean score across the 6 questions instead. The higher the score, the higher the immersion and presence.

4.1.5 Results

Throughout this results section and in the following discussion we use abbreviations and color indications for the six text input methods we tested: **Head Pointing (HP, red), Controller Pointing (CP, yellow), Controller Tapping (CT, blue), Freehand (FH, green), Discrete (DC, orange) and Continuous Cursor (CC, light blue).**

Method	WPM	Error Rate (%) (corrected)	User Experience	Physical Demand	Frustration	Subjective Performance
Head Pointing (HP)	III: $10.20_{\pm 1.91}$	II: $1.15_{\pm 2.14}$	II: $0.67_{\pm 1.16}$	IV: $41.43_{\pm 0.00}$	II: $41.07_{\pm 0.00}$	II: $32.86_{\pm 0.00}$
Controller Pointing (CP)	I: 15.44 ± 2.68	I: $0.97_{\pm 1.19}$	I: $1.17_{\pm 0.78}$	III: $37.86_{\pm 27.82}$	I: $28.10_{\pm 24.42}$	I: $28.33_{\pm 20.88}$
Controller Tapping (CT)	II: $12.69_{\pm 2.27}$	III: $1.94_{\pm 2.22}$	IV: $0.56_{\pm 1.17}$	VI: $51.90_{\pm 26.05}$	III: $42.86_{\pm 32.12}$	III: $38.81_{\pm 21.67}$
Freehand (FH)	IV: $9.77_{\pm 4.78}$	VI: $7.57_{\pm 7.69}$	III: $0.55_{\pm 1.18}$	V: $46.43_{\pm 26.28}$	IV: $50.71_{\pm 27.85}$	IV: $40.00_{\pm 29.41}$
Discrete Cursor (DC)	VI: $5.31_{\pm 1.05}$	V: $2.79_{\pm 3.02}$	VI: $-0.40_{\pm 0.88}$	II: $30.71_{\pm 24.15}$	VI: $62.14_{\pm 23.64}$	VI: $54.05_{\pm 28.31}$
Continuous Cursor (CC)	V: $8.35_{\pm 1.58}$	IV: $2.15_{\pm 2.93}$	V: $-0.07_{\pm 0.92}$	I: $28.81_{\pm 21.62}$	V: $57.38_{\pm 28.44}$	V: $47.86_{\pm 24.22}$
$F_{(5,120)}$	36.28	7.00	5.95	3.26	3.76	3.02
p	< 0.01	< 0.01	< 0.01	< 0.02	< 0.01	< 0.04
η^2	0.60	0.23	0.21	0.11	0.14	0.10

Table 4.3: Objective measurements and subjective feedback ratings with significant differences between the text input methods. The best method per scale is visualized in dark green, the second in green. Furthermore, the ranking for each scale is represented by Roman numerals.

Task Performance

The task performance metrics include quantitative measurements such as speed and accuracy (or precision). These metrics indicate to what extent users are able to cope with the task and interaction method. They are computed per participant and input method as the average over the five trials. WPM ranged between 5.31 ($SD = 1.05$) for DC and 15.44 ($SD = 2.68$) for CP (see Figure 4.5). A univariate ANOVA showed significant differences between the methods regarding WPM ($F(5, 120) = 36.28$, $p < 0.01$, $\eta^2 = 0.60$). Bonferroni corrected pairwise comparisons showed significant differences between all methods, except FH-HP, FH-CC, and HP-CC. So, FH, CC and HP can be seen as one group with relation to speed. Overall, the average corrected error rate was low with CP as the best ($M = 0.01$, $SD = 0.01$) and FH the worst method ($M = 0.76$, $SD = 0.08$). There was a significant effect between the input methods ($F(5, 120) = 8.20$, $p < 0.01$, $\eta^2 = 0.26$).

User Preference

User preference metrics consist of subjective feedback concerning the UX, task workload, immersion and motion sickness.

Averaged over all input methods, the UX was rated at 0.41 ($SD = 1.13$) on average on a scale between -3 (very bad) to 3 (excellent). A univariate ANOVA showed significant differences between them ($F(5, 120) = 5.95$, $p < 0.01$, $\eta^2 = 0.21$). The pointing techniques were rated best in total (CP:

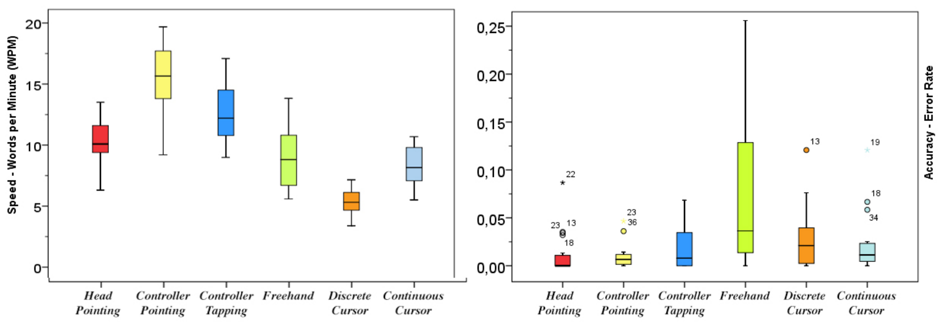

Figure 4.5: *Left:* Speed measurements, given in words per minute (WPM). *Right:* Corrected error rate measurements, given in percent (%).

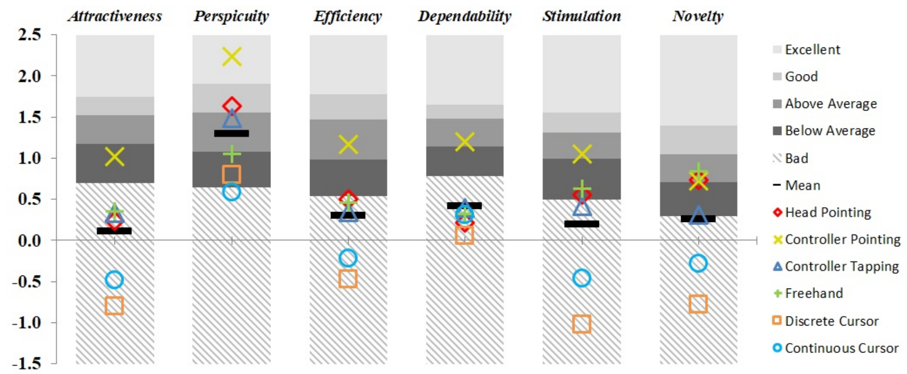

Figure 4.6: User Experience Questionnaire (UEQ) ratings with respect to comparison benchmarks.

$M = 1.17$, $SD = 0.78$; HP: $M = 0.67$, $SD = 1.16$). CP outperformed all other methods across the UEQ subscales, even ‚excellent' in terms of perspicuity ($M = 1.61, SD = 0.50$), except for the novelty aspect, where FH had a slightly better rating (see Figure 4.6).

The overall task workload was rated at an average of 47.10 ($SD = 20.94$). On average, CP was rated the best ($M = 38.43$, $SD = 22.32$) and DC ($M = 52.67$, $SD = 18.78$) the worst (see Table 4.3). We found no significant differences in overall task workload between the six input methods using univariate ANOVA ($p = 0.35$). But considering the NASA-TLX subscales, a multivariate ANOVA showed significant effects and differences between the methods regarding physical demand ($F(5, 120) = 3.26$, $p < 0.02$, $\eta^2 = 0.11$) and frustration ($F(5, 120) = 3.76$, $p < 0.01$, $\eta^2 = 0.14$), as well as performance ($F(5, 120) = 3.02$, $p < 0.04$, $\eta^2 = 0.10$).

The motion sickness total score was 19.8% on average ($SD = 1.32\%$). On average, HP was rated worst ($M = 22.29\%$, $SD = 14.40\%$) and FH best ($M = 17.99\%$, $SD = 8.91\%$). A multivariate ANOVA with all four MSAQ factors (G: gastrointestinal, C: central, P: peripheral, S: sopite-related) as dependent variables and single task as factor was conducted. It showed that there were no significant differences between the single tasks regarding the MSAQ factors (G: $p = 0.85$; C: $p = 0.47$; P: $p = 0.94$; S: $p = 0.81$).

The overall SUS count was 1.42 ($SD = 1.77$) and SUS mean was 4.33 ($SD = 1.19$) on average. However, the SUS counts for CT ($M = 1.62$,

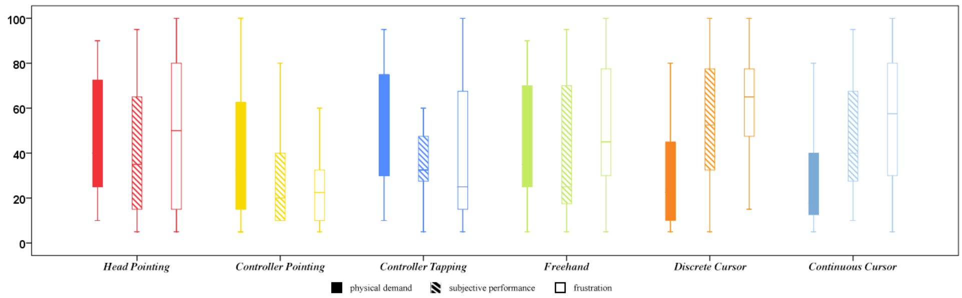

Figure 4.7: NASA subscales with significant differences between the six text input methods. Non-significant subscales (mental demand, temporal demand, and effort) are not shown for better clarity.

$SD = 2.04$) and FH ($M = 1.57$, $SD = 1.78$) are slightly higher on average than the rest. Nevertheless, the SUS mean is slightly higher for CP ($M = 4.50$ on average, $SD = 0.86$) than for CT ($M = 4.40$, $SD = 1.30$). There were no significant differences between the methods regarding immersion ($p = 1.0$).

4.1.6 Discussion

VR hardware (HMDs and controllers) are now widely available and affordable. But the question of new VR approaches to text entry has still not been widely explored. This research has thus attempted to explore this area by investigating task performance and user preferences in VR. In this section, we discuss the results, together with our observations and present a design decision tool that will be beneficial to future researchers and designers wanting to build on our findings.

Task Performance

Our results show that 15.4 WPM on average for CP can compete with the comparable QWERTY based pointing approaches from prior work [272, 192]. Although CP outperformed all other methods with regard to speed, tracked hand-held controllers are still required for this method. The controller-less alternative HP was with 10.2 WPM within the scope of the related VR head pointing approach from Yu et al. [335] and is comparable to the 6-13 WPM for speech input [120, 96]. Furthermore, HP was even faster than other head pointing techniques in non-VR studies [73, 17], which could be explained by the benefits VR involves.

However, character selection by pointing is usually not appropriate when realistic interaction is required, and controller or hand pointing can be imprecise due to natural hand tremor [149]. Concerning the isomorphic candidates, FH performance measures were relatively low, because the hardware we used couldn't deliver a satisfying experience. Keyboard input using optically tracked fingers implies crucial challenges (occlusion, accurate tracking sensitivity), so FH is strongly confounded by the accuracy of the Leap motion. Overall, we couldn't confirm the assumption that FH would have an effect on performance. Nevertheless, the technical and physical limitations of finger tracking techniques, especially for the Leap Motion device we used, have still a crucial impact on the task performance. In consideration of the fact that the accuracy attainable by the human hand is around 0.4 mm on average, the Leap Motion achieved 1.2 mm on average, whereas comparable consumer controllers (e.g. Kinect) were not able to achieve this accuracy [323]. But with 9.8 WPM on average, FH was faster compared to the related studies with 6 WPM [96].

User Preference

User preference is an aspect, which is not considered sufficiently within text entry in VR research. In the following discussion, we want to make the subjective feedback and observations more transparent.

User Experience Considerable methods due to their good usability and experience ratings are CP, HP and FH. But only the ratings for CP can be seen as above average to good, whereas HP and FH are rather below average. Because CT uses the same input device as CP and performed worse in all other measures (speed, accuracy, etc.) it can be disregarded if tracked hand-held controllers are available. Even though FH performed better in WPM than other studies found [32, 96], which could be explained by its naturalism and realism. However, the DC and CC methods were rated worse, so we would only consider them if the other methods are not possible at all. FH as the most natural way to type text on a keyboard has a positive impact on UX, especially because participants liked the novelty and stimulation of the method compared to the others. In summary, we can claim that using tracked hand-held controllers (e.g. CP or CT) result in better UX, while pad-based cursor techniques (CC, DC) should be completely disregarded concerning user preference.

Workload Due to Bowman et al. [31], natural (or isomorphic) interaction and especially mid-air writing provide little additional productivity, but could make the task more complicated and unnecessarily cumbersome. In addition, gestural interaction normally involves more muscles than other interaction techniques [15]. Considering the workload ratings of our mid-air and isomorphic techniques, FH and CT had a negative effect on physical ease. Here, participants needed to interact with the virtual keyboard in mid-air at an uncomfortable height, which consequently resulted in higher workload. So, if physical demand is a decision-making factor, mid-air techniques shouldn't be considered, due to potential gorilla arm fatigue.

But HP also resulted in high physical demand ratings and participants complained about neck pain and slight dizziness, which could have been slightly reduced by using eye tracking instead. CP was rated better, but still worse regarding physical demand even though the user needs to lift her arms at least to waist level. However, as the duration of the experiment, more precisely the single tasks, was not very long, the infamous 'gorilla arms' couldn't become a severe problem. Moreover, while pointing-based techniques do not enforce mid-air interaction, there is still a potential to suffer from hand tremor when using controllers at waist level, or neck pain if only head pointing is involved. Two participants complained about slight hand tremor after performing the CP task. However, due to the significantly better frustration and performance ratings, HP and CP should be always preferred to mid-air or pad-based text entry methods, if physical demand or realism can be neglected.

Nevertheless, tracked hand-held controllers and bare-hand input combined with appropriate feedback cues can help to make spatial relationships seem more concrete to the user and enhance presence by simulating physical interaction [114]. Although those problems can be bypassed when using

Vive controllers within their larger tracking space, there is still one potential drawback to it all: It is still unknown, whether current VR controllers can match the immersive quality of virtual hands and fingers visualization. Altogether, the experimental results would appear to give priority to all other methods over CC or DC. But disregarding performance and all other measures, physical demand ratings for pad-based methods were significantly lower than all others. When using non-tracked hand-held controllers (e.g. gamepad, joystick [333] or smartphone [7]), the user doesn't have to lift her arms or move her head. Only the danger of suffering a 'texting thumb' remains. However, this doesn't apply for pad-based methods using a d-pad on the HMD itself (e.g. Samsung Gear VR).

Motion Sickness & Immersion HP resulted in slightly (not significantly) higher motion sickness than all other methods. Of course a comparison regarding user preference of related non-VR techniques needs to be done in order to say more. However, there were no significant differences between the tested methods regarding immersion and motion sickness, which indicates that text entry in VR has no impact on immersion or motion sickness. Nevertheless, two participants stated that they would have preferred to have the keyboard on a virtual desk instead of typing on a floating virtual keyboard, which broke the immersion for them.

Design Guidelines for Text Entry in Virtual Reality

This section includes a set of general design guidelines for text entry in VR using a virtual keyboard, highlighting the major points discussed earlier in the design space, the experimental results and lessons we learned while moving from the task of text entry to the design and development.

When designing for experiences in VR a new set of design considerations comes into play than when designing for 2D interfaces. To help upcoming VR designers and developers of VR text input to create experiences that don't frustrate or make users feel nauseous, we created the following decision support tool based on Gonzales et al. [96] to guide the work (see Figure 4.8). Based on our findings, their tool for text input needs to be updated.

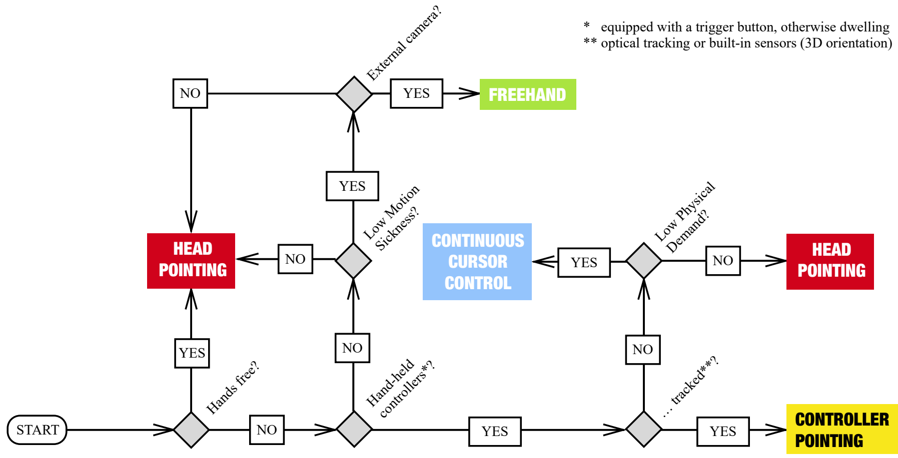

Figure 4.8: Decision support tool for VR text input on a virtual keyboard. Discrete Cursor is not considered, because of the bad results across all measurements. Controller Tapping performed slightly worse than Pointing, so it is not considered due to its higher technical requirements.

The decision which input device to choose will restrict the choice of techniques, if additional devices (e.g. gloves, cameras, keypads, etc.) or device tracking are needed. Nevertheless, hand-held controllers with at least one button to trigger input are recommended because of their robustness and familiarity. However, if the first choice 'User needs their hands free' is answered with yes, or no hand-held controller is available, Head Pointing using dwelling time should definitely be considered in addition to speech recognition. But speech has major drawbacks like recognition problems, privacy issues, and error-correction problems [333]. Regarding task performance metrics including speed and accuracy, as well as user preference, Controller Pointing performed significantly better than all other methods. However, if no tracked hand-held controller is available, the next method to choose would be HP, except when low physical demand is of particular importance. Only then is the Continuous Cursor the right choice.

4.1.7 Conclusion

In this paper, we have studied text entry in VR using a virtual keyboard and discussed the design space including criteria for assessing VR text entry methods. We have introduced six candidates that span different parts of the design space and evaluate their text entry performance and user preference. Although the general conclusion is to choose Controller Pointing for text entry in VR, its usage is dependent on certain criteria and limitations (e.g. tracked hand-held controllers). In addition, isomorphic keyboard interaction, as in the Freehand method, performed badly, even though it had promising UX results. To sum up, and putting our findings together with related work and our design space, in this paper we present an example decision support tool in the form of a flowchart, so that the results can be easily used by future VR designers and researchers.

Text entry is an essential part of HCI and there is still much research needed. Design annotation (e.g. for 3D artists or architects), filename entry or parameter setting, and communication between users are just a few applications for text entry in VR. Future VR systems (e.g. diaries, shops or social networks) may be designed to enable the user to stay in VR for longer times and therefore longer text entry needs to be feasible, too.

4.2 Pseudo-haptic Controls for Finger-based Menu Interaction in Virtual Reality

Nowadays, the major part of our population is familiar with GUIs (Graphical User Interfaces) on desktop computers or smartphone apps [217]. Over the last decades, menu metaphors have been widely established when designing 2DUIs and engraved into the user's mind: e.g. buttons, switches, or sliders. With the rising popularity of VR applications, planar UI elements have been directly adapted from 2DUIs and used for menu control in VR (e.g. Leap Motion Input Module[35]). Here, the common case is to render planar 2D controls in a VR 3D environment. But previous research has taught us not to assume that transferring conventional UIs into VR environments would be accepted by the users [161]. Prior work also discussed the benefits of using direct vs. indirect manipulations [328, 271]. Direct manipulations are hard to imagine without proper tactile feedback, in particular for controller-free and mid-air finger-based gestures (i.e. without haptic gloves or hand-held controllers). In these cases, the remaining feedback channels (auditive, visual) have to compensate for the lack of haptics. In general, VR enables UI designers to benefit from its multi-dimensionality, and in particular the stereoscopic view provides users with "pseudo-haptic" feedback for direct manipulations. As a consequence, research questions arise as to:

> *Why to use planar controls at all? Where do they have advantages and limitations concerning task performance and users' preferences compared to a pseudo-haptic UI? Which kind of button layout is the most efficient for menu control in VR environments among two common possible arrangements?*

In this paper, we present our study to answer which UI approach is best suited to our criteria based on the VRUX model, i.e. performance (speed, accuracy) and preference (UX, task workload, motion sickness, immersion). To this end, we first designed and set up two different UIs: a simple planar UI as baseline, and a pseudo-haptic UI based on physical metaphors.

[35] https://gallery.leapmotion.com/ui-input-module/

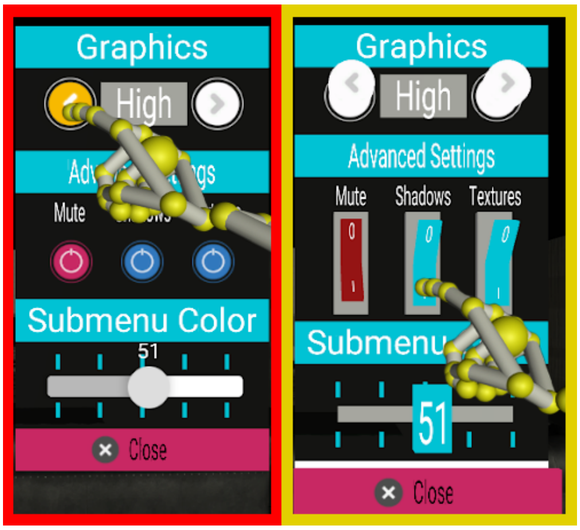

Figure 4.9: Appearance of the same menu controls (buttons, switches, sliders) in the two UIs used in the second experiment: planar (left, red), pseudo-haptic (right, yellow).

Both interfaces were designed for the same purpose: allow users of a VR environment to choose items from a menu using mid-air finger-based interaction. In particular, for the pseudo-haptic interface, we exploited the three-dimensionality of VR, in combination with visual and aural feedback, to provide the pseudo-haptic feedback. The planar UI is based on common VR UIs and is implemented using the standard controls from the Leap Motion SDK and its Interaction Engine. Finally, we used the two interfaces to carry out two subsequent within-subject experiments.

In the first experiment, the participants were asked to perform a selection task in a VR environment using two UIs (planar, pseudo-haptics) to access three button layouts (horizontal, vertical, circular). Here, we measured task performance (speed, accuracy) as the main part of the VRUX evaluation metrics (see Section 3.4. The study design was based on the menu selection study by Kulshreshth et al. [152], but using VR hardware for output and Leap Motion mounted on the HMD for input. The experimental results showed that the pseudo-haptic UI was more accurate on average, while the circular layout outperformed all other layouts. The same participants performed the second experiment after a short break, where we used the same apparatus and investigated what influence our two UIs have on the users' preferences (UX, task workload, motion sickness and immersion). Here, the pseudo-haptic UI was preferred in all aspects. Hence, the main contributions of this paper are as follows:

- **A study on mid-air finger-based VR menu control** with respect to *task performance* using different UIs (planar, pseudo-haptic) and layouts (horizontal, vertical, circular).
- **A study on mid-air finger-based VR menu control** with respect to *user preference* using different UIs (planar, pseudo-haptic).
- **Investigation of potential mechanisms and design guidelines** based on the results and observations during the studies.

4.2.1 Evaluated User Interfaces

Designing usable and effective UIs is very challenging for VR system developers and human factors specialists. For instance, a menu interface for controlling system states or changing object properties can be rendered as a GUI on a 2D desktop, but also as a flat object in a 3D environment. Our designed interfaces do not require any tracked hand-held controllers, but rather the tracking of the fingers, e.g. using gloves or Leap Motion [248]. In our approach, we decided to use the Leap Motion mounted on the front of the HMD. It is worth mentioning that the Leap Motion does not provide accurate tracking due to its technical limitations [323] and could therefore have a negative impact on performance and preference results. We have chosen the Leap Motion as it is currently the most affordable low-cost high-efficiency solution available for finger input and visualization in VR.

Planar User Interface

The concept of our planar and adapted 2D menu controls supports direct manipulation and serves as the baseline in our studies. It mimics a typical planar interface as is common for touchscreens or smartphones. The user interacts with the menus by touching the virtual elements with the virtual finger. As an interactive object, e.g. button or slider, is touched, a sound is

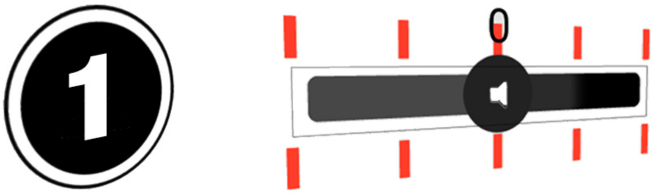

Figure 4.10: Planar UI Menu Control Elements (button, slider)

played as aural feedback and the color of the object changes. As there is no physical resistor, the virtual finger can point through all parts of the menu.

Planar Button In our planar UI, buttons are represented by simple images without depth (see Figure 4.10). We distinguish between three types of buttons: *Simple*, *Stepper*, and *Switch* buttons. With the *Simple* button, a simple touch triggers an action according to the usual functionality of a standard button. The *Stepper* acts like the simple button, but can be used to increment or decrement a value. Unlike the others, this button type supports a continuous-press function, i.e. the value is increased or decreased as long as the button is pressed. The *Switch* has two different states: *on* and *off*. Touching the *Switch* toggles its state, and its color is updated accordingly.

Planar Slider The concept of our planar *Slider* control includes two parts: a *handle* indicating the currently selected value, and *ticks* indicating the selectable values (see Figure 4.10). The handle can be moved and dragged while touching it with the finger. Here we distinguish between *discrete* and *continuous* control. In addition to visual feedback, a sound is played whenever the handle has been touched and every time the value changes.

Pseudo-haptic User Interface

The pseudo-haptic UI was created with the aim of maximizing affordance. All controls stick out of the menu in order to suggest and underline their three-dimensionality. We assumed that pseudo-haptic widgets based on physical metaphors, such as knobs or switches, could help to address the common issue of lacking feedback. According to their physical metaphors and affordances, their shapes should suggest the required actions in order to interact with them, e.g. pushing a protruding button or pulling a lever. Physical haptic cues suggest that they are solid, interactive and touchable.

Pseudo-haptic Button In our pseudo-haptic UI, buttons consist of a fixed in space trigger circle and a protruding disc which is movable only orthogonally in depth. Pushing the disc through the circle triggers a button press, which is also confirmed to the user through visual and auditory feedback, i.e. color and sound. Here, a continuous press is done by holding the disc behind the trigger circle, which can be used instead of multiple and repetitive single button presses. While common virtual buttons let the finger or controller pass through when pressing them, those pseudo-haptic buttons are pushed along the pre-defined axis (see (a) in Figure 4.11). According to its underlying physical metaphor, a button moves back to its original position after the finger has been released, like hitting piano keys.

Pseudo-haptic Switch In contrast to the planar UI, in which the button control includes switch behavior, our concept of pseudo-haptic switches offers three types of switches: *Toggle*, *Rocker*, and *Slider* switches (see (d,e,f) in Figure 4.11). *Toggle* switches are comparable to the switches which are used in cockpits of airplanes. By flipping the bi-stable switch from one side over its center, it automatically moves to the end of the opposite side, changes its color and state, plays a short "click" sound and triggers an action. The *Rocker* switch is similar to the *Toggle*, but its visualization and feedback is based on the behavior of conventional light switches. *Slider* switches are similar to our 3D sliders and are based on smartphone slider controls.

Pseudo-haptic Slider & Wheel Similar to the sliders in the planar UI, our *pseudo-haptic Slider* consists of a handle and tick marks, which appear along the top and the bottom of the control (see (b) in Figure 4.11). Here, the user can manipulate the slider value, which is displayed on the handle, by simply pushing the handle to the left or right like a real physical object. In addition to the pseudo-haptic *Slider*, our pseudo-haptic UI also includes a pseudo-

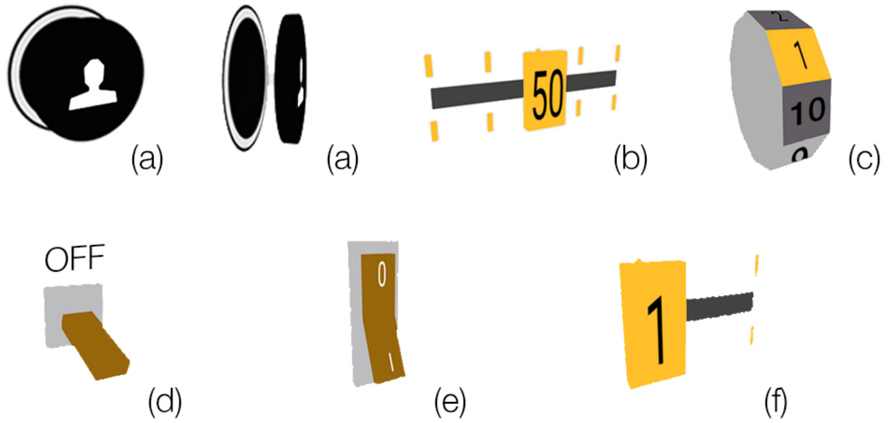

Figure 4.11: Pseudo-haptic UI Menu Control Elements (a: button, b: slider, c: wheel, d: toggle, e: rocker, f: slider switch).

haptic *Wheel*, which is based on common picker controls in mobile UIs and is visualized by a decagon-shaped wheel, which can be used horizontally or vertically (see (c) in Figure 4.11). To turn the *Wheel* using a finger, the applied force when touching a wheel's face rotates the wheel, which makes it easier to scroll through larger datasets, e.g. when choosing year of birth.

Advantages and Limitations

The main advantage of using customized planar UIs in VR is simplicity, acceptance and familiarity, due to the widespread use of smartphones and desktop computers. But the major limitations could be unintended inputs resulting from the small "gulf of execution" [223], even worse if too many UI elements are too small and too close to each other. Overall, we expected that the pseudo-haptic UI would be perceived as very engaging, interactive and attractive to use, but less efficient. Furthermore, it utilizes pseudo-haptic feedback through a combination of physical metaphors and feedback substitutions. Ultimately, we see the major benefit of our pseudo-haptic UI concept with regard to the user's preference due to the isomorphic aspects including familiarity, predictability and intuitiveness given by the physical metaphors [161]. But the pseudo-haptic UI might lag behind in performance, due to the additional effort for pushing a button or turning a wheel.

4.2.2 Experimental Setup

We conducted both experiments in controlled laboratory conditions. All participants performed both experiments one after the other, with a short break (~ 5 min) in between, within one session (~ 60 min). The order of the experiments was counterbalanced. In addition, all participants were part of the same experimental setup for both experiments.

Apparatus

The VR system used an Oculus Rift CV1 as a head-mounted display (HMD) and ran on a Windows 10 machine with Unity 5.5.4, which was also used to fill out the questionnaires and control the experiment. For finger tracking, we used the Leap Motion device mounted on the front of the HMD. Here, we used the Leap Motion SDK for Unity (Core Assets 4.1.6) for implementing all interface elements, as well as the Leap Motion Interaction Engine Early Access Beta[36]. The participants were standing in the center of the tracking area while performing the tasks for both experiments. For auditory feedback, the participants got audio feedback from the HMD headphones when interacting with the control elements.

Participants

A total of 31 unpaid participants (11 female) volunteered for this experiment, aged between 18 and 47 years ($M = 24.90$, $SD = 6.34$). After the experiments, they were asked to enjoy a VR experience in return for their participation. All participants were right-handed. 41.9% had a visual impairment (glasses or contact lenses), but none were color blind. As the Oculus Rift CV1 allows the user to wear glasses, no further adjustment was needed. On average, 61.3% of the participants had prior VR experience.

[36] https://goo.gl/MT96bk

4.2.3 Experiment 1: Task Performance

Although both experiments had the same experimental setup (participants and apparatus), they differed in their hypothesis, study designs, VEs, tasks and procedures. In the first experiment, two mid-air finger-based UIs (planar, pseudo-haptic) and three layouts (circular, horizontal and vertical) for menu control arrangement were compared with respect to task performance (speed and accuracy). Most VR applications require the user to interact with a menu interface including mainly buttons. Thus, we explored the conditions in a short selection task based on the study by Kulshreshth and LaViola [152]. Our main hypotheses were defined as:

$H_{1.1}$ The *Planar* UI is faster and more accurate than the Pseudo-haptic UI.

$H_{1.2}$ The *Circular* layout is faster and more accurate than the other layouts.

Design

The experiment was a within-subjects design, with two independent variables (user interface, layout) with two levels for the user interfaces (planer, pseudo-haptics) and three levels for the layouts (circular, horizontal, vertical), and two dependent variables related to task performance (speed, accuracy) based on the study design of Kulshreshth and LaViola [152] to ensure comparability and validity. All 10 trials per condition were counterbalanced using Latin square order. Aside from training, this amounted to: 31 participants $\times$ 2 user interfaces $\times$ 3 layouts $\times$ 10 trials $=$ 1860 trials.

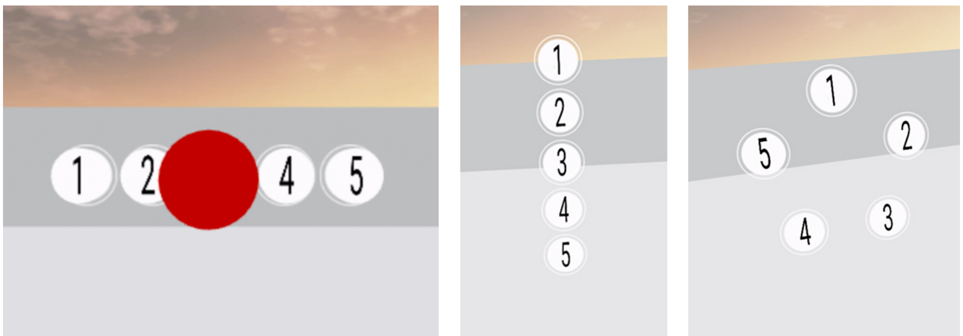

Figure 4.12: The three tested button layouts (from left to right: horizontal, vertical, circular) within the VE of the first experiment.

Virtual Environment

To avoid unnecessary or random distractions and to help the participant to focus on the selection task in order to be as fast and accurate as possible, we used a minimalistic environment, intended to be perceived as neutral and not distracting but still immersive [32]. The scene itself consisted of a gray floor, a single gray wall in the back and a standard sunset skybox to avoid visual distractions by a high contrast (see Figure 4.12).

A red sphere was placed at a comfortable distance ($\sim 20\%$ of the arm's reach) between the white numbered buttons and the user. This sphere served as a "call-to-action" button, i.e. before each trial the participant had to place a fist inside the sphere for $1s$ to confirm and start the trial, and consequently the time measurement. This also ensured that every trial started at the same position. The confirmation phase was interrupted immediately if the user extended at least one finger from the sphere. For the buttons, we used black numbers on a white background, and blue for the targets respectively, to make them as neutral as possible and distinguishable.

Task

The task goal for each of the two tested interfaces was to select, ten times consecutively, a blue target button among five possible buttons for each of the three button layouts, as quickly and accurately as possible. For every trial, one of the five circular numbered buttons randomly turned blue and became a target. A trial was started if and only if the participant confirmed it by holding a fist, i.e. no fingers extended, for one second within the call-to-action sphere. A button turned yellow as visual feedback and a sound was played when a button was pressed.

A trial, and the respective time measurement, ends if and only if the blue target button has been pressed successfully. Finally, it is worth mentioning that the experiment was designed for one hand only, more precisely the participant's dominant hand. They were not allowed to switch hands during the experiment, as otherwise this could influence the consistency of the arm fatigue ratings in the post-task questionnaires.

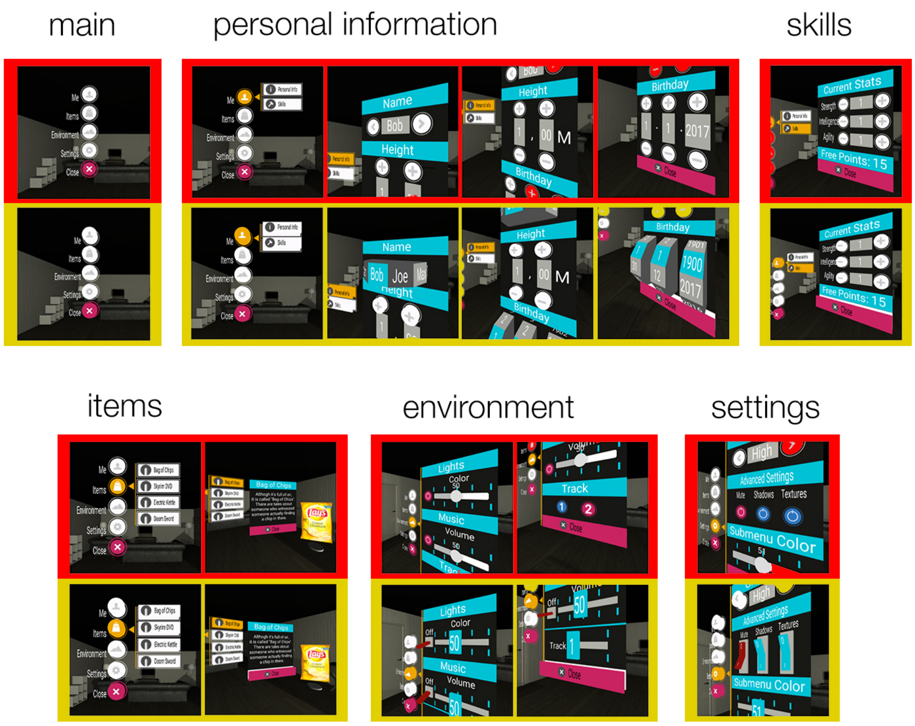

Figure 4.13: Planar (red) and pseudo-haptic (yellow) controls in the menu, as well as the five subtasks for each submenu used in the second experiment.

Procedure

At the beginning of the experiment, the participant was welcomed by the experimenter and was asked to fill out an informed consent form. Before each condition, the user interface and its interaction technique was explained and practiced in a training phase of five minutes maximum. During the training, the system was calibrated by the experimenter, i.e. adjusting the distance (∼80% of the arm's reach) and the height of the targets (center of the layout at the shoulder's height) to be more comfortable for the participant. Participants received only minimal instructions about the functionality of the different interaction techniques, so that no explicit conceptual model was assigned to them. The participant was instructed to confirm the trial start and to select the target button as quickly and accurately as possible. After each set of trials for one user interface, the participant was asked to take off the HMD and take a short 2-minute break.

4.2.4 Experiment 2: User's Preference

The two experiments were conducted in sequence with a break of five minutes in between. Here, our main hypotheses were defined as:

$H_{2.1}$ *Pseudo-haptic* allows for better UX than *Planar*.
$H_{2.2}$ *Pseudo-haptic* requires a lower task workload than *Planar*.
$H_{2.3}$ *Pseudo-haptic* allows for lower motion sickness than *Planar*.
$H_{2.4}$ *Pseudo-haptic* allows for higher immersion than *Planar*.

Design

The second experiment was a within-subjects design, with one independent variable (UI) with two levels (Planar, Pseudo-haptic), and four dependent variables related to the user's preference (UX, Task Workload, Motion Sickness, Immersion). All 26 trials per condition were counterbalanced using Latin square order. Aside from training, this amounted to: 31 participants $\times$ 2 UIs $\times$ 26 trials $=$ 1612 trials.

Task

Before letting the participants face the tasks, they get some time to become familiar with the interface again, as well as to explore the menu. In contrast to the first experiment, the users are allowed to use both hands. All participants performed all 26 tasks per interface; there were grouped into five subtasks for each submenu (Personal Information, Skills, Items, Environment, and Settings) plus the main menu (see an example set of controls in Figure 4.13). To ensure comparability in between, and to avoid undesirable effects (e.g. learning or fatigue), the trials per subtask, as well as the subtasks, were shuffled for each participant. In *Personal Information*, the name, height and date of birth can be set to predefined values. In *Skills*, the participant adds or removes points for attributes like strength, intelligence or agility. The *Items* trials are more like a search task, in which a certain object (e.g. a sword) can be rotated by touching it, and is equipped with a hidden code. In *Environment*, the lighting and music can be toggled and the color and volume can be changed by moving a slider. *Settings* consists of switches for toggling shadows and textures in the scene. Finally, it is worth mentioning that the first trial of every subtask includes pressing the correct menu button in the main menu. A subtask ends only if the participant has returned to the main menu by pressing a main menu or return button.

Virtual Environment

In the second experiment, we evaluated the users' preferences. Here, we wanted to provide the users a more realistic scenario of menu control in a VR application in contrast to the performance test, but still keep it as simple as possible. Therefore, we chose a virtual apartment as the environment, where the user manipulates the system states in realtime (e.g. rendering or shadow quality), as well as object properties or environment settings like in a realistic smart-home scenario (e.g. music volume, light color or intensity), by mid-air finger-based interaction with a VR menu interface. The apartment scene consists of several furnitures (kitchen and living room) and electronics like lamps, a TV and a stereo system (see Figure 4.14).

Procedure

As mentioned before, the participants take off the HMD after the first experiment has ended and take a break of five minutes before the second experiment starts. In a training phase before each tested interface, the experimenter gives a brief overview about the handling of all included widgets and menu controls. This training phase is guided by the experimenter to

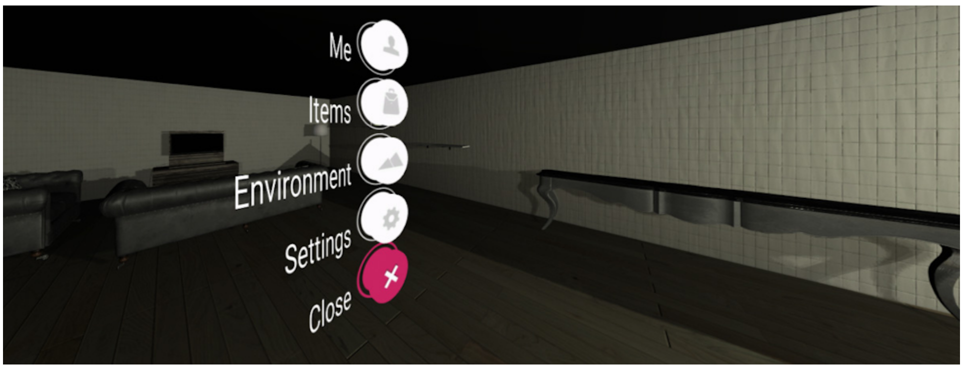

Figure 4.14: The pseudo-haptic UI in the VE used in the second experiment.

ensure that the participant has seen and tried out every control element in order to ensure she has understood the important aspects and differences between the interfaces and controls. Then the main phase of the experiment starts, i.e. 26 different trials per interface, which were shuffled in random order within the five subtasks. Every single trial proceeds as follows: (1) the experimenter reads the trial goal aloud, e.g. "Turn the lights on" or "Return to the main menu", (2) the participant confirms that she understood the instructions, (3) the participant starts carrying out the task, and (4) the participant notifies the experimenter that the trial has ended. After all trials of one interface are finished, the participant is asked to take off the HMD and to fill out the post-task questionnaires (UEQ [157], NASA-TLX [107], MSAQ [93], Slater-Usoh-Steed (SUS) [278]). Finally, after all tasks have been performed, the participant fills out a final post-study questionnaire to gather demographic data (age, gender, experiences).

4.2.5 Results

We evaluated two UIs (planar, pseudo-haptic) and three layouts (horizontal, vertical, and circular) for menu control in VR using mid-air finger-based interaction. Throughout this results section and in the following discussion we use abbreviations and color indications for the conditions we tested:

- *Planar UI (planar, red)* — the participant selects the planar target button by touching it with her finger.
- *Pseudo-haptic UI (pseudo-haptic, yellow)* — the participant selects the pseudo-haptic target button by pressing it with her finger behind the trigger circle.

Experiment 1: Task Performance

The task performance metrics include quantitative measurements such as speed and accuracy during the first experiment, which indicate to what extent users are able to cope with the task and interaction method.

Speed is given in seconds and ranged between 0.81 ($SD = 0.78$) for planar and 0.84 ($SD = 0.36$) for pseudo-haptic (see Table 4.4) on average, as well as 0.76 ($SD = 0.89$) for circular, 0.85 ($SD = 0.40$) for horizontal,

Experiment 1: PERFORMANCE	Speed (s) $[0,\infty[$	Accuracy $[0,\infty[$	**Experiment 2:** PREFERENCE	User Experience $[-3,3]$	Task Workload $[0,100]$	Motion Sickness $[0,1]$	Immersion $[1,7]$
Planar UI	$0.81_{\pm 0.78}$	II: $0.05_{\pm 0.29}$	Planar UI	II: $1.61_{\pm 0.87}$	II: $39.30_{\pm 21.65}$	II: $0.19_{\pm 0.10}$	II: $5.65_{\pm 1.26}$
Pseudo-haptic UI	$0.84_{\pm 0.36}$	I: $0.03_{\pm 0.19}$	Pseudo-haptic UI	I: $1.75_{\pm 0.89}$	I: $33.23_{\pm 18.48}$	I: $0.18_{\pm 0.10}$	I: $5.77_{\pm 1.34}$
$F_{(1,1854)} = ..$	1.62	4.09	$F_{(1,1610)} = ..$	10.20	36.71	4.03	3.98
$p < ..\ (\eta^2 = ..)$	0.20 (0.01)	0.05 (0.01)	$p < ..\ (\eta^2 = ..)$	0.01 (0.01)	0.01 (0.02)	0.05 (0.01)	0.05 (0.01)

Table 4.4: Objective performance measurements (speed, error rate) and subjective preference feedback ratings (user experience, task workload, motion sickness, immersion). The significantly best interface per scale is denoted in green. Furthermore, the ranking for each scale is represented by Roman numerals.

and 0.86 ($SD = 0.37$) for vertical layout on average. Selection accuracy was determined by the number of unnecessary interactions between the start and end of the task, i.e. 0 would be the best possible result, and higher is worse. The highest accuracy and the respective lowest amount of interaction on average was recorded for the pseudo-haptic interface (see Table 4.4); circular was the best layout ($M = 0.01$, $SD = 0.07$) and vertical the worst ($M = 0.07$, $SD = 0.32$). A univariate ANOVA showed no significant differences between the two UIs concerning speed, but a significant effect between the layouts ($p < 0.01$, $F_{(1,1854)} = 5.57$, $\eta^2 = 0.01$). Regarding accuracy, there was a significant effect between the UIs ($p < 0.05$, $F_{(1,1854)} = 4.09$, $\eta^2 = 0.01$), and the layouts ($p < 0.01$, $F_{(1,1854)} = 12.71$, $\eta^2 = 0.02$). Bonferroni corrected pairwise comparisons showed significant differences between circular and vertical layouts for speed and accuracy.

Experiment 2: User's Preference

We collected a variety of feedback to assess *UX*, *workload*, *motion sickness* and *immersion*, to ensure comparability with VR evaluations [161, 282].

The UEQ scales [157] cover classical usability (efficiency, perspicuity, dependability) and UX aspects (attractiveness, novelty, stimulation). The higher the score, the better. A univariate ANOVA showed significant differences between the two UIs regarding overall UX ratings ($p < 0.01$, $F_{(1,1610)} = 10.20$, $\eta^2 = 0.01$). Averaged over both UIs, UX for the pseudo-haptic UI was rated the best overall at 1.75 ($SD = 0.89$) on a scale between -3 (very bad) to 3 (excellent) (see Table 4.4). A multivariate ANOVA showed significant differences regarding the UEQ subscales (see Figure 4.15):

- Attractiveness ($p < 0.05$, $F_{(1,1610)} = 4.17$, $\eta^2 = 0.01$).
- Perspicuity ($p < 0.01$, $F_{(1,1610)} = 12.04$, $\eta^2 = 0.01$).
- Stimulation ($p < 0.01$, $F_{(1,1610)} = 17.11$, $\eta^2 = 0.01$).
- Novelty ($p < 0.01$, $F_{(1,1610)} = 20.41$, $\eta^2 = 0.01$).

NASA-TLX is a commonly used questionnaire to assess task workload based on six factors (mental, physical and temporal demand, effort, performance and frustration) [107]. The lower the rating, the lower the workload. When analyzing the workload using a univariate ANOVA, we found a significant difference between the two UIs ($p < 0.01$, $F_{(1,1610)} = 36.71$, $\eta^2 = 0.02$). The task workload of the pseudo-haptic UI was rated the best (see Table 4.4). Considering the subscales, a multivariate ANOVA showed significant differences between planar and pseudo-haptic (see Figure 4.15):

- Mental Demand ($p < 0.01$, $F_{(1,1610)} = 11.08$, $\eta^2 = 0.01$).
- Temporal Demand ($p < 0.01$, $F_{(1,1610)} = 66.92$, $\eta^2 = 0.04$).
- Performance ($p < 0.01$, $F_{(1,1610)} = 23.34$, $\eta^2 = 0.02$).
- Effort ($p < 0.02$, $F_{(1,1610)} = 6.30$, $\eta^2 = 0.01$).
- Frustration ($p < 0.01$, $F_{(1,1610)} = 36.50$, $\eta^2 = 0.02$).

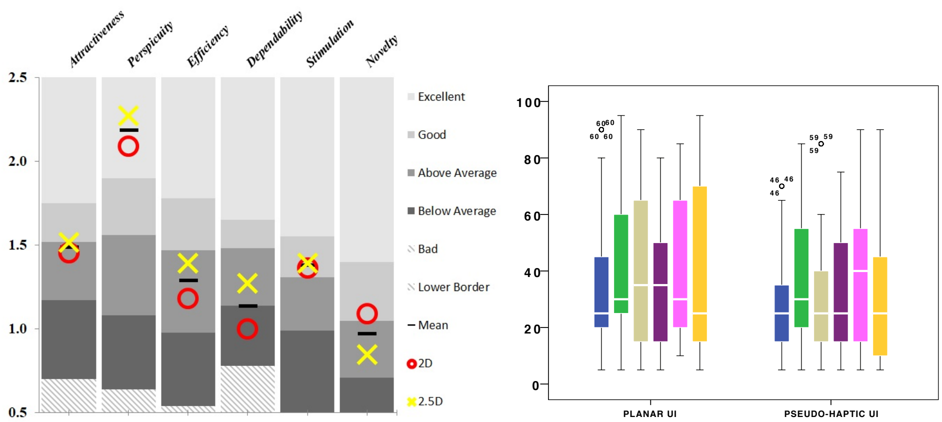

Figure 4.15: *Left:* UEQ results with respect to comparison benchmarks (see shaded boxes); for readability, the range is between 0.0 and 2.5, while the original ranged from -3 to 3. **Right:** NASA subscales (from left to right: Mental (blue), Physical (green) and Temporal Demand (beige), Performance (purple), Effort (pink), Frustration (orange)).

The Motion Sickness Assessment Questionnaire (MSAQ) assesses the motion sickness based on 16 questions rated on a 9-point scale [93]. The lower the score, the better. A univariate ANOVA showed significant differences regarding motion sickness between the planar and pseudo-haptic UI ($p < 0.05$, $F_{(1,1610)} = 4.03$, $\eta^2 = 0.01$). Overall, pseudo-haptic UI was the interface with the lowest sickness (see Table 4.4.

Finally, we used the common Slater-Usoh-Steed (SUS) questionnaire to measure the user's immersion in a VE [278]. The higher the score, the higher the immersion and presence. An univariate ANOVA showed a significant effect between the two Uis regarding immersion ($p < 0.05$, $F_{(1,1610)} = 3.98$, $\eta^2 = 0.01$). The pseudo-haptic UI was rated higher ($M = 5.77$, $SD = 1.73$) than the planar UI ($M = 5.65$, $SD = 1.54$).

4.2.6 Discussion

Current VR apps and their interaction techniques for menu selection are mostly based on guidelines and standards from non-VR areas. Therefore, we combined several metrics typically used for measuring (1) task performance, as well as (2) users' preferences based on the VRUX model (see Section 3.4).

Task Performance

The metrics evaluated in the first experiment indicate to what extent users are able to cope with the task, i.e. the UI and layout.

User Interfaces Overall, planar UI was the fastest on average, but not significantly, and pseudo-haptic was the most accurate interface, which rejects $H_{1.1}$. We did not expect these results for the pseudo-haptic UI, where the participant had to move her finger to the button and exert force to press the button through the trigger circle. This interaction requires more action than just touching the button to select it (see planar UI), and consequently takes more time. However, the results indicate that the pseudo-haptic UI design, which is more affordance-oriented and based on physical metaphors, is less error-prone than the planar button interaction. Nevertheless, the technical and physical limitations of finger tracking techniques, especially for the Leap Motion device we used, still have a crucial impact on the task performance. It should be considered that the accuracy attainable by the human hand is around 0.4 mm on average, while the Leap Motion achieved 1.2 mm on average, and comparable consumer controllers (e.g. Kinect) have not been able to achieve this accuracy [323].

Layouts Besides the UIs, a decisive factor concerning VR menu interfaces was the layout condition. Here, the circular layout turned out to be the fastest and most accurate solution to arrange buttons in a VR menu ($H_{1.2}$). This can be mainly explained by the fact that, initially, every button has the same distance to the hand, in contrast to non-circular layouts. Finally, the validity and comparability of our results are confirmed by the minimal difference of our speed and accuracy results from those of related studies [152]. The main difference with our work is that we use a HMD as output device. In addition, our results show that the circular layout is more efficient than the horizontal layout for VR, in contrast to the findings by Kulshreshth [152].

User Preference

In our concept, UX includes classical usability (Efficiency, Perspicuity, Dependability) and special UX aspects (Attractiveness, Novelty, Stimulation). The pseudo-haptic UI has the best overall UEQ score, which proves $H_{2.1}$. Both UIs have Attractiveness and Efficiency "above average", as well as "excellent" Perspicuity; planar was preferred over pseudo-haptic regarding Novelty. We assume that those results were affected by the familiar behavior of the physical metaphors. In contrast, planar UIs are known from "everyday devices" like smartphones and the corresponding drift away from analog pseudo-haptic metaphors in today's UIs (e.g. the evolution of buttons in UI/UX design[37]). Considering Stimulation, participants found it exciting and motivating to use the pseudo-haptic UI, which could be explained by its affordance-oriented design.

Task workload includes mental, physical and temporal demand, as well as subjective effort, performance and frustration ratings [107]. Regarding the

[37] https://goo.gl/UrqLq8

overall task workload, as well as every single subscale, the pseudo-haptic UI was rated best ($H_{2.2}$). The explanation for the results could be the same as in the first experiment. Concerning frustration, it is worth mentioning that many participants had problems using the planar slider, in particular when the task was to enter a precise value. Releasing the planar slider, micro movement (e.g. jitter caused by the Leap Motion or due to hand tremor) caused a change of the entered value at the last moment. This was perceived as very frustrating by the participants and might affect the ratings for the planar UI in a negative way. This limitation was also observed for the pseudo-haptic UI, but not so clearly dominant.

Finally, motion sickness ratings reflect how nauseous or unpleasant the user feels [93]. We also included immersion as part of the characteristics of VR [169]. As there was a significant effect in motion sickness and immersion between the UIs, both scores indicate that participants felt quite comfortable and the VE we used provided high immersion. Moreover, we decided to use the same environment and output device for all conditions, and short interactions to ensure that motion sickness would not negatively influence the performance and preference results. Pseudo-haptic UI achieved lower motion sickness ($H_{2.3}$) and higher immersion ratings ($H_{2.4}$) than the planar UI, but with low effect sizes.

Guidelines for VR Menus using Mid-air Finger-based Interaction

We wrap up this section with a set of general design guidelines for menu control in VR using a mid-air finger-based interaction, highlighting the major points observed during the studies, the results and lessons we learned while moving from the task to the design and development of VR UIs.

(1) *Focus on pseudo-haptic controls for user preferences and for performance-oriented menu interfaces regarding accuracy.*

As indicated by our experimental results, using pseudo-haptic controls is optimal for UX, task workload, motion sickness and immersion, and showed an overall benefit for interaction. However, if speed is of the highest importance, planar controls arranged in a circle could be considered, as they performed better on average, but not significantly.

(2) *When touching an element, the neighboring controls should be disabled.*

Avoiding accidental and unintentional touches of elements below the actual object of interest was found to be a major issue during our preliminary studies. Disabling all neighboring controls of an element upon initiation of the interaction mitigates this issue. Alternatively, restriction of the interaction to the finger serves the same purpose, but limits the freedom of interaction.

(3) *Movable control elements require fixation upon release.*

The moment when the finger or hand is removed from a movable control element, such as a slider, is crucial for successful completion of the action. When control elements are not fixed at that point, undesired changes in the input cause user frustration. For an optimal UX, it is essential to correctly address this phase of the interaction.

(4) *The absence of haptic feedback demands audio feedback.*

When there is no haptic feedback provided, users will have to rely on other modalities. Considering Chan et al. [45]'s findings, the lack of sound and any other visual aids will make it harder for users to estimate distances to virtual objects and thereby decrease the overall performance.

4.2.7 Conclusion

We have studied mid-air finger-based menu control in VR using Leap Motion mounted on an HMD. We designed and implemented two UIs including menu control elements. Although the general conclusion is to choose pseudo-haptic UIs for VR, its use is dependent on certain criteria and limitations, e.g. tracking. Our results and observations help designers to build more usable and effective VR UIs and move towards a stronger theoretical basis using principled guidelines. Future VR systems may be designed to enable the user to remain in VR for longer periods of time, demanding menu interfaces to increase usability and intuitiveness. Finally, the qualifying UIs and controls need to be evaluated in a longitudinal study in the context of interactive immersive VEs. In this work, an extensive library of UI widgets has evolved, which expand standard UI widgets available for development.

Chapter 5

Virtual Reality Concepts for Shopping Experiences

This thesis gives insights on applications and concepts for shopping experiences using VR. As discussed in Chapter 2, immersive VR systems have the potential to increase UX and decrease workload vice versa. At the same time, such VR systems require the developers and designers to be well prepared for VR-specific challenges to ensure good UX (see Chapter 3). The concept of VR combines digital content (e.g. search functionality of online shops) with physical aspects of the real world (e.g. product size). Current online shops may be functional and efficient, but do not offer enough of an immersive shopping experience [41]. Not only have technological changes in online or pop-up shops, and the current wave of digitization of retail brought economic benefits, they also caused a change in strategy, with retailers increasingly placing greater emphasis on customer satisfaction and the shopping experience. This is why it has already been invested increasingly in research in recent years and decades to improve the performance and usability of online shops' UIs. However, it is just as important for the performance of such UIs as it is for the customer's satisfaction and experience to provide the user with interactivity and information in an appropriate and supportive manner.

Online shops usually only offer ordinary 2D content (e.g. product photos or advertising videos) and use simple 2D interfaces, which are mainly

Figure 5.1: Important aspects and dimensions of shopping in VR.

used in a classic way with mouse/keyboard on the PC or via touch on smartphones or tablets. Here, the product sales are in the spotlight, and products must be found as quickly as possible for the sake of convenience and conversion rates. This focus comes at a cost as it leads to limited search functionality, confusion and product visualization [169]. While the common list-based approach using scrolling or page-based navigation can have good usability ratings, especially in the search for products, it abstracts from the actual "3D world" of a store and neglects the important aspect of UX and immersion, especially with increasing number of products and categories.

A VR shop could benefit from its third dimension and 3D interfaces such as 3D graphics or natural metaphors. It is claimed that shopping in VR offers a better shopping experience than two-dimensional e-commerce systems and that 3D applications are feasible for e-commerce [282]. Clearer content presentation and more adaptive UIs, which are designed for the tasks at hand, could lead to more positive consumer feedback and shopping experience. The use of VR systems in the retail sector has recently gained importance in the form of commercial applications and is becoming a new trend (e.g. eBay[38], Macy's[39], Saturn[40]). But there is a lack of user-friendly and intuitive UIs and interaction techniques [41], as well as a connection to previous findings from research on VR [180, 141] and HCI [169].

The remainder of this chapter is structured as follows (see Figure 5.2). Section 5.1 describes related work regarding shopping in VEs, with focus on

[38] https://vr.ebay.com.au/
[39] http://bit.ly/2ULX1f7
[40] http://locationinsider.de/saturn-startet-virtual-reality-shopping/

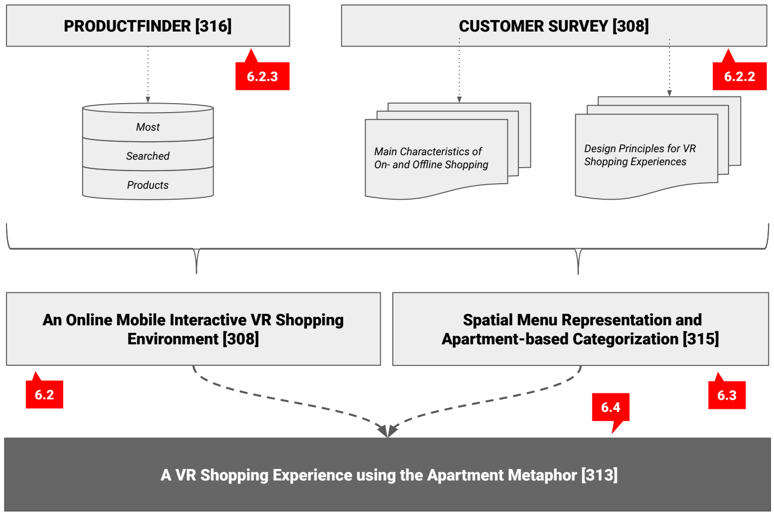

Figure 5.2: Chapter Overview.

off- and online shopping, VR in retail and a brief overview of commercial VR shop applications. As part of the user-centered design cycle, we conducted a customer survey (see Section 5.2.2) to understand the user and explore the main characteristics between on- and offline shops (see Section 5.2.2), followed by the definition of principles for designing interfaces for VR shops based on related work and our survey results (see Section 5.2.2). This dataset, characteristics and principles formed the basis for our experiments and the development of our VR shop prototypes.

In Section 5.2, we present a user study with a product search task based in a WebVR online shop using speech input in combination with VR output. In a subsequent study (see Section 5.3), we compared traditional linear store representation and categorization of online shops with a novel approach using the grid-based Apartment metaphor. As a concluding system (see Section 5.4), we adapted the Apartment metaphor to the representation a VR shop and explored selection and manipulation of virtual products in a third study. The VR shop prototypes described in this chapter were evaluated with respect to our VRUX model (see Section 3.4).

The concepts and findings of this chapter have been published previously in the following publications: [281, 290, 282, 289, 287]

5.1 Background

The investigation and evaluation of VR shopping approach from different domains. Specifically, we identified (1) menu representation and categorization for online shops, (2) VR in retail, (3) alternative VR shopping concepts, and (4) commercial VR shop applications.

5.1.1 Menu Interfaces for Online Shops

Prior works relevant to menu interfaces for online shopping include those addressing (1) menu representations, as well as (2) menu categorization which will be discussed here. Furthermore, we present an analysis of the menu structures of current popular online stores.

Menu Representation

Menu-guided interfaces have long been a prominent user interface component of software applications or websites and serve to structure the underlying amount of information hierarchically. Therefore, a variety of different menu representations have been explored in the past. Miller et al. [204] examined the effects of menu width and depth on speed and accuracy. Their results were confirmed by Zaphiris et al. [337], which showed that flat hierarchies are easier and faster to use and lead to greater orientation and satisfaction, and shorten interaction times. This has been confirmed by Zhang et al. [341] in their extensive study on web shop menus. While early work showed only a tendency towards menus at the top of the screen [212], top-positioned menus are faster for product search [341], while top and left-positioned menus were preferred. Based on these findings, we have developed a reference menu for our comparison study (see Section 5.3).

Cockburn et al. [55] investigated different approaches to improve the traditional linear menus in terms of performance and preference. They compared standard and shared menus [266], and showed that frequency split menus were the fastest. However, the use of frequency split menus in an online shop might be problematic as highlighting frequently selected but potentially unwanted objects would not lead to any improvement, and

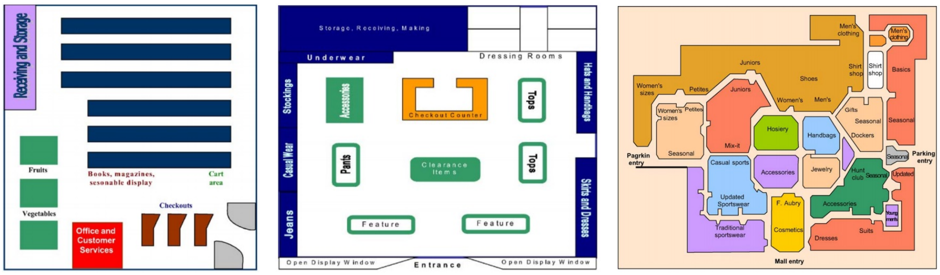

Figure 5.3: From left to right: Grid, Freeform, and Racetrack layout [315].

could lead to frustration and confusion [207]. Findlater et al. [82] used a prediction algorithm to determine menu items that the user probably needs and showed that they enable faster menu selection. As shopping is a changing process – both, the offer and the interests of customers are constantly varying – the use in online shops does not appear promising for the process of product exploration and is therefore not considered.

Ahlstrom et al. [6] compared different menu designs and their results show that a squared menu is faster and is preferred, which indicates that a spatial and rasterized arrangement could lead to better usability. Similarly, items are structured hierarchically in a grid in the *CommandMaps* menu [259], which is faster and requires fewer pointing activities compared to a ribbon menu interface (known from Microsoft Word), since the user no longer needs to change menu levels and benefits from spatial memory. In the same line of exploiting spatial memory also Vrechopoulos et al. [315] examined how different "brick-and-mortar" layouts can be transferred into a virtual retail environment. Their study again showed that a grid-based on a tree structure is easier to handle, which is furthermore supported by Griffith [99].

In summary, an adjustment of the menu representation can lead to higher performance [204, 82, 259] and/or better preference [337, 55, 6]. Deeper menus, i.e. with many layers, increase complexity and slow down interaction times, so that the width (number of items in the layer) is preferable to depth [156, 225, 337]. Moreover, it is advisable to minimize the menu depth without increasing the width extremely and on condition that the semantic data allow such a distribution [204]. The reference menu used in this work is based on [341], representing the state-of-the-art of traditional linear menus in current online shops. Furthermore, we developed a spatial menu representation in the line of prior work [315, 99] (see Section 5.3).

Menu Categorization

Besides the menu layout, especially a meaningful and comprehensible categorization of the menu items in e-commerce environments is important for efficiency [136, 264] and user-friendliness [310]. Larson and Czerwinski [156] recognized the importance of a semantic and integrated this fact into their research on width and depth of menus. Usually one differentiates between hierarchical (or faceted) categories and automated clustering [110]. While fully automated clustering according to the similarity of words or phrases has the advantage here having a quick structuring of information collections, it often leads to logical inaccuracies in contrast to the manually created hierarchies, which are preferred by users [110]. Resnick and Sanchez [249] confirm the influence of high-quality menu labels.

Adam et al. [3] presented a new categorization and representation scheme to enable intuitive menu navigation. Their spatial Apartment metaphor maps the mental model of an apartment with different rooms (living room, kitchen, etc.) to a structure of a smart home control interface. The top level of categorization corresponds to the rooms, followed by the devices and finally the tasks that contain the potentially possible system tasks of the selected device. The positive effects of the Apartment metaphor on performance and preferences serves as a basis and motivation for the menu categorization of our online shop prototype (see Section 5.3) and VR apartment shop prototype (see Section 5.4).

State-of-the-art Analysis

Even though related work recommends to abandon linear store-based menu interfaces, most current online shops are still employing these. Table 5.1 lists a selection of popular online shops and their used categorization and representation. While we do not claim this to be the most representative selection it contains some of the most frequented stores. All of them use a linear representation in which the items are arranged horizontally or vertically. Some of them also integrate a multi-column menu display, e.g. the *IKEA* interface contains a two-column menu at the top level. There are however significant differences in the menu width from 3 to 24 items, in contrast to the depths between 3 and 5 levels. This is in line with findings from previous work [156]. In current online shop menus, the individual items are predominantly text-based, while current research recommends more graphical methods. Only three of the eight online shops (*REWE*, *CARREFOUR*, *IKEA*) additionally integrate icons illustrating the associated text label, while three menus (*CONRAD*, *REWE*, *TESCO direct*) also show the number of sub-level items.

Besides the representation, the logical meaning of the underlying categorization is important. All considered menus in Table 5.1 use mixed categories, i.e. following different sorting strategies within a menu level. Mostly, the categories are based on a combination of assortments and themes known from physical shops, e.g. "beverages" refers to an assortment and "baby" to a theme. Studies have shown that such a mixture of categories can be unclear to the users, since the labels do not clearly describe the underlying information space, which is essential to give the user an overall impression of the search space [110]. Especially occasional or new customers

online shop	**categorization**	**representation**
amazon.com	product range, theme	$(4, 20)$
conrad.com	product range, theme	$(4, 8)$
ikea.com	product range, theme, room	$(3, 24)$
carrefour.fr	product range, theme	$(3, 17)$
rewe.de	product range, theme	$(3, 12)$
tesco.com/groceries	product range, theme	$(4, 11)$
tesco.com/direct	product range, theme	$(5, 12)$
zalando.com	product range, theme, target group	$(4, 3)$

Table 5.1: Overview of state of the art online shops for groceries, electronics, furniture or fashion. First number is the menu depth, second is the width of the top-level. Accessed on August 21st, 2018.

of a new or existing shop interface could be particularly frustrated, resulting in a negative effect on performance and preference. In the worst case, it could lead to shopping attempts being canceled and the shop not being visited again [189]. Our analysis indicates that current online shops do not follow findings on menu optimization and tend to use classical methods.

5.1.2 Virtual Reality in Retail

In the last decades, more and more virtual shopping environments have emerged, but mostly very simplistic. Some researchers claim that VR shopping environments provide a „more feel experience" compared to 2D e-commerce environments [342, 46] and 3D applications for e-commerce are feasible [18]. Sanna et al. presented a virtual shopping environment generated from preferred products the users chose before [257]. This all indicates that personality and feasibility are crucial factors and should be addressed when designing VR shopping experiences. AWE3D and ADVIRT [48] are architectures for adaptive 3D e-commerce websites. Here the store layout, organization, and appearance were adapted to the personalized rules of the user. Nevertheless, these 3D shop concepts only focus on adaptability, not availability, for VR technology. But the technology to develop e-commerce applications for immersive VR already exists. Laver et al. developed a grocery shopping simulator to investigate the interaction between the user and the program [159]. Despite they focused on usability, their results indicate that VR could have a strong impact on immersion. Although more and more VR shopping environments have emerged recently, they are still very simple and immature. In most instances, physical stores are merely virtualized and digitized, i.e. 3D models of the products are placed in a 3D representation of a typical existing store (see Figure 5.1.2).

Currently, a major problem of online shops is the lack of clarity, realism and immersion. Buffa et al. describe further advantages of 3D virtual stores in comparison to physical stores [41]. They state that customers benefit from less time-consuming shopping, 24/7 opening hours and more product

[37] https://invrsion.com/

[38] https://vr.ebay.com.au/

[39] https://goo.gl/h22ezQ

Figure 5.4: ShelfZoneVR[41], eBay[42], and Macy's VR[43].

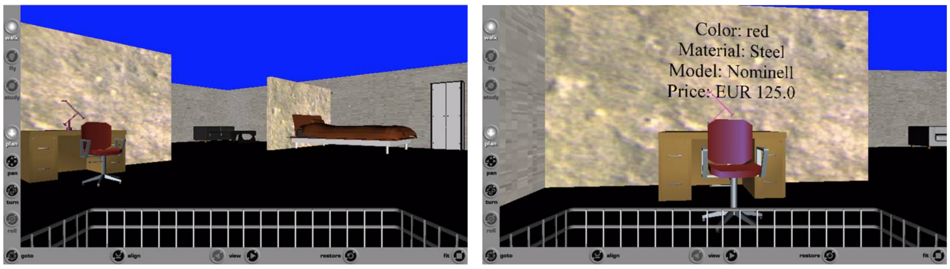

Figure 5.5: Views of the VR furniture shop Shop-WISE [67].

information. Adding a third dimension could fill these gaps in the virtual shopping experience. In their point of view, customers will then benefit from always-open unstaffed warehouses, time-saving shopping and multi-modal product information. This applies for online shops, but not for brick-and-mortar stores. Overall, these findings have provided the basis on which we investigated the benefits of online versus offline shopping and combined them in the form of interactive VR shopping environments.

The essential characteristics of VR are interactivity, immersion, and presence [169]. Bhatt examines the feasibility of bringing VR to e-commerce sites and concludes that the balance between the three characteristics is necessary and dependent on the circumstances [20]. For example, in the fashion industry, immersion is crucial, whereas in the financial sector presence is far more important. However, as we focus on types of goods found in conventional stores types (groceries, electronics, clothes, furniture, etc.) immersion and presence could be combined into one characteristic. Walsh et al. concluded that VR could address limitations of web-based shopping applications [319], expanding the range of e-commerce possibilities, which indicates that VR has the potential to create novel and rich shopping experiences. To legitimate the need of VR in retail, Lee et al. compared the UI of a VR shopping mall with an online shop [169]. Their results indicate that online customers remain passive observers, whereas in a VR shopping mall customers are engaged in the inspection and control of the 3D visualized target products. Moreover, VR customers can experience the product dimensions more richly and engage in a more interactive shopping activity.

In Shop-WISE [67] the user is able to pick up 3D products and inspect them (see Figure 5.1.2). Besides that, Shop-WISE allows searching for products by text input and moves the user to the desired product after it has been selected from a results list. We adapted the Shop-WISE approach and extended it using 3D model representations of the search results [282], so the user gets better clarity, whether the desired product has been found, which can also persuade users to buy the product [109].

5.1.3 Alternative Shopping Concepts in VR

Ogier et al. converted a stock-on-shelf store, which offers potential advantages like increased immersion, to a list-based store interface and compared them in a game-like setting [228]. The major drawback of stock-on-shelf interfaces is that the in-game size of the purchasable stock needs to fit in the in-game spatial representation of the store's shelves. But the store size cannot be increased without affecting the narrative. So they conclude that list-based shop interfaces are not appropriate for VR applications; even simple list-based UI elements can be disruptive and cause motion sickness [228].

In large retail stores, products of the same category are located in one area. Thus, they lack search functions, and one of the most mentioned disadvantages of physical stores is the issue of product search [282]. Anecdotal evidence suggests that customers store products of different categories at similar places at their home, e.g. food and cutlery in the kitchen. Magic Home [320] introduces a concept prototype featuring a VR furniture store, where customers walk inside a local physical store and try out the furniture they want to buy. In addition, customers can get a preview of the furniture inside a virtual representation of their home, which is connected to the store. So the customers can decide how well the furniture fits inside their home using the advantages of the virtual world, while for example sitting on a couch in the physical store. This approach inspired our *Apartment* metaphor [289], which was evaluated and used in the experiments in Section 5.3 and 5.4.

5.1.4 Commercial VR Shop Applications

Commercial solutions are also appearing on the market. A very recent example, albeit not yet market-ready, is the VR department store created by eBay[44] in collaboration with the shopping chain Myer. Customers can browse through eBay's product categories using head pointing in a mobile app and a mobile VR headset (e.g. Google cardboard). The products are represented by rotating 3D models, along with some side information (e.g. delivery date and price). The major drawback is the limited interactivity, as only head pointing with dwell time for selection is used (see Section 2.4.2). This can limit performance and UX because of higher workload [286].

Another commercial approach for a VR shopping application is ShelfZoneVR[45], which is a retail space simulator reproducing physical shops. Customers can freely move through the store using Point and Teleport (see Section 2.4.3) and grab the products with the HTC Vive controllers using the Virtual Hand technique (see Section 2.4.2). Here, the shopping environment is an exact representation of a physical store, where the shelves are filled with products entities. As mentioned before, the presence of multiple items of the same kind (stock-on-shelf) makes the store more complicated and increases unnecessarily the needed virtual space and store size [228]. However, we adapted their realistic product selection [287] to evaluate and compare it with a non-isomorphic selection technique (laser beam [27]).

[44] https://vr.ebay.com.au/

[45] https://invrsion.com/shelfzone

5.2 Experiment 1: An Online Mobile Interactive Virtual Reality Shopping Environment

In this work, we designed and implemented an immersive online shopping environment in VR. We tried to maintain the benefits of online shops (see Figure 5.8), like search functionality and availability, while simultaneously focusing on shopping experience and immersion. By touching the third dimension, VR provides a more advanced form of visualization, which can increase the customer's satisfaction and thus shopping experience. A case study of a first VR shop prototype was conducted and evaluated with respect to the VRUX model (see Section 3.4). The results showed that usability and UX of our system is above average overall. In summary, searching for a product in a VR online shop using speech input using a HMD proved to be the best for performance (speed, error rate) and preference (usability, UX, immersion, motion sickness).

5.2.1 Introduction

Current online shops may be functional and efficient, but do not provide enough of an immersive shopping experience [41]. This work focuses on developing an immersive VR online shopping environment that includes the major advantages of offline and online shopping (see Figure 5.8). The main goal of the work was to do the next step in investigating interaction in mobile VR shopping environments by taking a closer look on the influence of input (head pointing, speech) and output (desktop, HMD) on task performance and user's preference and behavior. In order to provide a realistic setting, we evaluated two common hands-free interaction techniques using smartphone VR in a comparative user study.

Technological changes, in the form of online shops, pop-up stores and digitization in general, have not only brought economic benefits to the retail sector. There is also a change in strategy underway, in which retailers will put more and more emphasis on satisfying the customer. Therefore, a lot has been invested in research to improve performance and usability of online shops' user interfaces. Nevertheless, it is equally important for the

Figure 5.6: Screenshot of our Virtual Reality Online Shopping Environment

effectiveness of such user interfaces as well as for the customer's satisfaction and shopping experience to provide the user with the desired information in an appropriate and supportive way.

Using VR systems in the shopping area has grown in its importance and has emerged to a new trend to create virtual stores. Retailers initiate these VEs to take customers along experiences. Although VR shopping is still at the very beginning, this could fill the gaps of common online shopping (e.g. customer satisfaction, and experience) and become one of the more popular ways to shop online. With the fast-growing market of VR hardware and the knowledge that retailers have gathered in the last decades through online shopping, the process is mutually beneficial.

5.2.2 Customer Survey

In the run-up to our work, we conducted an exploratory customer survey ($N = 41$, 14 female, aged between 18 and 58 years) with a focus on positive and negative aspects of on- and offline shopping, as well as shopping behavior and frequency. Based on the results of this survey, we designed and developed our VR shop concepts addressing the issues and including the benefits of on- and offline shopping. They were recruited in a collaborated German retail hypermarket and volunteered uncompensated.

On- and Offline Shopping Frequency and Behavior

We asked them on a scale from 0 to 5 (0: never, 5: daily), how often they go shopping on- and offline. The results indicate that customers prefer offline ($M = 3.91$, $SD = 0.62$) than online shopping ($M = 2.45$, $SD = 1.02$), with 98% offline at least once per week, 83% several times a week, but only 7% daily (see Figure 5.7).

We further asked them about their shopping behavior on a scale from 0 to 5 (0: never, 1: only offline, 5: only online) regarding food, textiles, clothes, furniture, and electronics (see Figure 5.7). Food ($M = 1.29$, $SD = 0.55$) and furniture ($M = 1.62$, $SD = 0.86$) are bought more locally than online, as well as textiles ($M = 2.45$, $SD = 1.13$) and clothes ($M = 2.55$, $SD = 0.97$) mostly locally but occasionally online. Asked on a scale from 1 to 5 (1: doesn't

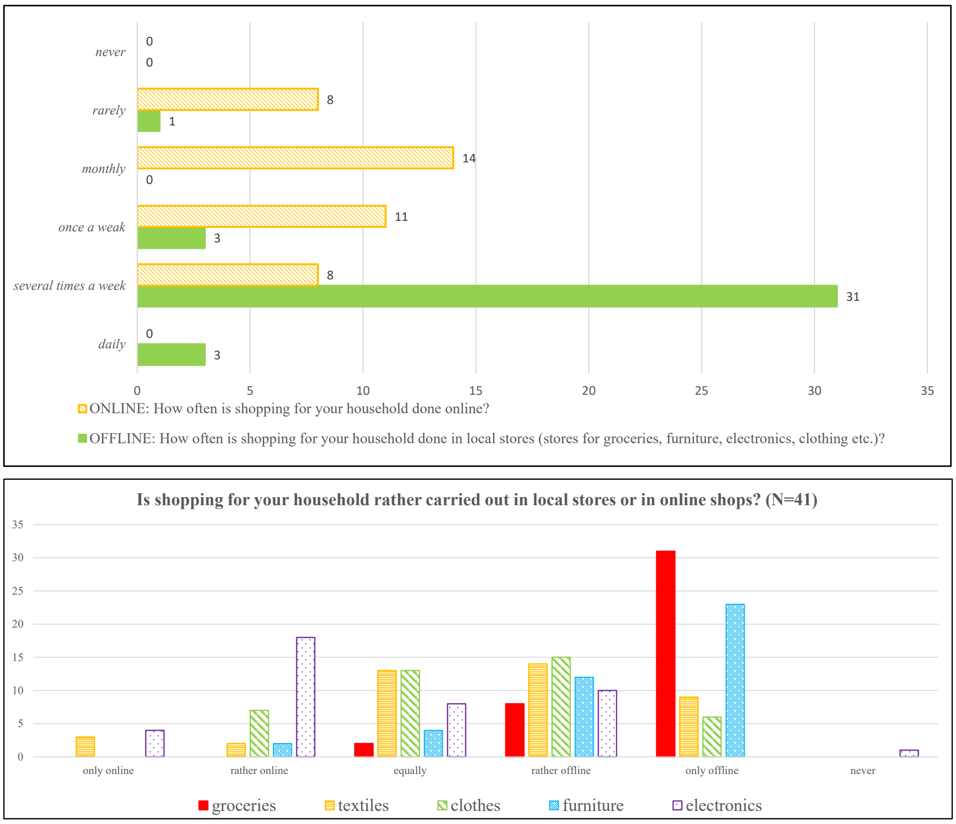

Figure 5.7: Shopping Frequency and Behavior Online vs. Offline with regard to groceries, textiles, clothes, furniture, and electronics (N=41).

apply at all, 5: applies fully), whether they find it easy to find a desired product in offline and online shops, over 78% preferred online ($M = 4.05$, $SD = 0.85$) compared to offline ($M = 3.45$, $SD = 0.89$).

Main Characteristics of On- and Offline Shopping

Regarding the positive aspects of online shops, 68% of the participants mentioned the search functionality (see Figure 5.8). In each case, 12% mentioned that the search functionality is often badly implemented, and the display of products in lists or tiles is unstructured and confusing. As benefits of offline shops, most mentioned employees (51%), signs (34%), and the use of departments and shelves as categories (24%).

In summary, today's online shops are generally acceptable in their search quality and functionality, but lag behind because of its often confusing and unstructured interfaces, which finally results in lower UX and customer's satisfaction in general according to participants' comments. This is also

	positive aspects (N=102)	**negative aspects** (N=30)
ONLINE (N=60)	• **search functionality (28)** • **filter / categories (10)** • ratings / feedback / tests (8) • experience (1) **∑ = 47**	• **confusing / unstructured (5)** • **bad search function (5)** • too big range of products (2) • no filters (1) **∑ = 13**
OFFLINE (N=72)	• employees (21) • **signs (14)** • **shelf names / departments (10)** • habit (10) **∑ = 55**	• **store too large / lack of clarity (8)** • **not sorted properly (4)** • **not sorted (4)** • too frequent changes in sorting (1) **∑ = 17**

Figure 5.8: This figure illustrates the main characteristics of off- and online shopping mentioned by the participants of our customer survey (N=132, with N as the number of mentioned items).

reflected by the product's dimensions and representation in ordinary online shops. The majority of participants of the customer survey reported that the product representations in the form of colored 2D pictures in scrollable lists are hard to understand, i.e. the user doesn't get a clear sense of its size, shape, or weight. In addition, the user is forced to scroll the list or navigate through several pages, so the workload (effort, frustration, etc.) and UX (attractiveness, stimulation, etc.) may thereby be impaired.

Designing Interfaces for VR Shopping Experiences

When designing UIs for VR shopping experiences, the main task of the application should be clear, so that the designer can move from task to design. The main benefits should be respected while avoiding the mistakes that surround on- and offline shops according to customer survey findings (see Figure 5.8). Therefore, we provide a list of potential guidelines based on our survey results that could be used to inform the design of VR shop UIs:

i. **CLARITY.** The scene should be as simple as possible, but keep the feeling of being in a store looking for products. Therefore, signs, shelves and departments as visual categorizations and cues can help the customers to be better oriented and feel more immersive. And an increasing number of products in a store requires customer-friendly filters (e.g. allergens) and a proper categorization.
   But: *Too many and confusing categories and a lack of clarity of the filters may also influence the performance and experience. Do not overfill the shelves with products. Keep in mind that in virtuality, there is no need to overfill the shelves with duplicates of products.*

ii. **EFFICIENCY.** Provide the user with a user-friendly and efficient search functionality. Here, search filters can help to select products by their attributes and features or by their description texts.
   But: *If you are using text-based methods, there will be a strong need of auto-completion, spell and synonym checker. Otherwise the usability and performance might be influenced negatively. Overall, text entry is unsuitable but also unavoidable in everyday VR scenarios. Nonetheless, the focus should*

not only be on task performance, as workload and motion sickness have been proven to be as decisive as pure efficiency for optimal UX ratings.

iii. **ORIENTATION.** Use signs, shelf names or numbers in combination with a proper and useful categorization to give the users a better orientation in the virtual market. In addition, depending on the size of the market, the products and their shelves should be sorted properly in visually distinguishable departments.
But: *Do not overload the user with disturbing colors, unintuitive names or „too much information". A clearly arranged assortment makes it easier and faster for the user to find and select the desired product. Furthermore, using HMDs for output are preferred over 2D screens, e.g. for object manipulation in a furniture arrangement application [284].*

iv. **PERSONALITY.** An essential factor is the personal aspect of a virtual shop. The customers of offline stores prefer the presence and interactivity of employees. More precisely, the ability to talk to an assistance and ask questions like „where can I find ...?" or exploring interactive floor maps [290] is preferred instead of typing text on a computer or terminal. If available, use speech to fill this gap. The speech interface can serve as the personal part of the environment.
But: *Today's state-of-the-art speech-based interfaces are still in an infancy stage. So the developers should always keep in their minds that the speech recognition ratio can have a crucial influence on the shopping experience. In addition, customers stated to avoid speech and preferred text input in public.*

v. **QUALITY.** Provide the user with sufficient information about the products (e.g. feedback, ratings and tests). Instead of providing textual detailed information about the product's dimensions, shape or weight, it can be visualized using interactive 3D object representations and gravity in the virtual environment.
But: *Incorrect size, shape or weight will influence massively the user's impression of the product and environment itself.*

5.2.3 Concept

For the future of buying products online, retailers could have a huge benefit using VR shops, in which a 3D rendering of products is created for customers to view from every side. This could lead to more satisfied customers as well as an ability to showcase detailed items, e.g. to compare them with their 3D rendered counterparts.

As ordinary online shops offer 2D content, they use simple 2D interfaces with hyperlinks, labels, icons and menus. Here, only the products are important, and you need to find them quickly, for the sake of convenience and conversion rates. The products are mostly displayed in a list or grid, i.e. data browsing is done by scrolling or paged-based navigation. While this approach can have high usability ratings, in particular for finding products, it abstracts away from the „3D world" of stores, and disregards the value of UX and immersion, especially with an increasing number of products.

A VR shop instead could profit by its third dimension and 3D interfaces such as 3D graphics, natural metaphors or avatars [281]. More vivid content representations might give a more positive consumer response. But modern media requires more complex interaction techniques, which consequently cause higher levels of user instrumentation [100]. By touching the third dimension, VR provides a more advanced form of visualization, which can increase the customer's satisfaction and thus shopping experience. It allows more natural user interfaces (e.g. camera control using head rotation) than the usual mouse/keyboard interaction in a desktop environment.

But from a technological point of view, visualizations and interfaces of VR, shopping worlds, as well as customer interactions in VR were rarely studied as well as the development of VR online shopping platforms. Thus, there are only a few results that give insights on how such technologies provide compensation for the lack of multi-sensory input and output (see look-oriented interaction with objects as in 360° product views, spatial interaction with fresh food counters [92] or interactive clothes booths) or ways to enable multi-sensory interactions online.

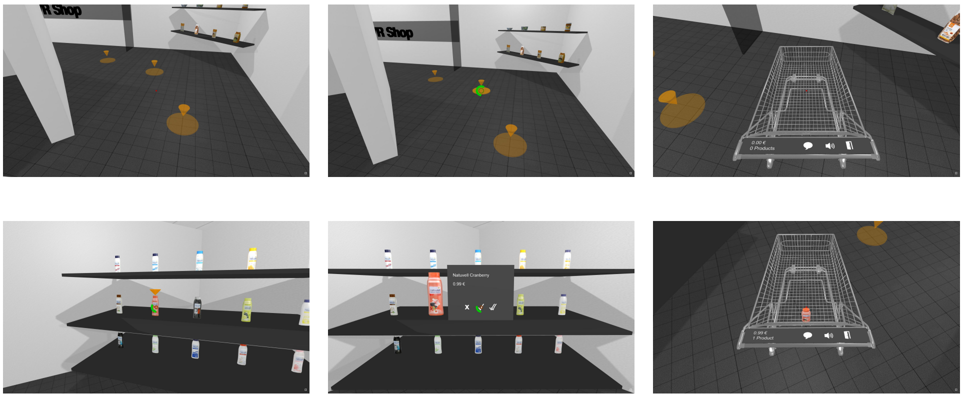

Figure 5.9: **Head Pointing.** Here, the user looks at a nearby waypoint in order to be moved. Then, a product is placed into the shopping cart after it has been selected. Finally, the total sum and amount of items is displayed on the shopping cart's handle.

Controls and Interfaces

As we focus on a mobile online WebVR application without controller and hand interaction, our system provides two different input modes: *Speech Input* and *Head Pointing*. According to the customer survey results (see Section 5.2.2), the majority of the participants valued the search functionality of online shops. But on the contrary, this can even have the opposite effect if it is error-prone. Therefore, and because we want to avoid text entry in VR, our *Speech Input* method should enable VR customers to find their desired products in an efficient way. The speech interface is located world-fixed on a wall in the scene and provides the functionality to search for products and list the results (see Figure 5.10). The user starts by gazing at the button with a speech bubble icon and speaking the search term into the microphone. After the search results are displayed, the user can select the desired product by gazing for five seconds in order to be moved to the desired shelf. When the user is looking at a result, the front view of the respective product is displayed in the middle of the search interface for better decision-support.

Another way to search for a product, and for the more exploratory characters among the users, we chose a navigation through *Head Pointing* and waypoints (see Section 2.4.3). Here, the user is moving around using the way-points and looking for it in the shelves (see Figure 5.9). Our system

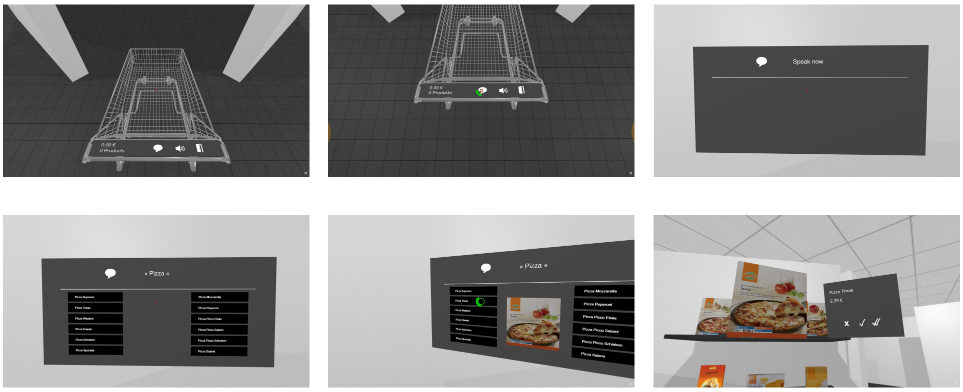

Figure 5.10: **Speech input.** Here, the user starts by gazing at the speech button, and speaks the search term into the microphone. After the search results are displayed, the user can select the desired product in order to be moved to the desired shelf.

does not include eye-tracking hardware, as it is the today's state-of-the-art for common VR installations and applications. But different concepts exist (e.g., PupilLabs' HTC Vive or Oculus DK2 integration[46], or other commercial products like SMI[47]), so it is just a matter of time until the first mobile VR hardware equipped with eye-tracking will be affordable and usable. As we wanted to focus on affordable, commodity and state-of-the-art mobile VR scenarios, we decided to use head pointing interaction instead of eye gaze.

Whereas the procedure of this input mode is straightforward, i.e. the user moves around in the VE by using (looking at) the move plates and looking around to explore the market, it has the limitation that in the worst case scenario the user has to move through the whole VE until the desired product is found. The advantage is that efficiency should increase proportional to the spatial knowledge of the market layout, i.e. once the environment is well-known, it is easy to find products by moving straight towards them. Nevertheless, it can take time to find the product in the desired shelf.

Concerning the output devices, the system automatically detects which output devices are currently available and displays the rendered scene to them. One way to use it is in a *Desktop* setting with an attached monitor (see Figure 5.11). Of course, the render quality is better than on a mobile

[46] `https://pupil-labs.com/blog/2016-02/eye-tracking-for-vr-and-ar/`
[47] `https://bit.ly/2XZMM8K`

Figure 5.11: This figures show the **output devices** we used for our prototype. On the left: 22 inch display with Microsoft Kinect 2 for speech input and a standard mouse for camera control. On the right: a common smart-phone VR HMD and LG Nexus 5X.

device, but applications in the *Desktop* setting could be lacking in immersion and UX. Another way is to use a Head-mounted Display (HMD) for mobile VR, which is cheap and affordable, in contrast to high-end HMDs (see Figure 5.11). Whereas in a *Desktop* setting, *Head Pointing* and *Speech Input* have to be tracked by an extra device (like Microsoft Kinect), both come for free using smart-phone sensors and a built-in microphone. Another benefit is handiness, because smartphones need not a tethered connection, so they can be used nearly everywhere. However, the most crucial limitation for smart-phone VR is the render quality and computing performance, which is not comparable with the power of a PC.

Product Dataset

Before we started to investigate shopping in VR, we wanted to build a realistic basis for our studies considering product data, store layout and orientation in on- and offline shops. We developed the *ProductFinder* to gather data, e.g. the most searched products and store preferences.

Orientating oneself and finding products in physical retail stores is a well-known problem, in contrast to search functionality of online shops. Common modern retail stores have up to 10,000 square meters and offer more than 100,000 products. We introduced the ProductFinder [290], an intelligent product information system for situated interactive public displays in retail environments in order to equip physical retails stores with a search functionality and make them comparable with online shops. The ProductFinder dataset including the most searched products build the fundament of our VR shop studies and make them closer to reality.

The *ProductFinder* terminal was deployed from September 2014 to August 2016, in a so called hyper-store with over 10,000 square meters and over 100,000 products. We logged interactions of customers with the devices and gathered data over 300,000 database queries and 1,200,000 touches. All data is anonymous. The following analysis is narrowed down to data from 2015, preventing seasonal distortions. We assumed a break of 60 seconds between two touches to the screen of a device as a new session (meaning a new user). We observed altogether 30,453 sessions, lasting between 18 and 71 seconds (50% of data; $M = 55.79, SD = 62.64$). Only 1,530 sessions (5.02%) consisted only of one touch to the screen, most sessions (50%) consisted of 9-30 touches. Most user interaction took place on weekend (on Sunday the store was closed) and in the evening (5pm - 8pm).

On average 1.76 ($SD = 1.34$) unique search terms were entered per session, resulting in $41,078$ search terms at all. Most searched products in 2015 were: salt (458, 1.11%), chips (420, 1.02%) and mustard (298, 0.73%).

Figure 5.12: This figure illustrates the *ProductFinder* terminal in the retail environment. The user interacts via direct touch.

most searched products	
chips	1,83 %
mustard	0,83 %
milk	0,77 %
salt	0,71 %
oil	0,54 %
spices	0,48 %
noodles	0,42 %
eggs	0,41 %
bread	0,41 %
ketchup	0,39 %

most searched departments	
for the food cupboard	32,42 %
breakfast	9,19 %
wine & spirits	7,06 %
spices	6,49 %
dairy products	4,99 %
household goods	4,97 %
crisps	4,18 %
confectionery	3,80 %
cosmetics & toiletries	3,39 %
fresh ready-meals	3,35 %

most searched shelves	
mayo / mustard	4,87 %
country selection	4,18 %
baking ingredients	3,39 %
spices	2,48 %
tea / spreads	2,23 %
butter / margarine	2,21 %
soups / fund	2,08 %
sugar / flour	2,01 %
dried fruit / condensed milk	1,97 %
spirits	1,86 %

Figure 5.13: An extract of the most popular searched terms in one year.

Most of the time, the people were looking for terms that could be matched directly to shelf names (e.g. mustard 79.48%), instead of departments (e.g. wine 15.49%) or even products (e.g. a specific product with 5.03%).

Virtual Environment and Product Placement

The VE of our VR shop was designed to be as simple as possible, but still close to a real market (see Figure 5.14). While in reality shelves are overfilled with products and the environment is decorated all over with advertisements, our approach contains only a clearly arranged assortment for quick and easy selection. This should solve one of the most common problems the customers complained of in the customer survey, namely that they struggle more with the lack of clarity the larger the market is. The environmental colors were chosen to be as neutral as possible to prevent positive or negative impact. Nevertheless, the attention of the user should always be attracted to the products, therefore they are colorful and natural, unlike the colorless surrounded elements. The contrast is emphasized by a subtle glossy reflection of the products.

Existing market layout and floor planning data provides information about dimensions, position and size of different types of store elements (walls, shelves, departments), so they could be easily parsed and rendered. Shelves contain at least one product, products are evenly distributed, and the space between each rack is computed by the market is height and the number of tiers (levels in the shelf). A product is represented by a virtual 3D model and is placed on its corresponding shelf level according to the

Figure 5.14: The VE includes the shopping cart, a move plate (orange), the speech interface, and shelves equipped with products on the right.

floor plan data. After a product has been selected, a detail view is displayed, including a larger 3D model of the product, side information like price and description, and buttons for adding it to the cart or closing the detail view.

Offline shopping experience in HCI is either context-aware [142] or social [90, 210]. Based on findings of Black et al. [22] concerning context- aware shopping experience, we integrated a virtual shopping cart into our design. The cart is always located in front of the user and holds all previously selected products, which can also be removed. A control panel at the cart's handle includes the total price, a button to open the speech interface and one to reset the scene. Besides those store elements, waypoints are displayed as move plates and spread over the whole market. In order to prevent unintentional movements, the user only sees nearby surrounding move plates. Thus for every possible user position, it is ensured that there at least two move plates in range.

5.2.4 Pilot Study

Before turning to the first main study, we describe a preliminary experiment we conducted to establish the trial time limit, appropriate product locations, different movement types and feedback modes. We conducted this study with four unpaid university students, without giving the participants feedback regarding their trial time. The pilot study task was to process items on a shopping list after a short exploration phase in the environment.

One participant was concerned about the speech interface's location, which initially was dynamically positioned at the user's position in his FoV. He or she complained about a lack of realism, so the speech interface was finally positioned at a fixed location in the virtual market. Another participant asked for functionality to add a product twice to the shopping cart. The participants had two options to move through the scene, by moving along a path or teleporting. In the path movement setting, the user is moved over a minimal path along way-points to the target position. The teleportation is based on concepts of dream research and uses the metaphor of closing the eyes. After the user has been teleported to the target position, the „eyes" were opened again. Here, the participants preferred the path movement, so it was used in the main study.

With regard to the product highlighting, all pilot participants preferred the cone highlighting method, where a cone was placed above the product the user is currently looking at, fading out all other products. So in the end, the cone highlighting was chosen as the standard feedback mode for product highlighting. As we used real products from a common German retail store, product names can be very complex to detect from speech or even pronounced by the participants. So, with regard to the planned speech input method, we used the pilot study to remove product names to be searched by the participants in each task in the main experiment, which could not be recognized easily or impossible to be recognized using the Web Speech API from Google.

5.2.5 User Evaluation

We conducted this study to gather insights into task performance (speed, error rate), user preferences (usability, UX, motion sickness, immersion) with respect to our VRUX model (see Section 3.4). Furthermore, we explored unmet needs of an online mobile interaction VR shop. The study provides us with:

- **Metrics.** Objective and behavioral performance data that provides a baseline to measure future improvements of VR online environments.
- **Customer insights.** Actionable insights on how to optimize usability, satisfaction, and experience for the customers.
- **Actionable improvements.** Concrete recommendations for improvements based on research findings.

The purpose of this experiment was to evaluate the end-to-end experience of users as they interact with our system using two different input (*Speech Input* vs. *Head Pointing*) and output (*Desktop* vs. *HMD*) methods.

Hypothesis

H_1 The task is performed faster with *Desktop* for output.

H_2 *HMD* is preferred by the user in terms of UX and usability.

H_3 *Speech Input* is more efficient (speed, error rate) than *Head Pointing* only and is preferred by the user (usability, UX).

H_4 *HMD* with *Speech Input* outperforms the others in all aspects.

Participants

A total of 16 unpaid participants (13 male, 3 female) volunteered in this experiment, aged between 17 and 55 years ($M = 23.88$, $SD = 8.52$). 69% of the participants had a university background; the rest were staff members and high-school graduates. The overall experience with VR and desktop applications, as well as the acceptance of VR shopping, were rated on a Likert-scale from 1 to 5. The level of experience with VR applications was very low overall ($M = 1.25$, $SD = 1.34$), whereas it was very high with desktop applications ($M = 4.25$, $SD = 0.93$). When asked whether they would go shopping in VR in the near future, the answers varied widely ($M = 2.88$, $SD = 1.31$). Most of the participants tend to buy more often per month offline ($M = 7.56$, $SD = 5.48$) than online ($M = 2.93$, $SD = 2.23$).

Apparatus

For the two desktop tasks a standard desktop computer was used with an i7 CPU, 16 GB RAM and a Nvidia GeForce GTX 980Ti graphics card. The display screen was 22 inches with a resolution of 1920x1080 pixels. In the *Desktop/Pointing* task, a standard mouse/keyboard setup was used for input (see Figure 5.15). The system had to be connected to the web at all times, because the application itself was a web page, which permanently sent logs to the a web server and was initially loaded from it. For displaying the WebVR content, a Google Chrome web browser and Microsoft Windows 10 was used. The *Desktop/Speech* task was performed with the same setup, but without the keyboard and mouse and instead using the built-in microphone of a Microsoft Kinect 2 for the speech input (see Figure 5.15) and the Web Speech API[48]. For both VR tasks LG Nexus 5X smartphone with a display size of 5.2 inches was used, which was put into Elegiant 3D VR glasses (see Figure 5.15). For using the system on the mobile device, only the Google Chrome App has to be installed, in combination with at least Android 6.0.

[48] `https://cloud.google.com/speech/`

Figure 5.15: Left: **Desktop setting.** Here, the participant controls the camera by moving the mouse with her/his dominant hand. The speech was recorded by a Microsoft Kinect 2, placed under the display. The keyboard shown in this picture had no effect and was not used in the experiment. Right: **HMD setting.** Here, the user controls the camera by moving her/his head. Speech was recorded by the smart-phone's microphone.

Design

The experiment was a within-subjects design, having two independent variables with two levels, respectively:

- Input Method (*Head Pointing*, *Speech Input*)
- Output Method (*Desktop*, *HMD*)

The input method conditions were counterbalanced using a Latin square. Aside from training, this amounted to 16 participants $\times$ 2 input methods $\times$ 2 output methods $\times$ 4 search terms $= 256$ trials. The four products were placed all over the virtual shop, and in order to ensure equal conditions for every participant, all trials started at the same position. Furthermore, the participants received only minimal instructions about the functionality of the different interaction types, so that no explicit conceptual model was assigned to them. The six dependent variables were as follows:

- *Performance* (Task Completion Time, Error Rate)
- *Preference* (Usability, UX, Motion Sickness, Immersion)

It is also important to mention, that every trial had to be performed within a time limit of 120 seconds (see Section 5.2.4); otherwise the trial was counted as failed. The average duration of the whole experiment per participant was ~ 60 minutes, including introduction and questionnaires.

Task

The first experiment consisted of four different tasks representing all combinations of two output (*Desktop* vs. *HMD*) and two input (*Head Pointing* vs. *Speech Input*) modes. Every participant had to perform each of these four tasks with the goal to search for four specific products within the virtual shopping environment one after another, select them, and put into the shopping cart provided for this purpose.

In *Head Pointing* mode, the participants performed a search task in order to find a certain product. As is standard practice, this product search task had two stages. The participant could only interact with waypoints or products at near range. However, after they stood in front of a shelf where they expected to find the desired product, the second stage of the search task began. Here, they searched for the product in the shelf's racks. When the participant had put the correct product in the shopping cart, the trial was completed successfully.

In *Speech Input* mode, the navigation part of the search task was omitted and replaced by a speech input task. Moreover, the participants then would not be able to move using the waypoints, so they were not displayed anymore. They started in front of a speech interface panel. After a product search was initiated by speech input, the search results were displayed. Since a search result has been selected, the participant was moved to the corresponding shelf along the shortest path.

In summary, each of the four tasks consisted of four trials, one per product. A trial was completed if the correct product was in the cart within the time limit of 120 seconds; otherwise the trial was marked as failed.

Procedure

At the beginning of the experiment, the participant was welcomed by the experimenter. After that, the participant was to perform all four tasks in Latin-square order, which lasts about 20-30 seconds in average for one of four trials per task, i.e. about 5-10 minutes for all four in- and output task combinations in average. Before each task, the task goal and the handling of the interaction technique were introduced to the participant within a short warm-up phase with a maximum of five minutes.

After the warm-up phase in the VE and after each trial, the experimenter reset all states in the VE, so that the participant could always start from the same position. When all four trials of a task were performed, the participant was asked to fill out post-task questionnaires collecting the subjective feedback: a System Usability Scale (SUS) [36], User Experience Questionnaire (UEQ) [157], Motion Sickness Assessment Questionnaire (MSAQ) [93] and (only for VR) Presence Questionnaire [278]. Those questionnaires took about 5 minutes after each of the four tasks. Furthermore, there was a 2-minute break after each set of post-task questionnaires and before the next task.

Finally, after all 16 trials were performed and all post-task questionnaires were filled out, the participant was asked to fill out a final questionnaire collecting demographic data. Overall, the experiment took about 60 minutes per participant in total.

5.2.6 Results

In the following we use abbreviations for the four tasks we tested: **D** for *Desktop*, **HMD** for *Head-mounted Display*, **1** for *Head Pointing*, and **2** for *Speech Input*. Put together, they are: **D1** (*Desktop/Pointing*), **D2** (*Desktop/Speech*), **HMD1** (*HMD/Pointing*), **HMD2** (*HMD/Speech*).

Task Performance

The task performance metrics include quantitative measurements such as speed and accuracy during the first experiment. These metrics indicate to what extent users are able to cope with the task, i.e. the interaction and output methods. They are computed per participant and condition as the average over the four trials per task.

The **task completion time** is the elapsed time for the user to complete a trial within a search task. In the *Head Pointing* mode, it is defined by the elapsed time between the first pointing interaction with a way-point and the moment the searched for product was added to the cart. In the *Speech Input* mode, the time measurement started with the first glance on the speech balloon symbol on the search interface panel, and ended when the searched for product was added to the cart.

The overall task completion time of all trials ($N = 64$) was on average $24.05s$ ($SD = 11.68$), with significant differences between the four tasks ($p < 0.01$, $F(1, 202) = 4.461$). **HMD2** was the fastest task with a mean time of $22.80s$ ($SD = 6.44$), followed by **D2** with $23.00s$ ($SD = 12.27$), **D1** $23.45s$ ($SD = 8.29$), and **HMD1** with $31.07s$ ($SD = 11.10$). When comparing the input mode, *Speech Input* was faster on average ($M = 22.90s$, $SD = 9.67$) than *Head Pointing* ($M = 25.83s$, $SD = 14.10$). But concerning the output modes, the **D1** and **D2** were faster, with a mean time of $22.85s$ ($SD = 13.45$), than the ones using *HMD* for output ($M = 25.46s$, $SD = 9.05$). Furthermore, an univariate ANOVA analysis were conducted with task, input mode and output mode as factors, and the elapsed time as dependent variable. A significance was found regarding the task ($p < 0.01$, $F(3, 202) = 4.461$), as well as a significant effect of completion time with regard to the input mode ($p < 0.05$, $F(1, 202) = 5.662$) and output ($p < 0.05$, $F(1, 202) = 5.990$). In addition, an interaction was found between the input and output modes ($p < 0.01$, $F(1, 202) = 6.610$) with regard to speed.

The **error rate** is defined by the ratio of the number of times that participants failed to find the desired products within a maximum period of 60 seconds. It is also worth to mention that if the speech recognition has failed, i.e. the spoken command wasn't detected correctly, the trial would have not been automatically counted as a fail. But, however, only in a very small minority of cases, which was not worth to mention due to our choice of product names after the pilot study (see Section 5.2.4), the time limit exceeded because of recognition fails instead of „not finding the product".

The overall error rate of all trials ($N = 64$) was on average 21% ($SD = 0.41$), with significant differences between the four tasks ($p < 0.01$, $F(1, 252) = 26.235$). **HMD2** had the lowest error rate of 2% ($SD = 0.13$). While **D2** had an error rate of only 8% ($SD = 0.27$), **D1** followed with 22% ($SD = 0.42$), and finally **HMD1** with a 50% failure rate ($SD = 0.50$). *Speech Input* had a notably lower error rate of 5% ($SD = 0.21$) against *Head Pointing* with

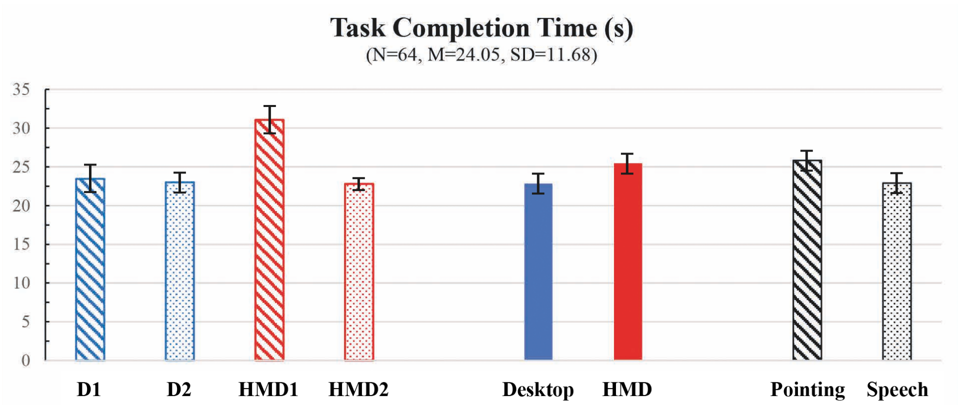

Figure 5.16: Results of the task completion time measurements. The values are given in seconds (s). D1 (Desktop/Pointing), D2 (Desktop/Speech), HMD1 (HMD/Pointing), HMD2 (HMD/Speech).

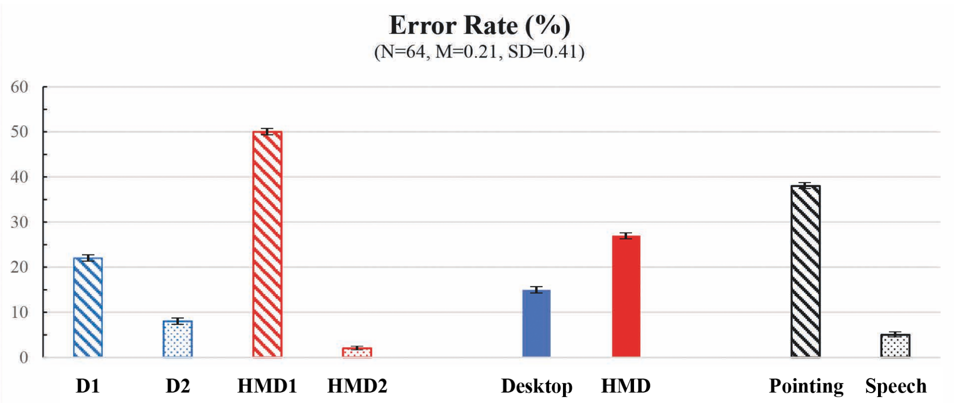

Figure 5.17: Results of the error rate measurements. The values are given in percent (%). D1 (Desktop/Pointing), D2 (Desktop/Speech), HMD1 (HMD/-Pointing), HMD2 (HMD/Speech).

38% ($SD = 0.49$). Considering output, *Desktop* had less failures ($M = 15\%$, $SD = 0.36$) than both *HMD* tasks ($M = 27\%$, $SD = 0.45$). An univariate ANOVA showed significant differences of error rate in general (not the recognition rate) regarding input ($p < 0.01$, $F(1, 252) = 53.480$), and output ($p < 0.01$, $F(1, 254) = 7.761$), and finally an interaction was found between input and output ($p < 0.01$, $F(1, 252) = 17.463$).

User Preferences

For user preferences metrics, we collected a variety of subjective feedback to assess *UX*, *Usability*, *Motion Sickness* and *Immersion*, important in order to ensure comparability with the VRUX (see Section 3.4)

We chose the User Experience Questionnaire (UEQ) [157] as an end-user questionnaire to measure **UX** quickly in a simple and immediate way. Overall, the UX was rated at 1.15 ($SD = 0.85$) on average on a scale between -3 to 3. **HMD2** was rated with a value of 1.43 ($SD = 0.83$) on average, followed by **D2** with 1.16 ($SD = 0.62$) on average. Nevertheless, **D1** had an average of 1.05 ($SD = 0.97$) and **HMD1** of 0.96 ($SD = 0.96$). With regard to the input methods, *Speech Input* was rated with 1.29 on average ($SD = 0.73$) and *Head Pointing* with 1.00 ($SD = 0.95$), but also without significant differences ($p = 0.19$, $F(1, 60) = 1.79$, $\eta^2 = 0.03$). Finally, concerning the output, *HMD* was rated with an average of 1.19 ($SD = 0.92$) and *Desktop*

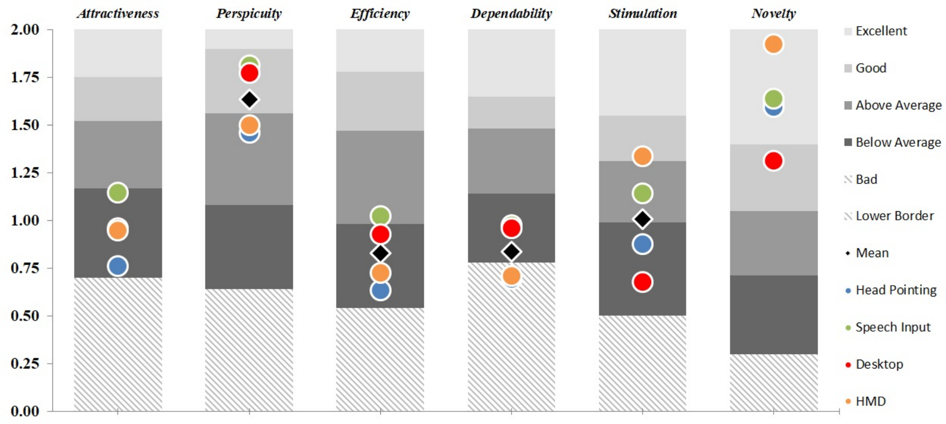

Figure 5.18: Results of the User Experience Questionnaire (UEQ)

with an average of 1.10 ($SD = 0.80$). However, there were no significant differences between the tasks ($p = 0.45$, $F(1, 60) = 0.90$, $\eta^2 = 0.04$) nor the output modes ($p = 0.68$, $F(1, 60) = 0.17$, $\eta^2 = 0.01$). However, the data was subjected to a factor analysis, which resulted in the construction of a 26-item questionnaire including six factors, see Figure 5.18. A multivariate ANOVA with all six factors as dependent variables showed a significance effect for stimulation with regard to the single tasks ($p < 0.05$, $F(3, 60) = 2.61$, $\eta^2 = 0.12$). Regarding the input, we found no significant differences, but an effect between the outputs for stimulation ($p < 0.02$, $F(3, 60) = 6.61$, $\eta^2 = 0.10$) and novelty ($p < 0.03$, $F(3, 60) = 5.59$, $\eta^2 = 0.09$).

The System Usability Scale (SUS) [36] is likely the most popular questionnaire for measuring attitudes toward system **usability**. It is a reliable and valid measure of perceived usability. Furthermore, it performs as well as or better than commercial and homegrown internal questionnaires.

The overall SUS score including all four tasks was on average 72.19 ($SD = 15.92$). **HMD2** had the best SUS score in average ($M = 77.50$, $SD = 13.21$), followed **D1** ($M = 75.00$, $SD = 17.18$), **D2** ($M = 70.94$, $SD = 12.84$) and **HMD1** ($M = 65.31$, $SD = 17.39$). Regarding the input modes, *Speech Input* was rated higher (**D2** and **HMD2**; $M = 74.22$, $SD = 13.21$) than *Head Pointing* (**D1** and **HMD1**; $M = 70.16$, $SD = 17.89$). Here, a univariate ANOVA pointed out a significance regarding each task ($p < 0.01$, $F(3, 60) = 7.82$), and a significant effect with regard to input mode ($p < 0.05$,

$F(1, 252) = 4.56$). The *Desktop* tasks (**D1** and **D2**) had with regard to output mode an average score of 73 ($M = 72.97$, $SD = 15.25$) and the *HMD* tasks had an average score of 71 (**HMD1** and **HMD2**; $M = 71.41$, $SD = 16.41$), without significances between them ($p = 0.43$, $F(1, 60) = 0.623$).

For measuring **motion sickness**, we asked the participants to fill out the Motion Sickness Assessment Questionnaire (MSAQ) [93], which is a valid instrument for the assessment of motion sickness.

The experimental results show a total score of 5% on average ($M = 0.05$, $SD = 0.08$), with **HMD1** as the highest motion sickness rating ($M = 0.09$, $SD = 0.10$), followed by **HMD2** ($M = 0.08$, $SD = 0.08$), **D2** ($M = 0.03$, $SD = 0.04$) and **D1** with the lowest score ($M = 0.02$, $SD = 0.04$). *Head Pointing* had an average score of 6% ($M = 0.06$, $SD = 0.08$) compared to *Speech Input* with an average of 5% ($M = 0.05$, $SD = 0.07$), but without significant difference between them ($p < 0.01$, $F(1, 60) = 61.23$). Regarding the output, *Desktop* received a higher score of 2% ($M = 0.02$, $SD = 0.04$) than *HMD* with a lower average of 8% ($M = 0.08$, $SD = 0.09$), with significant differences ($p < 0.01$, $F(1, 60) = 45.16$).

A Multivariate ANOVA with all four MSAQ factors (G: gastrointestinal, C: central, P: peripheral, S: sopite-related) as dependent variables and single task as factor was conducted. It showed, that there are significant differences regarding the single tasks (G: $p < 0.01$, $F(3, 60) = 22.79$; C: $p < 0.01$, $F(3, 60) = 21.15$; P: $p < 0.03$, $F(3, 60) = 3.29$; S: $p < 0.01$, $F(3, 60) = 18.93$).

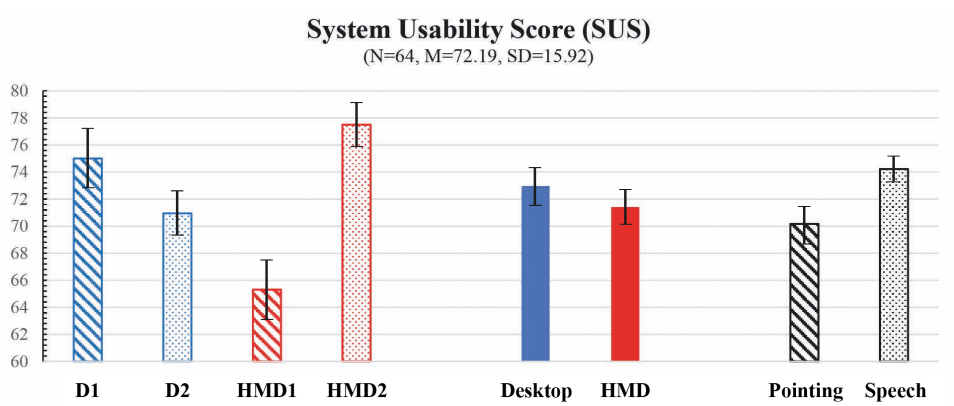

Figure 5.19: System Usability Score (SUS). D1 (Desktop/Pointing), D2 (Desktop/Speech), V1 (VR/Pointing), V2 (VR/Speech).

Another Multivariate ANOVA with all four MSAQ factors as dependent variables and input and output as factors was conducted. There were significant differences regarding the output (G: $p < 0.01$, $F(1, 60) = 67.10$; C: $p < 0.01$, $F(1, 60) = 61.23$; P: $p < 0.01$, $F(1, 60) = 6.81$; S: $p < 0.01$, $F(1, 60) = 54.93$). But apart from that, there were no significances found regarding the input (G: $p < 0.33$, $F(1, 60) = 0.97$; C: $p < 0.27$, $F(1, 60) = 1.25$; P: $p < 0.10$, $F(1, 60) = 2.85$; S: $p < 0.18$, $F(1, 60) = 1.86$). There was also no interaction found between the input and output types (G: $p < 0.59$, $F(1, 60) = 0.30$; C: $p < 0.33$, $F(1, 60) = 0.98$; P: $p < 0.65$, $F(1, 60) = 0.21$; S: $p < 0.93$, $F(1, 60) = 0.01$).

In order to measure the **immersion** and presence of the VE, we used the common immersion questionnaire for VR applications from Slater et al. [278]. There are no results provided for **D1** and **D2**, because the SUS (Slater Usoh Steed) questionnaire is designed primarily for VR applications. So it was only filled out by the participants after the VR tasks. Moreover, immersion measures are not the forefront of desktop applications and this part of the evaluation was more a plausibility test for our proof of concept including the VRUX. So these results will serve as a basis for future studies exploring interaction in VR shopping environments.

However, the 'SUS Count' shows the mean of the SUS count of '6' or '7' scores amongst the 6 questions. Here, the SUS count for **HMD2** (*Speech Input*) is slightly higher (1.13) than for **HMD1** (1.00). The 'SUS Mean' uses the mean score across the 6 questions instead. Both *HMD* tasks have nearly equal SUS means, whereas **HMD2** (*Speech Input*) has an average of 2.76 ($SD = 0.91$), and **HMD1** (*Head Pointing*) 2.76 ($SD = 0.86$), which was slightly variable. Here, there was a significant effect concerning the input methods (($p < 0.05$, $F(1, 60) = 4.271$). The overall immersion score of the VR shopping environment was 2.78 ($SD = 0.88$).

5.2.7 Discussion

In the following, the results will be discussed with a detailed consideration of the objective performance and subjective preferences results with respect to our VRUX model (see Section 3.4).

Task Performance

The analysis of the speed results demonstrate that searching for a product in the VE by *Speech Input* performs faster than using only the user's *Head Pointing*, which partly confirms H_3. So it seems to be easier for the users to simply pronounce the product name and let the system search for its location. In *Speech* mode, the system provides a list of search results in realtime and the user only needs to choose one of these, instead of scanning the whole market for the desired product. Although there was an exploration phase before each task, searching only via *Head Pointing* faces an additional handicap of spatial knowledge of the environment.

Concerning the two outputs, *Desktop* was more efficient than *HMD* with regard to speed and error rate (H_1). On the one hand, this can be explained by the different means of pointing control (mouse vs. head) and the graphical resolution due to the technical setups. Then again, it is to be expected that *Desktop* tasks are faster on average than using *HMD*, because despite its more natural use of *Head Pointing*, the participants were used to control the camera by mouse in 3D scenes. But in practice, this is contrary to usability and UX ratings. The bad performance of *Head Pointing* using *HMD* could be also explained by the technical limitations of using WebVR on a smart-phone browser and the still error-prone sensor input of mobile devices. Maybe using a more sophisticated HMD and hardware would provide different results. However, this experiment focused on commodity, affordable and state-of-the-art VR hardware.

User Preferences

Based on research, a SUS score above 68 would be considered above, and any **usability** rating below 68 is below average [258, 93]. Due to this rating, the overall SUS Score of our VR shop application in the experiment is with 72.2 slightly above the average of 68. Therefore, we can state that the participants understood the content, and it helped them accomplish their task. The measured SUS score also illustrates that the system adequately communicates what users are required to do with the application, namely find products in the store.

Since differences in **UX** scores for each different tasks, as well as for the input and output modes, hardly exist, all scores are still moderate on average. The hardware is still a bottleneck, so special VR smartphones (e.g. Google Pixel) with a browser specifically developed for WebVR purposes could solve most of the user issues. Because there was also no significance found between the output mode regarding UX and usability, H_2 could only be proved partly. On closer inspection, *HMD* outperforms the alternative especially in terms of novelty and stimulation, two of the six UEQ factors [157]. This underlines once again that our approach of a hands-free VR shopping can be considered to have higher usefulness than a common 3D desktop variant. Whereas *Head Pointing* is the commonly used input method for mobile VR applications, our results show that *Speech Input* is preferred by the user regarding usability and UX (H_3). It is worth noting that the results are highly dependent on the speech interface and its recognition, as well as the data quality of the product names to be searched. Thus, *HMD* with *Speech Input* reached the highest usability and UX scores, which can explained by its combination of in- and output (H_4).

The results of the MSAQ indicated that using *HMD* causes more **motion sickness** than *Desktop*, which is emphasized by the found significances. Although this can been seen as obvious, because 3D desktop applications should not cause motion sickness anyway, we have considered it to be very important to use *Desktop* as a baseline especially for motion sickness. However, with regard to the input method, *Head Pointing* causes more motion sickness than *Speech Input*, which can be explained by the worse physical demand due to the higher amount and frequency of head movements. Another indicator for higher motion sickness can be the longer time the participants stayed in the environment using *Head Pointing*. However, the overall high ratings are largely due to WebVR and its technical limitations, like resolution, rendering performance of the smart-phone and wearing comfort of the HMD.

Finally, the results of the **immersion** questionnaire showed that the intensity of immersion is not affected by the input method, because both analyzed tasks had the same value. Overall, the values are located around the middle of the scale, so we could assume that the users had a good feeling of presence when using the *HMD* output. The explanation for that is twofold. On the one hand, the render quality and computational power of the mobile device is worse than on a desktop monitor. Low-res textures, restricted lights, and shadows can reduce immersion, which could be solved by using more powerful devices. On the other hand, the inaccuracy of the mobile device sensors can cause a temporary latency of the head movement, which could be avoided by better-calibrated sensors.

5.2.8 Conclusion

Current online shops may be functional and efficient, but do not provide enough of an immersive shopping experience. So this experiment focused on developing an immersive virtual shopping environment that includes and combines the major advantages of online and physical stores and tries to tear down limitations of e-commerce and VR. In order to provide a realistic setting, we evaluated hands-free interaction techniques using smart-phone VR in a comparative user study. Because the smartphone itself is already available for most users, it is the cheapest and most affordable scenario for VR online shops. Our system enables the user to search for products by head pointing only or in combination with speech input in a VE.

Finally, it could be confirmed that the overall usability of the system is slightly above average, but not significantly; the *HMD* setup with *Speech Input* was especially preferred. Regarding efficiency, *Speech Input* was more efficient than using *Head Pointing* only, but it should be mentioned that proper data quality is essential to provide optimal recognition. Our concept includes availability and sustainability instead of only adaptability [48].

Most of the existing online shops and other virtual applications try to simulate conventional stores without addressing the limitations of those shops. Therefore, online shopping needs to be reconceptualized and designed from a different perspective. Most shopping trips already start at home when customers create a shopping list. This can be done by physically or mentally inspecting their supplies at home. Therefore, we conducted another experiment and propose an apartment as a shopping environment, where products are located where an average buyer would expect them.

5.3 Experiment 2: Spatial Menu Representation and Apartment-based Categorization for Online Shops

This work builds on the results of the ProductFinder [290] and our mobile VR shop prototype (see Section 5.2) and aims to explore, design and implement better and intuitive categorization schemes and menu representations for online shops, that enrich and improve the shopping experience. We utilize the Apartment metaphor, in which products are categorized into rooms and furniture representing departments and shelves. Furthermore, we developed a realistic and interactive map-based spatial menu representation based on prior research findings. In a comparative user study, we evaluated our new menu categorization and representation in comparison with the current standard in online shops, based on real data from a local retailer [290]. The results show that the Apartment metaphor in combination with a spatial representation outperforms all other conditions with regard to all tested variables of task performance (success rate, task completion time) and user preferences (UX, usability, workload).

5.3.1 Introduction

Nowadays, online shops are a well established and indispensable part of our everyday life. They often offer a better product availability, time savings and higher comfort compared to physical stores since purchases can be made from the comfort of ones home [282]. These factors have lead to an enormous growth that is anticipated to continue [208, 269]. While in 2017, only 10.1% of worldwide purchases were made online this is expected to grow to 15.5% by 2021[49].

While the design and interaction of online shops have changed and improved over the last years, due to a higher focus on UX and usability [121, 166, 327], they are still mainly focused on product presentation and purchase transactions. Especially the search bar functionality is used to improve efficiency of customer searches has received most attention [235], while product search and explorations using a menu interface has been

[49] https://tinyurl.com/y8akourk

largely neglected. Contrary to common belief, product searches via a search bar has been proven to be not necessarily preferred by the user nor is it generally more effective [156]. Especially when an online shopper does not know the explicit product name or is simply exploring the store, e.g. for a gift, the search bars might not be a suitable solution.

Although the benefit of visualizations used in online shops is well known, menu representations are still mainly text-based. Furthermore, the underlying categorization is often inconsistent throughout different online shops, which makes it difficult to understand the underlying classification [239]. Therefore, we explored the usage of the Apartment metaphor [3] as a menu representation for online shops. It exploits the users familiarity with an apartment environment and categorizes products based on their association to rooms and furniture to make them easier and faster to explore for the user. We present the development of the categorization of the different products as well as spatial visualization showing the floor plan of an apartment or a store to explore the different categories. Our comparative evaluation shows that an Apartment metaphor based online shop menu outperforms classic linear store-based menu structures in task performance, usability and UX. Hence, this work contributes to the further development of online shop menus by:

- **A study of an online shop prototype** was conducted and evaluated four combinations of menu representation and categorization based on related work with respect to task performance and user preference.

- **User insights and actionable improvements** were provided for designing and developing future shopping environments.

5.3.2 Concept

In our approach, we wanted to explore the usage of the Apartment metaphor for categorization in combination with a spatial representation as a new menu type for online shops. Therefore, we re-categorized products from a set of previous categories (departments and shelf names), which are more likely to be found in a physical market, to residential categories as part

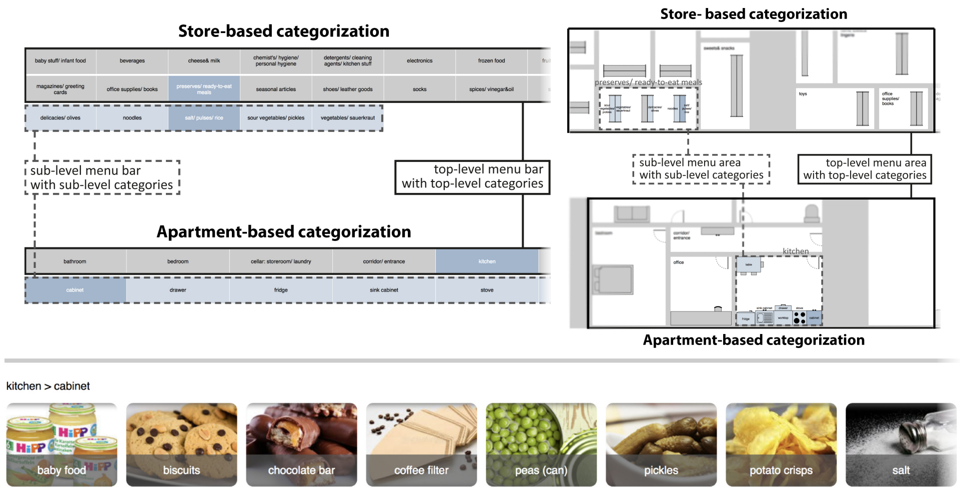

Figure 5.20: Linear (left) and Spatial (right) representations, and product area relating to the sub-level "cabinet" under the top-level "kitchen".

of a pilot study. Based on these categories, we developed a spatial menu representation, as an alternative to the traditional linear menu.

To test and compare these four menu types and combinations of both representations and categorizations, they were integrated into an online shop prototype. As the basis for our menu, we selected a set of 36 products, that represent the core areas of online trading. These products were selected from a local hypermarket with an associated online shop and is based on its frequently searched products [290] from various traditional product areas such as food, clothing, electronics and others.

Representations

Representations of menus vary in the arrangement of the items on the screen. This can affect the search time to find a desired menu item visually, as well as the time to point and click [6, 47, 309]. We examined two different representations here: *Linear* and *Spatial*. The depth and width of both menus depend on the underlying categorization (see Figure 5.20).

The *Linear* menu is most common in today's web interfaces such as online shops [341]. Here, the menu items are arranged to form either a horizontal or a vertical line, usually with text labels. Since horizontal linear menus are recommended at the top of the screen [341], this combination is used as a reference menu representation.

Our *Spatial* menu representation is influenced by earlier findings on the grid arrangement [6, 259] and floor maps [200], which are often used in shopping malls, e.g. as printed maps or orientation points. We use a spatial representation – a map –, based on real environments, where menu items are arranged according to the position they occupy in the real world either inside a store or an apartment depending on the used categorization (see Figure 5.20). Thereby we hope to exploit the users spatial memory and improve the performance of the menus [104, 253].

Categorizations

The categorization determines the semantic structure of the menu and is usually structured hierarchically. In this paper, the categorizations are based on a three-level hierarchy (or menu depth) with top, sub and product levels. Overall, two different categories are investigated in this work: a traditional *Store*- and an *Apartment*-based categorization.

The traditional *Store-based* categorization with different departments and themes, serves as the reference point for the evaluation as it represents the de-facto standard in current online shops [341]. The top-level categories of the hierarchy correspond to the product range that is typically separate departments, such as "milk & cheese" or "beverages". The subordinate categories describe shelves in these delimited market areas, e.g. "lemonades", followed by the product level as the lowest (see Figure 5.20).

The *Apartment* metaphor used in our approach uses the fact that users are familiar with the structure of an apartment and the items (products) placed in it based on everyday habits and experiences [3]. As with the store-based categorization, this is also based on a three-level hierarchy: rooms (e.g. kitchen), furniture (e.g. refrigerator), and product (e.g. mustard). In order to develop these categories and define the corresponding product assignments a pilot study was carried out to find out where users assume that these products should be located inside an apartment.

5.3.3 Pilot Study

To create the needed reclassification of the selected products from store- to apartment-based categorization, we conducted a pilot study, which was divided into two phases: categorization and product allocation inside the Apartment metaphor, and creating product groups to serve as the logical basis for the experimental design.

Phase 1: Categorization and Product Allocation

In total, 42 participants (20 female) between 18 and 58 years ($M = 31.12, SD = 12.35$) volunteered in the pilot study, which consisted of two sessions. In the first qualitative session, we conducted a semi-structured on-site interview and asked the participants where, i.e. in which room on or in which furniture they store or would store each of the selected products, followed by a demographic questionnaire. Most participants lived in a two-room apartment ($M = 2.17, SD = 0.85$) (excluding bathroom, kitchen or hallway) with an average of two inhabitants ($M = 2.38, SD = 1.19$). All of them have already shopped online, most of them at least once a month ($N = 34$).

The result of the first session was a preliminary set of locations (rooms and associated furniture), which form an essential part of the Apartment categorization, as they form the basis for the product allocation. A total of seven rooms were considered, named by over half of the participants: bathroom, kitchen and bedroom ($N = 41$), hallway ($N = 37$), living room ($N = 35$), pantry/cellar ($N = 32$) and office ($N = 30$). After defining the rooms used for product allocation, the furniture within these rooms had to be specified. The data collected by the interview was qualitative, so characteristic keywords were chosen to organize the answers given. Furniture that is very similar in use has been combined, e.g. "built-in cupboard" ($N = 1$), "linen cupboard" ($N = 2$), "chest of drawers" ($N = 2$) and "cupboard" ($N = 11$) were combined under "cupboard". In the second session, after we conducted and analyzed all 42 interviews, the same participants were asked to map the 36 products to room-furniture pairs using an interactive web application (see Figure 5.21).

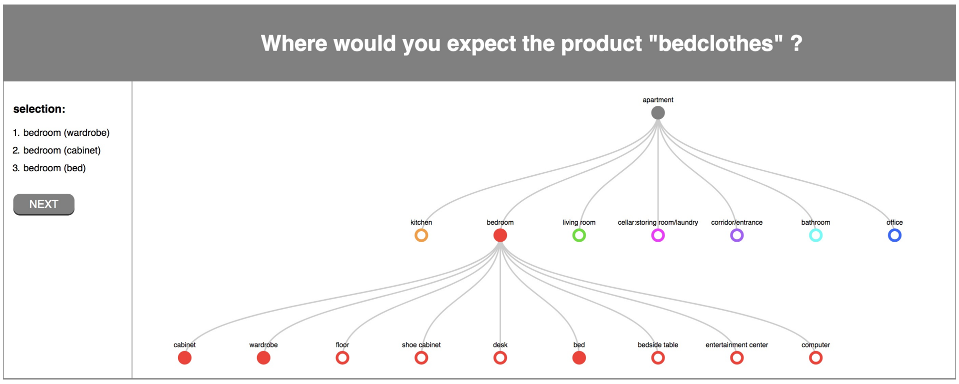

Figure 5.21: The tree representing the hierarchy of the apartment categories with room nodes on the first level and furniture nodes on the second level.

Phase 2: Product Groups

After selecting a suitable product set in the conceptual process, each product was assigned a level of difficulty to ensure better comparability. This was necessary to control learning effects within a categorization during the main study. This second phase of the pilot study took place in the local hypermarket, which provides us with the selected product data set, floor plan, as well as the departments and shelf names. Each participant ($N = 30$, 14 female) was provided with a worksheet, where the products had to be assigned to the store-based top categories. For each product, all answers were identified and statistical values were calculated for the error rates, i.e. the percentages of incorrect product searches.

The error rate over all products was 21.67% ($SD = 27.20$) on average, which shows that there is considerable potential for improvement, at least for this particular retailer, since expectations often do not correspond to reality. The average statistical values resulting from the short classification led to two comparable product groups being formed in the following experiment to eliminate learning effects in relation to the different categories. Therefore, all 36 products were divided into two groups of 18 each with comparable error rates (21.1% vs. 22.2%).

5.3.4 User Evaluation

We conducted this experiment to compare the developed apartment metaphor for online shopping with more common menu representations in respect to task performance, user preferences, and unmet needs. We evaluated two menu representations (*Linear* vs. *Spatial*) and categorizations (*Store* vs. *Apartment*). Our main hypotheses were defined as:

H_{1-1} The task can be performed more efficiently using *Apartment* .

H_{1-2} *Apartment* is preferred over *Store* categorization.

H_{2-1} The task can be performed more efficiently using *Spatial* .

H_{2-2} *Spatial* is preferred over *Linear* representation.

Participants

For the second experiment 24 different unpaid participants (12 female) were recruited from the university's campus; they were aged between 20 and 33 years ($M = 25.3$, $SD = 3.6$). Most of the participants live in a two room apartment ($Median = 2, M = 2.04, SD = 0.86$), which do not include bathroom, kitchen or hallway, with two inhabitants ($Median = 2$). Seven participants live with their parents/family (29.17%), six in a partnership or shared apartment (25%) and five live alone (20.83%). On a 7-point scale from never to daily with regard to online shopping frequency, most participants regularly purchase online, i.e. 62.5% at least several times per month, and all participants shop online at least once per month with computer ($N = 24$), compared to tablets ($N = 7$) or smart-phones ($N = 14$).

Apparatus

The experiment was conducted on a MacBook Pro connected to a 24-inch monitor. A standard wireless mouse was used as input device with medium speed settings (see Figure 5.22). The software was displayed in Google Chrome. HTML, CSS and JavaScript were used to implement the different menu interfaces in the prototype. Additionally, the JavaScript D3 library was used for data visualization purposes in the spatial menu condition. A database was set up using XAMPP, data exchange was realized using PHP.

Design

The experiment was a within-subjects design, with two independent variables with two levels each (representation: linear and spatial; categorization: store- and apartment-based) and five dependent variables related to the performance (task completion time, success rate) and preference of users (UX, usability, workload). All conditions were counterbalanced using a Latin square. In order to eliminate learning effects concerning the different categorizations, the total set of 36 selected products was split into two comparable and equally difficult groups with 18 products respectively. Aside from training, this amounted to: 24 participants $\times$ 2 representations $\times$ 2 categorizations $\times$ 18 product searches $= 1728$ trials.

Task

During the main study in this second experiment, each participant performed a series of 18 search trials using a combination of the two representations (linear and spatial) and categorizations (store and apartment). The goal was to find and select a specific product and confirm the selection. The top-level categories were the departments of a local store for store-based and rooms for apartment-based categorization. The sub-level represented the shelves for store-based and furniture for apartment-based categorization (as developed in our pilot study). While the top- and sub-level visualizations differed depending on categorization and representation, the product level

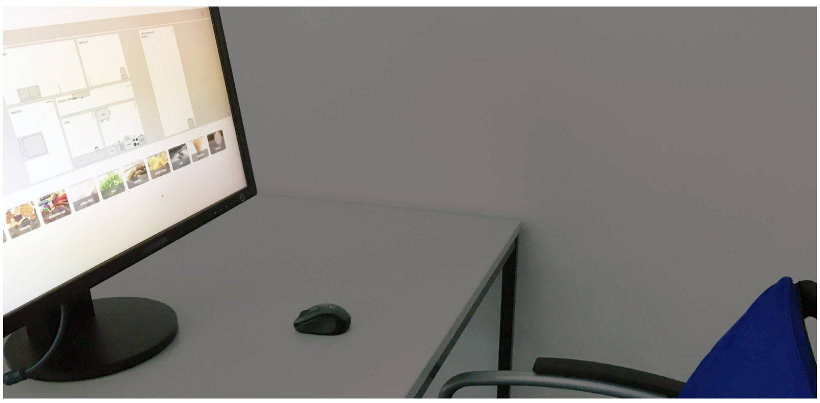

Figure 5.22: Experimental Setup.

was displayed similar over all conditions by a product image and text label. A trial was successfully completed when the correct target product was selected and confirmed within a time limit of 30 seconds.

Procedure

After welcoming by the experimenter, the participant was introduced by an informed consent form. Each participant used all four menu types in Latin square order to search for products in the prototypical online shop. Before using a particular type of menu, the participant was introduced to the tested condition by watching a demonstration video that showed an example search task step by step. Then a set of 18 search trials was carried out in random order. Before each trial, the name and image of the target product appeared for five seconds. Then the product had to be found in the three-level menu and selected by clicking on it and confirming the selection. After each trial set per menu type, three post-task questionnaires (UEQ [157], SUS [36], NASA TLX [107]) were filled out by the participant to collect user preference ratings. The entire process was then repeated for the other three menu types. Afterwards, a final post-study questionnaire was answered, which included demographic questions.

5.3.5 Results

Throughout this results section and in the following discussion we use abbreviations, fill patterns and color indications: **Linear (striped), Spatial (solid), Store-based (orange), Apartment-based (blue)**.

Task Performance

The task performance is measured quantitative through task completion time and success rate. These metrics indicate to what extent users are able to cope with the menu interfaces. They are computed per participant and condition as the average over the 18 trials.

The **success rate** describes the ratio between the number of successful and the total number of product searches. A product search is considered successful if the correct product has been selected and confirmed within

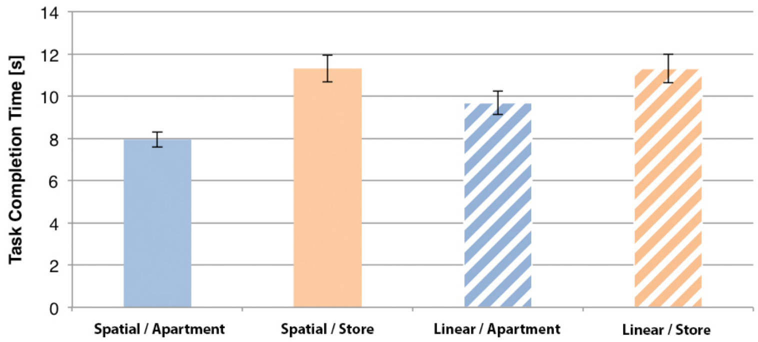

Figure 5.23: Speed measurements (seconds) of successful trials.

the maximum execution time of 30 seconds. *Spatial/Apartment* achieved the highest average success rate ($M = 98.61, SD = 11.72$), and *Linear/Store* the lowest ($M = 69.44, SD = 46.12$). An univariate ANOVA showed significant differences with regard to success rate between all four conditions ($F_{3,1724} = 60.71, p < 0.01, \eta^2 = 0.10$). Furthermore, there were significant differences regarding success rate between the representations ($F_{1,1724} = 34.96, p < 0.01, \eta^2 = 0.02$) with *Spatial* ($M = 90.05, SD = 29.96$) better than *Linear* ($M = 80.44, SD = 39.69$), as well as between the categorizations ($F_{1,1724} = 144.94, p < 0.01, \eta^2 = 0.08$) with *Apartment* ($M = 95.02, SD = 21.76$) better than *Store* categorization ($M = 75.46, SD = 43.06$).

Task completion time was measured as the elapsed time in seconds to complete a single product search. The timer started when the countdown reaches zero and stopped automatically when the correct product has been selected and confirmed. In this analysis, we only included successful product searches. We found significant differences for task completion time between the single conditions ($F_{3,1469} = 33.96, p < 0.01, \eta^2 = 0.07$) with *Spatial/Apartment* as the fastest with 7.96s ($SD = 3.80$) on average, and both *Linear/Store* ($M = 11.31, SD = 5.89$) and *Spatial/Store* ($M = 11.31, SD = 6.17$) the lowest. Furthermore, an univariate ANOVA showed significant differences for speed between representation ($F_{1,1469} = 9.35, p < 0.01, \eta^2 = 0.01$) with *Spatial* ($M = 9.47, SD = 5.28$) faster than *Linear* ($M = 10.38, SD = 5.77$), and categorization ($F_{1,1469} = 78.12, p < 0.01, \eta^2 = 0.05$) with *Apartment* ($M = 8.79, SD = 4.82$) faster than *Store* ($M = 11.31, SD = 6.04$).

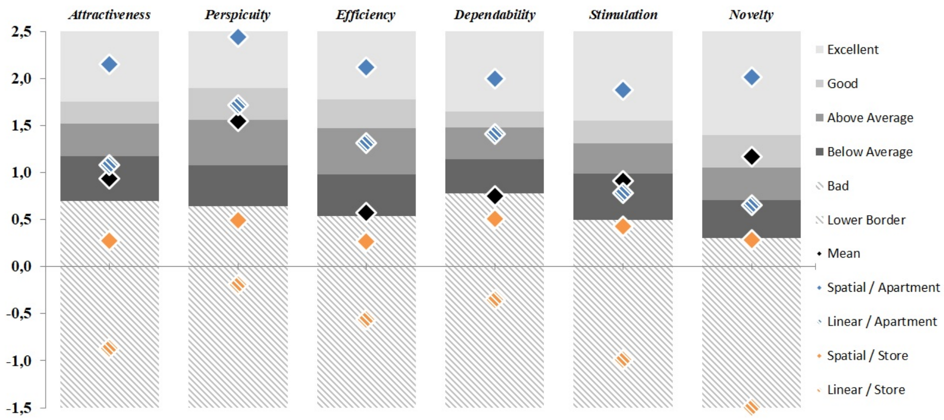

Figure 5.24: UEQ results with respect to comparison benchmarks (see shaded boxes). To make it easier to read, this figure shows a detail part between -1.5 and 2.5, while the original ranges between -3 and 3.

User Preferences

We chose the UEQ [157] as an end-user questionnaire to measure **UX** in a quick and straightforward way. On a scale between -3 and 3 the overall UX, *Spatial/Apartment* achieved the highest score of 2.10 ($SD = 0.53$) on average, and *Linear/Store* the lowest ($M = -0.76, SD = 1.15$), with significant differences between all four conditions ($F_{3,92} = 30.61, p < 0.01, \eta^2 = 0.50$). Furthermore, representations ($F_{1,92} = 22.46, p < 0.01, \eta^2 = 0.20$) and categorizations ($F_{1,92} = 69.17, p < 0.01, \eta^2 = 0.43$) also differed significantly regarding the overall UX score. *Spatial* was rated higher with an average of 1.24 ($SD = 1.30$) than *Linear* ($M = 0.20, SD = 1.50$) with respect to representation, whereas *Apartment* was rated higher ($M = 1.63, SD = 1.02$) than *Store* ($M = -0.19, SD = 1.33$) with respect to categorization. However, the data was also subjected to a factor analysis, including the six UEQ factors. *Spatial/Apartment* outperformed all other menu interfaces across the UEQ subscales, even ‚excellent' in terms of all subscales with respect to the UEQ benchmarks [261], followed by the *Linear/Apartment* , *Spatial/Store* , and finally *Linear/Store* (see Figure 5.24). The *Apartment* was rated higher on average than *Store* , as well as *Spatial* over *Linear* .

The SUS [36] is one of the most popular questionnaire for measuring attitudes toward system **usability**. It is a reliable and valid measure of perceived usability. *Spatial/Apartment* had the best score with 89.17 ($SD = 8.16$) on average, and *Linear/Store* ($M = 50.73, SD = 26.66$) the worst. A univariate ANOVA pointed out a significance regarding each condition ($F_{3,92} = 20.62, p < 0.01, \eta^2 = 0.40$), between the representations ($F_{1,92} = 55.34, p < 0.01, \eta^2 = 0.38$), and a significant effect between the categorizations ($F_{1,92} = 6.37, p < 0.05, \eta^2 = 0.07$). Comparing the categorizations, *Apartment* had a higher average usability score ($M = 85.05, SD = 13.95$) than *Store* (M=56.35, SD=23.54). With regard to the representations, *Spatial* ($M = 75.57, SD = 19.89$) was rated higher than *Linear* ($M = 65.83, SD = 26.93$).

The **task workload** of the tested menu representations and categorizations was assessed with NASA TLX [107]. On average, *Spatial/Apartment* was rated the best ($M = 22.10, SD = 10.78$) and *Linear/Store* ($M = 61.04, SD = 18.69$) the worst. An univariate ANOVA showed significant differences between the four menu types ($F_{3,92} = 22.22, p < 0.01, \eta^2 = 0.42$), the representations ($F_{1,92} = 7.58, p < 0.01, \eta^2 = 0.08$) and categorizations ($F_{1,92} = 58.93, p < 0.01, \eta^2 = 0.39$). *Spatial* was rated lower ($M = 35.68, SD = 22.06$) than *Linear* ($M = 45.96, SD = 24.27$), whereas *Apartment* achieved lower scores ($M = 26.49, SD = 16.19$) than *Store* ($M = 55.15, SD = 21.17$). We conducted a multivariate ANOVA with regard to these factors and found significant differences for all factors except physical demand between the four conditions, only for temporal demand, effort and frustration between the representations, and for all factors between the categorizations:

5.3.6 Discussion

In the following, we discuss the results (task performance and user preference) as well as the participants' feedback and comments. Furthermore we also discuss the limitations of this study.

Task Performance

The average success rate of about 99% of the spatial apartment-based menu is significantly higher than all other tested conditions, and contrasts with the linear store-based menu with the lowest rate of 69%. The speed results are based on successful product searches only and show that the task was executed faster with spatial apartment-based menus than with all other menus. This suggests that its intuitive categorization and spatial representation help the user to better understand the underlying information space. In addition, the visual cues in the spatial menus actually seem to facilitate the visual search process.

The clear differences in task performance indicate that a spatial grid-based menu in conjunction with an apartment-based categorization was more efficient than all other tested combinations, which proves H_{1-1} and H_{2-1}. Since the menu with the worst average task performance is the commonly used menu type in today's online shops (see Section 5.1.1), our results show a remarkable potential for improvement.

User Preferences

The overall UX results and highest ratings in all six UX subscales show that there is a clear advantage of the spatial apartment-based menu over all other tested menus. In particular, the significantly higher ratings of "Perspicuity", "Efficiency" and "Dependability" speak for more understanding, user-friendliness and reliability. Here, too, the visual hints of the spatial representation seem to facilitate the search process. High ratings in "Attractiveness", "Stimulation" and "Novelty" indicate that the more realistic and vivid presentation of the apartment categories seems to lead to a new and appealing experience.

Similar trends can be observed in the usability results. Here, too, the two apartment-based menus have achieved significantly better results. In addition, spatial menus achieved significantly higher usability values than linear menus within the respective categorization. Since values around 68 can already be interpreted as average to moderate[50], the two apartment-based interfaces with an average score of 85.05 can be described as ‚excellent'. Whereas the spatial apartment-based menu even has a value of 89.17, which shows that comprehensibility is further supported by the illustrative character of the spatial representation. Overall, the results show that spatial

[50] www.measuringu.com/sus

NASA TLX subscale	Menu Type	Representation	Categorization
Mental Demand	(17.49, 0.01, 0.36)	-	(49.02, 0.01, 0.35)
Physical Demand	-	-	*(4.70, 0.05, 0.05)*
Temporal Demand	(17.34, 0.01, 0.36)	*(4.41, 0.05, 0.05)*	(47.51, 0.01, 0.34)
Performance	(11.01, 0.01, 0.26)	-	(29.01, 0.01, 0.24)
Effort	(15.13, 0.01, 0.33)	(4.73, 0.05, 0.05)	(40.65, 0.01, 0.31)
Frustration	(20.11, 0.01, 0.40)	(8.66, 0.01, 0.09)	(50.96, 0.01, 0.36)

Table 5.2: The values refer to this format: $(F(x, 92) = .. , p < .. , \eta^2 = ..)$, with x=3 for the menu types and x=1 for representations and categorizations.

apartment-based menus were more usable than the other tested menus.

The task workload results show that store-based menus have scored more than twice as many points (55.15) as the apartment categorization (26.49). This indicates that the classification by rooms and furniture is cognitively less demanding than by product worlds and shelves. The spatial menus also achieved significantly lower utilization than the linear menus. For the individual subscales of the NASA TLX the spatial menu leads to significantly less effort, frustration and mental demand, which indicates that the visual cues facilitate and accelerate the orientation process. The apartment categorization additionally minimizes the mental demand, since no complex and strenuous considerations were necessary. Overall, the new categorization and representation is less demanding and frustrating.

In summary, taking into account the results of user preference, it can be stated that there is a significantly higher preference with regard to UX, task workload and usability for the spatial apartment-based menu interface. Thus the hypotheses H_{1-2} and H_{2-2} can be accepted, since they are fulfilled in all aspects considered.

Observations and Comments

The participants' comments also confirm the overall impression of the previously discussed results. In the post-study questionnaire, the participants were explicitly asked for their opinion of the tested menus. Here, 23 out of 24 participants preferred the spatial apartment-based menu, only one the linear apartment-based menu. This choice was based on terms such as "intuitive", "easy", "entertaining", "clear" or "fast". The results of a pair comparison also showed that the combination of spatial apartment-based menus was preferred (98.61%), which faces the least preferred linear store-based one (15.28%). This was also confirmed by comments like "so hard" or "it will take a long time" for linear store-based and "this was cool", "great" or "very intuitive" for spatial apartment-based after the corresponding demonstration videos were shown. The majority of participants would like this combination to be integrated into current online shops.

Limitations

The major drawback of this study is the limited amount of 36 products tested. This applies in particular to the remarkably high ratings of preference questionnaires, which are often close to the optimal rating. Such ‚excellent' results rarely occur in practice and are probably due to the limited test conditions. The scope and thus the number of products and categories of real online shops is usually much larger and therefore more complex. It might well be the case that the apartment metaphor does not scale with a large amount of products in its current form. It might require another layers (e.g. including "Toaster" or "Fridge" categories).

Furthermore, online stores have a wider range of functions. The implemented prototype can therefore certainly not reflect the complexity of a real online shop, but forms the basis and new insights for a rethinking in the area of menus in online shops. In addition, the selected products are mainly based on a list of frequently sought-after products from a particular market. Thus it cannot be completely excluded that other products can be found more easily with the traditional categorization. Overall, expectations for measurements in a fully functional online shop should be realistically lowered overall. However, the clear significant differences show that spatial apartment-based menus should still be preferred to the others.

5.3.7 Conclusion

Current online shops may be functional and efficient, but the possibilities of increasing performance and user preferences are not exhausted yet. Especially when it comes to exploratory setting in which a user is trying to find something from a menu, those shops can be significantly improved. Even though related work recommends to abandon linear and store-based menu interfaces [6, 259, 200], most current online shops are still employing these. In this respect, we investigated two menu representations (Linear, Spatial) and categorizations (Store, Apartment) in an online shopping prototype.

The Apartment metaphor [3, 287] turned out to be an effective way to support consumers to quickly, easily understand and use the offered information by filtering out desired parts. The success rate was 42% higher and led to 42% faster search times than with with stored-based concept. Furthermore, spatial menus performed significantly better than linear menus with about 12% higher success rate on average. Excellent usability and UX ratings indicate that spatial apartment-based UIs increase understanding and reliability, while low workload indicates that the intuitive apartment could lead to less frustration. Hence, this experiment confirmed the previous approaches and demonstrated that a spatial representation in combination with a corresponding categorization leads to significant performance increases. While we do not claim absolute generalizability for all stores, this work highlights the potential for improvement.

5.4 Experiment 3: A Virtual Reality Shopping Experience using the Apartment Metaphor

In this work, we propose a VR shop concept where product placement is not organized in shelves but through spatial placement in appropriate locations in an apartment environment. We thus investigated how the spatial arrangement of products in a non-retail environment affects the user, and how the actual shopping task can be supported in VR. In order to answer these questions, we designed two product selection and manipulation techniques (grabbing and pointing) and two VR shopping cart concepts (a realistic basket and an abstract one) and evaluated them in a user study. The results indicate that product interaction using pointing in combination with the abstract cart concept performs best with regard to error rate, UX and workload. Overall, the proposed apartment metaphor provides excellent customer satisfaction, as well as a particularly high level of immersion and UX, and it opens up new possibilities for VR shopping experiences that go far beyond mimicking real shop environments in VR.

5.4.1 Introduction

Online shopping bypasses many disadvantages of conventional stores like limited opening hours and is more focused on functionality. However, this focus comes at a cost, leading to limited search functionality and product visualization [169]. Most online shops only present products using text and images, while customers of physical stores can interact with products and view them from every side. VR has the potential to create novel shopping experiences that combines the benefits of on- and offline stores [282].

Motivation

Based on the experiment on product placement in a virtual apartment (see Section 5.3), we designed and evaluated a VR shopping approach using the *Apartment* metaphor. We investigated two different shopping cart representations for our prototype: an isomorphic shopping basket known from physical stores, and a non-isomorphic concept. Furthermore,

Figure 5.25: Virtual Shopping Environment using Apartment Metaphor

we assumed that the basket would increase the feeling of presence, because of its intuitiveness and familiarity, whereas the virtual sphere as shopping cart representation would outperform on UX, workload and performance, mainly because it has no physics and the products could be added faster. We further explored two approaches for product interaction, i.e. selection and manipulation of products. This motivates the following research question:

> *Does the isomorphic concepts provide higher user preference due to their familiarity, or can the user adapt to the non-isomorphic methods?*

To answer this question, we conducted a study to evaluate the UX using two isomorphic and non-isomorphic interaction methods and shopping cart modes with regard to the VRUX model (see Section 3.4), including task performance and user's preference. In this study, the task goal was to search for a product in the virtual apartment, select and manipulate the product using different techniques (*Grab* vs. *Beam*) and put it into different types of shopping carts (*Basket* vs. *Sphere*). For each search trial, task performance has been measured including task completion time and error rate. Furthermore, each participant had to complete multiple questionnaires for each combination of carts and methods to measure the user's preferences concerning immersion, motion sickness, workload and UX. Based on those scores, we conclude that the laser beam selection (*Beam*) and a virtual shopping cart

(*Sphere*) was preferred regarding UX and workload, as well as being more efficient concerning error rate.

Since many existing VR stores try to simulate conventional store interactions, we instead compared an isomorphic shopping basket representation and virtual 3D product manipulation techniques with non-isomorphic approaches. The isomorphic concept represents shopping in a physical store holding a realistic shopping basket in one hand and grabbing products with the other. The non-isomorphic concept is designed to use the capabilities of VR, e.g. providing users with a "magic" laser beam for product selection and manipulation. Nonetheless, the results of the study indicated that our application had an overall good UX, which was best for the combination of both non-isomorphic concepts. Henceforth, we can assume that the user successfully adapted to the methods.

5.4.2 Concept

In this section, we describe the main parts of this work, including the implemented product selection and manipulation techniques, as well as shopping cart representations. Besides that, we describe system components like product representation, categorization and placement, navigation in the environment and the virtual apartment as the environment itself.

Interaction Techniques

We present two implementations of interaction techniques to select and manipulate products in a virtual shopping environment. Virtual 3D manipulation tasks, like product interaction in a VR shop, combine target selection and manipulation [29]. For the isomorphic shopping experience we used a selection and manipulation technique based on the hand metaphor [242], whereas the non-isomorphic experience uses an adaptation of an interaction by pointing technique based on a laser tractor beam metaphor [29, 229, 52, 145].

Grabbing the Product This concept utilizes a motion controller with a button, which triggers the interaction. In our case, the HTC Vive controller is held by the user's dominant hand and the grip button triggers the inter-

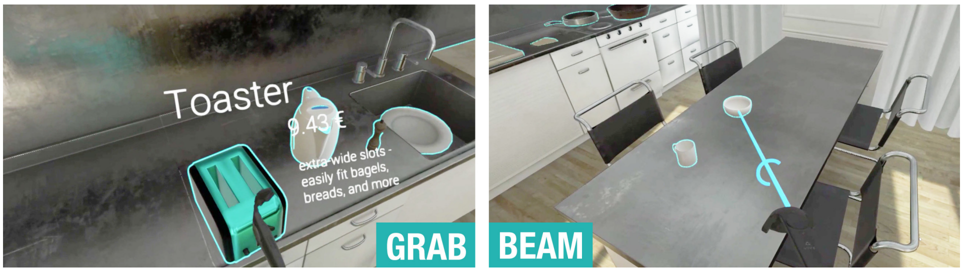

Figure 5.26: This figure shows the two implemented product interaction techniques (Grab, Beam).

action. When the controller intersects with a product, it gets highlighted, which is visualized by a colored halo and by displaying the product's description and price (see Figure 5.26). Now, the user can grab the highlighted product by pressing the grip button, which hides the controller model. While the button is pressed, the product can be manipulated by moving and rotating the controller. The product can be dropped by releasing the button, and the controller becomes visible again.

In summary, this is a classical example for isomorphic object selection and manipulation. Therefore, it should be intuitive and familiar for the user, because it represents everyday interaction with products in a store or at home. Nevertheless, the major drawback is that if objects are out of reach, it requires additional movement by the user and can therefore have a negative impact regarding performance and preference.

Laser Tractor Beam The non-isomorphic counterpart uses the concept of interaction by Controller Pointing [130]. Here, the interaction is initiated by pressing the trigger button of the HTC Vive controller in the user's dominant hand. A product gets selected and highlighted when the blue laser beam intersects with it (see Figure 5.26). After a short dwell time (3s), visualized with a small indicator below the ray, the product is moved towards the controller analogously to the Tractor Beam metaphor [229, 29]. If the tractor beam phase has not been interrupted by releasing the trigger button, the product stops at some distance to the controller. Then, its position and orientation is now linked to the controller (see Figure 5.26), as in the grabbing technique, but here the controller remains visible.

The main advantage of this method is that the user can interact with products out of reach without extra physical movement. However, this method may be less familiar and intuitive to the user compared to the grab interaction. Furthermore, the pointing to the desired product becomes more complex the further away it is or when multiple products are close to each other or occluded.

Virtual Shopping Cart Representations

In the following, we present two implementations of virtual shopping cart representations, in which the user can add selected products.

Realistic Shopping Basket Concerning context-aware shopping experience [23], we claim that an integration of a virtual shopping cart is of crucial importance. The concept of our isomorphic cart representation is based on a real-world shopping basket. Here, a virtual basket is attached to the controller held by the user's non-dominant hand (see Figure 5.27). Products can be added to the basket by placing them inside the basket. The total price of contained products is displayed on the handle of the basket. Unfortunately, this representation is not fully realistic, because we had to overcome the problem of large products that do not fit into the shopping basket. So, larger products shrink in size when they come near the basket, which allows the basket to store many different products of different sizes (e.g. plants or televisions). The scaling is initiated when the product reaches a certain radius around the basket. Consequently, the product is scaled up to its original size when it leaves the trigger area.

The number of products that can be stored inside the basket is still limited to its physical bounds. Nevertheless, this basket allows users to always have an overview of the current dimensions of their purchase, i.e. amount and sizes, in contrast to list-based carts in online shops. Furthermore, interaction with the products inside the basket is still possible, which allows the user to view the product information or remove a single product from the basket. We expect that this isomorphic concept of a virtual shopping cart representation will be more familiar and intuitive to the user, because of its similarity to an everyday shopping basket. Nonetheless, the physical properties of the basket may cause issues, such as accidentally losing products due to swinging of the basket.

Virtual Shopping Sphere The non-isomorphic approach uses a virtual sphere object containing a shopping cart icon, which is placed above the non-dominant controller (see Figure 5.27). The "adding of a product" works differently than using the basket, where the user places the product physically inside the basket. In this method, the user places the product inside the sphere and releases the selection button to add the product to the cart. Then, if a product has been successfully added to the virtual cart, it loses its physical properties such as gravity (in contrast to the basket). The products inside the virtual cart are organized circularly around the sphere, where the radius is proportional to the number of products. Here too, products are scaled down when placed inside the cart, and the products remain interactive.

The main advantage of this concept is the almost unlimited amount of products which can be stored and stacked around the sphere without gravity, because their sizes change dynamically if the number of products increases. Furthermore, the products are better organized than with the basket, where they are constantly "flying around" inside the basket because of their physics. Of course, the physics could have been disabled for the basket, but it should represent a realistic and isomorphic representation of a shopping basket. The non-isomorphic virtual shopping cart representation may be less intuitive and familiar for users, which could affect UX.

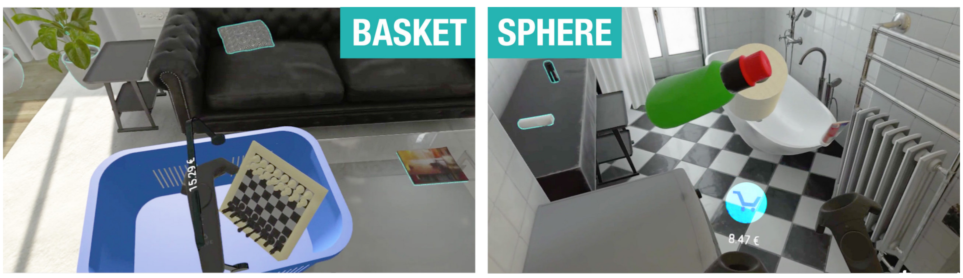

Figure 5.27: This figure shows the two implemented virtual shopping cart representations (Basket, Sphere).

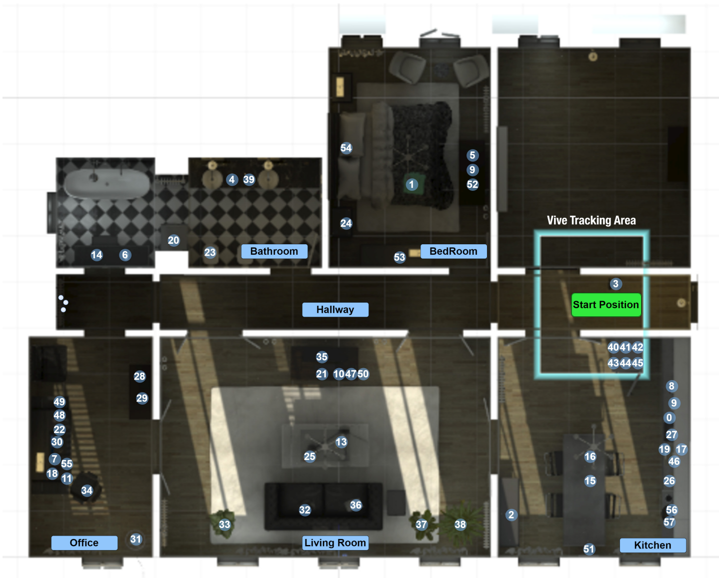

Figure 5.28: Apartment Categorization and Product Placement

Virtual Environment

In the design of our VR shop concept, we wanted to improve search performance and maximize immersion. We thus focused on recreating the experience of a believable apartment. The apartment should create familiarity for the users to help them navigate through it and find the products faster than in a physical store. In order to navigate the environment, we used the current standard concept for movement in commercial VR systems, namely the “Point and Teleport” method [33]. All products are placed at locations inside the apartment based on our online shop prototype using a spatial apartment-based menu interface (see Section 5.3). To avoid multi-placement and to make the planned experiment easier to replicate, we have chosen the most frequently mentioned product locations among the prior experiment results (see Figure 5.28).

Each product in our prototype is visualized by a 3D model representing its real-world counterpart. To separate the interactive products visually from the environment, they are highlighted with a blue outline (see Figure 5.25). We used physics for every product, including gravity, and set appropriate parameters for every single product to increase immersion. To overcome accidental displacements, dropped products outside the cart "re-spawn" in its original location after five seconds. Every product offers detailed information like its name, a short description and its price, which is displayed if the user selects a product or uses an implemented info ray to allow the user to view product information from afar and without triggering a selection. The yellow info ray is triggered by pressing the touch pad button on the user's dominant hand controller.

5.4.3 User Evaluation

We conducted an experiment to investigate VR shopping using the Apartment metaphor with respect to user performance, preference, and unmet needs. In this context, we evaluated two product interaction methods (*Grab* vs. *Beam*) and shopping cart representations (*Basket* vs. *Sphere*).

Hypothesis

- H_1 The task can be **performed more efficiently** using a tractor beam (*Beam*) and adding into a realistic shopping basket (*Basket*).
- H_2 Non-isomorphic interaction (*Beam*) is **preferred over** isomorphic interaction (*Grab*) with regard to **UX** and **task workload**.
- H_3 Non-isomorphic cart representation (*Sphere*) is **preferred over** the isomorphic shopping cart (*Basket*) with regard to **UX** and **task workload**.
- H_4 Non-isomorphic interaction (*Beam/Sphere*) is **preferred** with regard to **UX** and **task workload**.

Participants

For the experiment 16 unpaid participants (4 female) were recruited from the university's campus; they were aged between 21 and 33 years ($M = 24.82$, $SD = 3.22$). The overall shopping frequency on different devices (PC, smartphone, tablet, supermarket), disregarding the type of goods, was rated on scales from 0 (never) to 6 (several times daily). While they tend to shop more rarely on tablets ($M = 1.00, SD = 0.00$) or smartphones ($M = 1.80, SD = 0.75$), the majority prefer to shop in conventional stores ($M = 4.80, SD = 0.87$) or in online shops using a PC/laptop ($M = 3.60, SD =$

1.11). Furthermore, the average experience level with VR applications was rated rather low overall ($M = 1.9, SD = 0.88$). Finally, the participants were asked to rate the appropriateness of goods for VR shops on a scale from 1 (very relevant) to 6 (very irrelevant):

$$\begin{aligned} furniture: &\quad M = 2.20, SD = 1.17 \\ traveling: &\quad M = 2.40, SD = 1.69 \\ realestate: &\quad M = 2.40, SD = 2.16 \\ electronics: &\quad M = 3.70, SD = 1.56 \\ clothes: &\quad M = 4.30, SD = 2.06 \\ groceries: &\quad M = 6.00, SD = 1.35 \end{aligned}$$

Apparatus

The VR system used an HTC Vive and ran on a Windows 10 machine with Unity 5.5.4. A standard desktop computer was used with an i7 CPU, 16 GB RAM and Nvidia GeForce GTX 980Ti graphics. Besides experiment control, this PC was also used for filling out questionnaires by the participant. It is worth to mention that the frame rate of 60 fps was the same in all conditions. Two Vive controllers were used for interaction in the environment. The Vive lighthouses were installed about $2.5m$ above the ground in two opposite corners to span a maximum tracking area of approximately $4m \times 4m$. The participants were standing in its center while performing the tasks.

Design

The experiment used a within-subjects design with two independent variables having two levels:

- Product interaction (*Grab*, *Beam*)
- Shopping cart representation (*Basket*, *Sphere*)

Both conditions were counterbalanced using a Latin square. This amounted to 16 participants $\times$ 2 techniques $\times$ 2 carts $\times$ 10 product searches $= 640$ trials.

Overall, 60 different products were evenly distributed all over the virtual apartment according to their most probable location, based on the results of our second experiment (see Section 5.3). In order to ensure equal conditions for every participant, all trials started at the same physical and virtual position. The participants received only minimal instruction about the functionality of the different interaction types, so that no explicit conceptual model was assigned to them. The six dependent variables were:

- *Performance* (Task Completion Time, Error Rate)
- *Preference* (Workload, UX, Motion Sickness, Immersion)

Task

In each task, the participant performed a product search using a combination of the two product interaction methods (*Grab*, *Beam*) and the two shopping cart representations (*Basket*, *Sphere*), see Figure 5.4.2.

Each task starts with an exploration task followed by a search task. We chose an introducing exploration task without an explicit goal to browse the environment and obtain information about the rooms, orient the user to the world and build up knowledge. Besides that, in this training phase the user was able to get familiar with the interaction techniques and cart modes. Here, simple colored quads were randomly placed all over the environment to prevent memorizing the product locations. Each search task consisted of ten trials (product searches) in a row. Before each trial, the participant had to position herself in the center of the tracking area to ensure equal starting conditions. A trial was successfully completed when the target product was added into the cart within a time limit of 60 seconds, or counted as failed otherwise. The participant could travel through the apartment by using the standard Vive navigation techniques (natural walking within the tracking range, and teleportation).

Procedure

First, the participant was introduced to the experiment and signed an informed consent form. Then the experiment started with a 5-minute SteamVR tutorial to get familiar with the headset and the controllers, followed by a

5-minute exploration of the VE, i.e. the apartment without any products, shopping carts or interaction functionality. In the main part of the experiment, the participant had to perform all four tasks in Latin-square order. Each task started with a short training phase of up to five minutes, in which the participant could familiarize herself with the selection and manipulation technique, as well as shopping cart. Then, in the actual task, ten search trials had to be performed. Before each trial, the name of the target product appeared for five seconds. Then the target had to be found and added into the shopping cart.

After each task, the participant was asked to take off the HMD and fill out the post-task questionnaires to gather subjective feedback about the user's preferences, namely UEQ [157], NASA-TLX [107], MSAQ [93], and SUS [278]. A demographic questionnaire was filled out at the study's end.

5.4.4 Results

We use the same abbreviations as in the concept: *Grab* and *Beam* for the product selection techniques; *Basket* and *Sphere* for the shopping cart representations; *Grab/Basket*, *Beam/Basket*, *Grab/Sphere*, and *Beam/Sphere*.

Task Performance

The task performance metrics include quantitative measurements such as speed and accuracy during the third experiment. These metrics indicate to what extent users are able to cope with the task, i.e. the interaction methods and cart representations. They are computed per participant and condition as the average over the ten trials per task.

Task completion time was measured as the elapsed time in seconds to complete a single product search. The timer started when the countdown reaches zero and stopped automatically when the correct product has been added to the cart. We found no significant differences between the single tasks, the carts nor the selection techniques regarding speed: Basket lasted $16.84s$ ($SD = 10.11$) on average, whereas Sphere took $17.71s$ ($SD = 15.31$). Beam lasted $16.94s$ ($SD = 15.17$) and Grab $17.61s$ ($SD = 10.30$) on average.

Regarding the **error rate**, all participants successfully completed all trials (finding and added all correct products into the cart within the time limit), regardless of the selection technique or cart mode. When looking closer into the number of corrections (i.e. the number of times a wrong product was added to the cart and corrected before trial ended), a univariate ANOVA analysis showed significant differences regarding the number of corrections between the cart modes ($F_{(1,636)} = 20.64, p < 0.01, \eta^2 = 0.05$), but none for the selection. Sphere was best with no corrections, whereas Basket caused 0.24 ($SD = 0.73$) corrections on average.

User Preferences

For user preferences metrics, we collected a variety of subjective feedback to assess *UX*, *Task Workload*, *Motion Sickness* and *Immersion*, important in order to ensure comparability with related VR evaluations and the VRUX model (see Section 3.4).

We chose the UEQ [157] as an end-user questionnaire to measure **UX** in a quick and straightforward way. On a scale between -3 and 3 the overall UX was rated 1.40 ($SD = 0.64$) on average. Concerning the overall UX score, a univariate ANOVA showed significant differences between all four tasks ($F_{(3,636)} = 15.16, p < 0.01, \eta^2 = 0.06$). *Beam/Sphere* achieved the highest score ($M = 1.62, SD = 0.48$) and *Grab/Basket* the lowest ($M = 1.17, SD = 0.90$). Furthermore, cart modes ($F_{(1,636)} = 18.13, p < 0.01, \eta^2 = 0.03$) and selection techniques ($F_{(1,636)} = 27.28, p < 0.01, \eta^2 = 0.04$) also differed significantly regarding the overall UX score. *Sphere* was rated higher with an average of 1.50 ($SD = 0.54$) than *Basket* ($M = 1.30, SD = 0.71$) with respect to cart mode, whereas *Beam* was rated higher ($M = 1.53, SD = 0.46$) than *Grab* ($M = 1.28, SD = 0.75$) with respect to product selection technique.

However, the data was also subjected to a factor analysis, including the six UEQ factors Attractiveness (ATT), Perspicuity (PER), Efficiency (EFF), Dependability (DEP), Stimulation (STI), and Novelty (NOV); see Figure 5.29. Concerning these factors, a multivariate ANOVA showed significances between the cart modes (ATT: $F_{(1,636)} = 13.53, p < 0.01, \eta^2 = 0.02$; DEP: $F_{(1,636)} = 68.63, p < 0.01, \eta^2 = 0.09$; EFF: $F_{(1,636)} = 39.48, p < 0.01, \eta^2 = 0.06$), and the selection techniques (ATT: $F_{(1,636)} = 29.05, p < 0.01, \eta^2 = 0.04$;

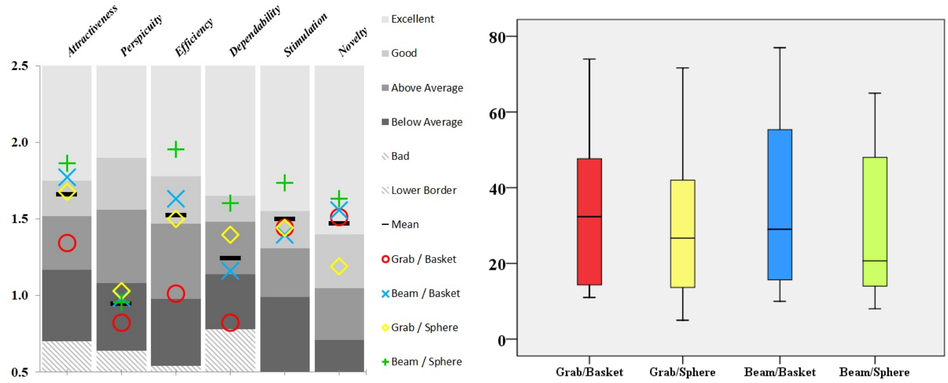

Figure 5.29: *Left:* UEQ results with respect to comparison benchmarks (see shaded boxes); to make it easier to read, this figure shows a detail part between 0.5 and 2.5, the original ranges between -3 and 3. *Right:* Results of the overall NASA TLX workload; to make it easier to read, this figure shows a detail part between 0 and 80; the original ranges from 0 to 100.

NOV: $F_{(1,636)} = 8.18, p < 0.01, \eta^2 = 0.01$; DEP: $F_{(1,636)} = 19.73, p < 0.01, \eta^2 = 0.03$; EFF: $F_{(1,636)} = 69.56, p < 0.01, \eta^2 = 0.09$). Interactions between the techniques and carts were also found for PER ($F_{(1,636)} = 22.82, p < 0.01, \eta^2 = 0.03$), NOV ($F_{(1,636)} = 5.48, p < 0.03, \eta^2 = 0.01$) and STI ($F_{(1,636)} = 4.86, p < 0.03, \eta^2 = 0.01$).

The **task workload** of the tested selection techniques and cart representations was assessed with NASA TLX [107]. Our system achieved an overall workload score of 32.42 ($SD = 20.38$) on average. Univariate ANOVAs showed no significant differences between the cart modes for overall workload, only between the single tasks ($F_{(1,636)} = 3.67, p < 0.01, \eta^2 = 0.02$) and selection techniques ($F_{(1,636)} = 9.05, p < 0.01, \eta^2 = 0.01$). *Beam/Sphere* was rated 29.77 ($SD = 19.44$) and *Grab/Basket* 36.26 ($SD = 22.49$) on average; (see Figure 5.29). As NASA TLX contains six subscales (MD: Mental Demand, PD: Physical Demand, TD: Temporal Demand, PF: Performance, EF: Effort, FR: Frustration), we conducted a multivariate ANOVA with regard to these factors. We found significant differences between the carts for FR ($F_{(1,636)} = 5.32, p < 0.01, \eta^2 = 0.02$), and between the techniques for PD ($F_{(1,636)} = 38.02, p < 0.01, \eta^2 = 0.05$), EF ($F_{(1,636)} = 14.17, p < 0.01, \eta^2 = 0.02$) and FR ($F_{(1,636)} = 21.66, p < 0.01, \eta^2 = 0.03$).

Motion sickness was measured with the well-established MSAQ [93] questionnaire. The system reached a total score of 16.83% ($SD = 6.53$) on average. There were no significant differences between the tasks (*Beam/Basket*: $M = 14.30\%, SD = 3.00$; *Grab/Basket*: $M = 15.00, SD = 4.20$; *Beam/Sphere*: $M = 15.14\%, SD = 3.00$; *Grab/Sphere*: $M = 15.28\%, SD = 4.13$) and the cart modes (*Basket*: $M = 16.71\%, SD = 6.31$; *Sphere*: $M = 16.95, SD = 6.74$) nor the selection techniques (*Beam*: $M = 16.89, SD = 6.00$; *Grab*: $M = 16.89, SD = 7.02$). A MSAQ consists of four categories (Gastrointestinal (G), Central (C), Peripheral (P), Sopite-related (S)); therefore, we conducted multivariate ANOVAs with regard to these factors. We found no significant differences between *Basket* and *Sphere*, but between *Grab* and *Beam* for G ($F_{(1,636)} = 25.48, p < 0.02, \eta^2 = 0.04$). Moreover, interactions between carts and techniques could be found for G ($F_{(1,636)} = 5.27, p < 0.03, \eta^2 = 0.01$).

The **immersion** was measured using the SUS questionnaire [278], where participants were asked to answer six questions on a scale between 1 and 7. Here, SUS Mean is the average across all six questions ($M = 5.06, SD = 0.95$), while SUS Count shows the amount of answers with 6 or 7 ($M = 3.07, SD = 2.19$). For SUS Mean and SUS Count, no significant differences between the selection techniques or carts exist, nor any between the tasks. Regarding SUS Mean, *Beam/Basket* was rated with 5.13 ($SD = 0.76$), followed by *Grab/Basket* ($M = 5.08, SD = 0.99$), *Grab/Sphere* ($M = 5.07, SD = 1.03$) and *Beam/Sphere* ($M = 4.96, SD = 1.00$). The SUS Count for *Grab/Basket* was rated 3.24 ($SD = 2.22$), followed by *Beam/Basket* ($M = 3.12, SD = 2.00$), *Beam/Sphere* ($M = 3.06, SD = 2.24$) and *Grab/Sphere* ($M = 2.88, SD = 2.28$).

5.4.5 Discussion

The study investigated two product interaction techniques (*Grab* vs. *Beam*) and two shopping cart representations (*Basket* vs. *Sphere*) in a VR shopping environment. In the following, we discuss the results of task performance (task completion time, error rate) and user preference (UX, task workload) with respect to our hypothesis.

Task Performance

The analysis of the error rate showed that all trials were successful, i.e. the participants collected all target products within the time limit. The experimental results of task performance showed no significant differences between the selection and manipulation techniques or cart modes, so H_1 has to be rejected. *Beam* was expected to be significantly faster than *Grab*, because the participant does not need to navigate to the product. Observations indicated that *Grab* was slower when the product lay on the ground or above head level, because the participant had to stretch or bend. Observations also indicated that the smaller trigger volume of *Sphere* might cause a speed loss. Here, the participant had to place the product inside the volume precisely, instead of let the product drop into the basket. Consequently, *Beam/Sphere* requires to increase *Sphere's* trigger volume for a better performance.

User Preferences

Concerning the UX ratings, *Beam* was preferred over *Grab* (H_2). This could be explained by the notably better ease of use and the lower physical demand of the tractor beam interaction. In addition, *Sphere* was preferred over *Basket* regarding UX, as expected (H_3). The gap in Perspicuity for all methods can be filled by introducing a conceptual model to the participant in the training phase or by clearer visual aids. These findings partly confirm that isomorphic interaction might be not the right choice for VR applications with regard to UX (H_4), which is also attested to prior work [26, 29].

The fully isomorphic combination (*Grab/Basket*) even achieved good Novelty and Stimulation ratings compared to the UEQ benchmark [157]. The participants were obviously naive to this "natural" interaction technique in particular the representation of a realistic shopping basket and thus experienced it as uncommon, stimulating and novel in VR. Apart from that, *Grab/Basket* was rated just below average in Dependability, and even got the lowest rating in Efficiency. This could be explained by the frustration when products are dropped unintentionally or unexpected behavior of the basket's physics. Overall, *Beam/Sphere* had excellent ratings for Attractiveness, Stimulation, Novelty and Efficiency, and good ratings for Dependability with respect to the UEQ benchmarks [261], as well as the highest overall UX score (H_4). This indicates that the non-isomorphic conditions are the optimal combination for purchasing products in a VR shop regarding UX (H_4). This combination performed slightly worse with regard to Stimulation and Novelty, in contrast to the excellent ratings for Beam/Basket.

As expected, *Beam/Sphere* had the lowest task load (H_4), and *Basket* turned out to be more frustrating than *Sphere* (H_3), whereas *Grab* was more frustrating and physically demanding than Beam (H_2). The NASA-TLX [107] results are in accordance with the UX ratings. The participants stated that using *Basket* was significantly more frustrating than *Sphere*, mainly because of the physics and limited space (or volume). So, as mentioned before, the physical behavior of the basket can cause unexpected behavior like unintended loss of products resulting in higher workload and frustration. Using *Grab* for product interaction was rated significantly more demanding and caused more frustration than *Beam*. The higher physical demand could be explained by the additional need to (physically) move to objects out of reach. Unintentional dropping of objects and picking them up from the ground additionally increased the frustration.

Limitations

As expected (see Section 5.4.2), some participants had problem to find certain products, e.g. tissues were expected to be in almost every room in accordance with the pilot study placements. However, we decided to place the products at the most frequent locations in order to make the main experiment more controllable and reproducible. In addition, some participants remarked that the "Point and Teleport" technique negatively influenced the immersion. But they also admitted that this method might currently be the best option to address the limited walking space. For some similar looking products like DVDs or books, it was hard to decide which they had to choose. Nonetheless, these issues could be easily addressed in future work using multiple product placements and other travel techniques.

5.4.6 Conclusion

Current online shops may be functional and efficient, but do not provide an immersive shopping experience, whereas physical stores lack efficiency and functionality [282] (e.g. customers are often frustrated when searching for products in offline stores). The proposed Apartment metaphor demonstrated the benefits of a combination of e-commerce and physical-inspired store environments (see Section 5.3). In this respect, we investigated two product interaction methods (Grab, Beam) and two shopping cart representations (Basket, Sphere) for an alternative VR shopping environment. The results show that *Beam/Sphere* outperforms the others in terms of error rate, UX and workload, whereas *Beam/Basket* was the fastest.

Overall, the experimental results indicated that our system was rated high for immersion and UX. To minimize motion sickness we would recommend to use *Beam/Basket* when designing a VR shopping environment, due to the better results regarding the gastrointestinal factor. Most of the participants enjoyed their experience with the VR Shop and showed interest in using it in the future. Hence, VR shopping has the potential to become a new shopping medium which combines the advantages of e-commerce and physical stores. However, this might not apply for all types of goods. The participants found the suitability for VR shops for electronics as above average, and very relevant for furniture, property or traveling. Therefore, we recommend that future VR shops should focus on other types of goods and product categories, such as furniture or travel.

Chapter 6

Conclusion and Outlook

In this last chapter, we will summarize the theoretical, technical and design contributions of this work. We will also identify possible directions for future work that are based on the results of this work but go beyond its scope. Finally, we will conclude this work with final remarks.

6.1 Thesis Summary

The goal of this thesis was to investigate shopping and system control in VR with a focus on VR-specific factors and metrics to make future VR applications, their interfaces and devices more comparable.

Chapter 2 describes theoretical background and related work for this thesis. It includes a definition of Virtual Reality, as well as a categorization of human factors in VR and current VR systems. In addition, this chapter also presents existing interaction techniques for selection, manipulation and navigation in VR derived from the related work.

In Chapter 3, we discussed related evaluation approaches and metrics and we developed a concept for evaluation of UX in VR applications.

With respect to this novel approach, we investigated two different directions of everyday VR: System Control in VR (Chapter 4), including an empirical analysis of selection-based text entry and a comparison of planar and pseudo-haptic UIs for finger-based menu control in VR; and Shopping in Virtual Environments (Chapter 5).

Chapter 4 includes the exploration of selection-based text entry in VR, finger-based UIs for menu control, as well as authentication in VR. In Section 4.1, we discussed the design space for assessing selection-based text entry in VR and evaluate six implemented methods that span different parts of the design space with respect to our VRUX model. Next, in section 4.2, we presented pseudo-haptic UI controls for mid-air finger-based menu control in VR environments and compared them with conventional 2D controls in a user study. Finally, we investigated potential design mechanisms and provide guidelines for mid-air finger-based VR menu control based on experimental results and observations during the studies.

Chapter 5 starts with a theoretical background focused on VR in retail, alternative shopping concepts, and commercial VR shop applications. As part of the user-centered design cycle, we conducted a customer survey (see Section 5.2.2) to understand the user and explore the main characteristics between on- and offline shops (see Section 5.2.2), followed by principles for VR shops (see Section 5.2.2).

Section 5.2 describes our design of an immersive VR online shopping environment. We tried to maintain the benefits of online shops, like search functionality and availability, while simultaneously focusing on shopping experience and immersion. Next, in Section 5.3, we presented a realistic and interactive map-based spatial menu representation based on prior research findings. In a comparative user study, we evaluated our new menu categorization and representation in comparison with the current standard in online shops, based on real data from a local retailer. In Section 5.4, we described details of a comparative study and its findings about the Apartment metaphor for representing a VR shop and we compared isomorphic selection and manipulation of products.

6.2 Contributions

In this section, we want to outline the theoretical, technical and design contributions of this thesis:

6.2.1 Theoretical contributions

In Chapter 3, we presented existing approaches and knowledge from the field of 3DUI and VR evaluation. Furthermore, we introduced our new Virtual Reality User Experience (VRUX) evaluation approach with focus on 3DUI- and VR-specific external factors. Here, we combined the metrics of evaluation of 3DUIs and the characteristics of VR (see Section 3.4). This allows a more refined and differentiated classification of interaction techniques and UIs for VR. By classifying common affordable and commodity VR systems based on their interactivity, comfort and graphics quality, we allow characterization of VR systems more fine-grained (see Section 2.2.5). In the course of this thesis, we examined the influence of consumer input and output devices in more detail regarding performance and preferences.

6.2.2 Technical contributions

While military, education and gaming topics are well investigated, basic topics such as shopping, text entry or menu control in VR are lagging behind. In Chapter 4, we presented a finger-based pseudo-haptic UI for menu control in VR based on physical metaphors, as well as text input methods for VR using consumer-ready hardware. Finally, we contributed advanced technical solutions for the creation of VR applications in the areas of shopping and system control. In Chapter 5, we presented two immersive VR online shopping environments to fill in the missing link between on- and offline shopping. Here, we designed and implemented those VR shops maintaining the benefits of online shops, such as search functionality and availability, while simultaneously focusing on shopping experience, clarity and immersion.

6.2.3 Design contributions

The technical aspects of VR applications, the context in which they are deployed and the exposure of their content to a large audience raise the need for a tailored design in terms of design guidelines and prototyping tools. When designing for experiences in VR a new set of design considerations comes into play than when designing for 2D screens. We allow future

VR designers and developers of VR shops to create experiences that does not frustrate or make users feel nauseous by providing principles to guide the work. In the Chapters 4 and 5, we contributed guidelines including actionable insights on how to optimize performance, usability, satisfaction, and experience for the users of VR. In this context, we provided main characteristics of on- and offline shopping (see Section 5.2.2), as well as a list of potential guidelines and lessons learned to inform the design of VR shop (see Section 5.2.2) and system control (see Sections 4.1.6 and 4.2.6). The design spaces and the evaluated methods will provide a solid baseline for comparison of future VR applications.

6.3 Future Work

Even though the results and contributions that we present in this work open up new possibilities, they also leave open challenges that can be addressed in future work. Our categorizations show that, on the one hand, little research has been done in the fields of UX evaluation. With regard to UIs and interaction techniques, the areas in the context of everyday VR, such as “Shopping in Virtual Environments” and “Text Entry and Menu Control”, still offer space for further investigations. In particular, the closer examination of the individual VR-specific factors, such as motion sickness or immersion, have only been sporadically examined, although they represent a convincing sphere of influence for VR experiences. In the following, we describe some exemplary scenarios for future work in the fields of shopping, text interaction and menu control in VR.

6.3.1 Shopping in Virtual Environments

In summary, our VR shop creates compelling virtual sensory richness through which customers can experience the value of the product information more richly and engage in a more active shopping activity, compared to ordinary online shopping applications. In an ordinary shopping mall, customers have to use a rather plain UI, leading to lower customer satisfaction. This might cause customers to become passive observers, merely observing the information. Whereas VR customers are engaged in the inspection and control of the 3D visualized target products, due to the virtual sensory richness driven by the 3D environment and UI. But information overload in VR is more likely and should be avoided. Therefore, different layouts and representations of VR shops should be explored and compared, like graph-based approaches, or even more abstract concepts like searching in a virtual apartment. So, another aspect what should be studied in more detail in the future is the visualization of a virtual store. Our mobile VR shop prototype was based on existing layout data, but due to performance issues of WebVR and smartphones we chose a smaller retail space of about 180 square meters within one single floor.

Besides different store layouts, future studies should explore differences between the store size in all three dimensions, i.e. different number of floors and sizes of the market area. Because people tend to think two-dimensionally, and even the front-back axis is more accessible than the left-right axis [124, 324]. So, it might be easier for customers to orientate in virtual stores or malls with less floors and less turnings to left or right.

6.3.2 Interaction with Text in Virtual Reality

In this thesis, we have studied text entry in VR using a virtual keyboard and discussed the design space including criteria for assessing VR text entry methods. Text entry is an essential part of HCI and there is still much research needed. Design annotation (e.g. for 3D artists or architects), filename entry or parameter setting, and communication between users are just a few applications for text entry in VR. Future VR systems (e.g. diaries, shops or social networks) may be designed to enable the user to stay in VR for longer times and, therefore, longer text entry needs to be feasible, too. Finally, the qualifying techniques need to be evaluated in the context of interactive immersive VEs.

Since development of the first personal computers, research has addressed the issue of pleasant and efficient reading on computer screens [72]. After the development of new devices with different output conditions, e.g. e-book readers, tablets and smartphones, reading research turned on to these devices [256]. The constant growing development and availability of technology in various areas of VR brought new challenges and opportunities regarding reading in this field. VR UI designers should be careful with too bright and overloading colors, because the user's eyes are only a couple centimeters away from the screen. In addition, too thin fonts or fonts with serifs can be very hard to read, because there are simply not enough pixels for a clear view of the text. Low display resolution, use of lenses and short eye-display distance demands specific adaptions of text presentation in VR, while new possibilities arisen. For example, users may benefit from encapsulation in another world and may read texts on the beach listening to the sea, but actually sitting back home in a sofa.

6.4 Concluding Remarks

In this thesis, we investigated holistically interaction in VR. VR technology might be a catalyst for new experiences, but we should keep in mind that it could not solve all our problems. Moreover, interactions with 3D virtual objects in particular need to be designed carefully and in a human-centered way. With the theoretical analysis, the novel evaluation model for UX in VR and the provided design space and guidelines, this thesis builds a basis for the generalized and structured design and development of UIs for everyday VR applications. With the work described in this thesis, our findings and the lessons we learned through the design and deployment of our solutions, we want to underline the importance of VR-specific and human-centered design of interactivity for future VR experiences. Because in the future – not necessarily even a distant future – technology aspects will be no longer excuse for bad usability of VR applications.

List of Figures

Bibliography

[1] C. Abras, D. Maloney-krichmar and J. Preece. User-Centered Design. In *Bainbridge, W. Encyclopedia of Human-Computer Interaction* (2004), vol. 37, pp. 445–456.

[2] J. Accot and S. Zhai. Beyond Fitts' Law: Models for Trajectory-based HCI Tasks. In *Proc. of the ACM SIGCHI Conf. on Human Factors in Computing Systems* (1997), CHI '97, pp. 295–302.

[3] S. Adam, K. S. Mukasa, K. Breiner and M. Trapp. An Apartment-based Metaphor for Intuitive Interaction with Ambient Assisted Living Applications. In *Proc. of the 22Nd British HCI Group Annual Conf. on People and Computers: Culture, Creativity, Interaction* (2008), BCS-HCI '08, pp. 67–75.

[4] B. Adelstein, L. Li, J. Jerald and S. R. Ellis. Suppressed Perception of Head-Referenced Image Motion during Head Movement. *Proc. of the Human Factors and Ergonomics Society Annual Meeting 50* (10 2006).

[5] R. Agarwal and E. Karahanna. Time Flies when You're Having Fun: Cognitive Absorption and Beliefs about Information Technology Usage. *MIS Q. 24*, 4 (2000), pp. 665–694.

[6] D. Ahlström, A. Cockburn, C. Gutwin and P. Irani. Why It's Quick to Be Square: Modelling New and Existing Hierarchical Menu Designs. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems* (2010), CHI '10, pp. 1371–1380.

[7] J. Ahn and K. Min. VirtualPhonepad: A Text Input Tool for Virtual Environments. *Advances in Artificial Reality and Tele-Existence* (2006), pp. 58–64.

[8] I. G. Angus and H. A. Sowizral. Embedding the 2D Interaction Metaphor in a Real 3D Virtual Environment. In *IS&T/SPIE's Symp. on Electronic Imaging: Science & Technology* (1995), Int. Society for Optics and Photonics, pp. 282–293.

[9] S. Aukstakalnis and D. Blatner. Silicon Mirage: The Art and Science of Virtual Reality. *Peachpit Press* (1 1992).

[10] S. Azenkot and S. Zhai. Touch Behavior with Different Postures on Soft Smartphone Keyboards. In *Proc. of the 14th Int. Conf. on Human-computer Interaction with Mobile Devices and Services* (2012), MobileHCI '12, pp. 251–260.

[11] M. Bachynskyi, G. Palmas, A. Oulasvirta and T. Weinkauf. Informing the Design of Novel Input Methods with Muscle Coactivation Clustering. *ACM Trans. Comput.-Hum. Interact. 21*, 6 (Jan. 2015), pp. 30:1–30:25.

[12] D. R. Badock and J. G. May. Vision and Virtual Environments. In *Handbook of Virtual Environments - Design, Implementation, and Applications*, Kelly S. Hale and Kay M. Stanney, Eds., 2nd ed. CRC Press, Boca Raton, Fla, 2014.

[13] J. P. Baker and A. P. Paplinski. Virtual Reality Tracking System. Tech. rep., Robotics and Digital Technology, Monash University, 1994.

[14] R. Balakrishnan and G. Kurtenbach. Exploring Bimanual Camera Control and Object Manipulation in 3D Graphics Interfaces. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems* (1999), pp. 56–62.

[15] T. Baudel and M. Beaudouin-Lafon. Charade: Remote Control of Objects Using Free-hand Gestures. *Commun. ACM 36*, 7 (July 1993), pp. 28–35.

[16] J. Bauer and A. Ebert. *Mobile Devices for Virtual Reality Interaction. A Survey of Techniques and Metaphors*. Springer International Publishing, Cham, 2015, pp. 91–107.

[17] N. Bee and E. André. Writing With Your Eye: A Dwell Time Free Writing System Adapted to the Nature of Human Eye Gaze. In *International Tutorial and Research Workshop on Perception and Interactive Technologies for Speech-Based Systems* (2008), pp. 111–122.

[18] C. H. Bei Bei, N. S. Chaudhari and J. C. Patra. Design of a Virtual Shopping Mall: Some Observations. In *Int. Conf. on Cyberworlds* (Nov 2005), CW '05.

[19] N. Bevan. What is the Difference Between the Purpose of Usability and User Experience Evaluation Methods. In *Proc. of the Workshop UXEM* (2009), vol. 9, pp. 1–4.

[20] G. Bhatt. Bringing Virtual Reality for Commercial Web Sites. *Int. Journal Human-Computer Studies 60*, 1 (2004), pp. 1–15.

[21] F. Biocca. Will Simulation Sickness Slow Down the Diffusion of Virtual Environment Technology? *Presence: Teleoperators & Virtual Environments 1*, 3 (1992), pp. 334–343.

[22] D. Black, N. J. Clemmensen and M. B. Skov. Pervasive Computing in the Supermarket: Designing a Context-Aware Shopping Trolley. *International Journal of Mobile Human Computer Interaction (IJMHCI) 2*, 3 (2010), pp. 31–43.

[23] D. Black, N. J. Clemmensen and M. B. Skov. Pervasive Computing in the Supermarket: Designing a Context-Aware Shopping Trolley. *Int. Journal Mobile Human-computer Interaction 2*, 3 (July 2010), pp. 31–43.

[24] B. Bolte, F. Steinicke and G. Bruder. The Jumper Metaphor: An Effective Navigation Technique for Immersive Display Setups. In *Proc. of Virtual Reality International Conf.* (Laval, France, 2011).

[25] C. W. Borst and R. A. Volz. Preliminary Report on a Haptic Feedback Technique for Basic Interactions with a Virtual Control Panel. In *Proc. of the Int. Conf. on EuroHaptics* (2003).

[26] D. A. Bowman, J. Chen, C. A. Wingrave, J. F. Lucas, A. Ray, N. F. Polys, Q. Li, Y. Haciahmetoglu, J.-S. Kim, S. Kim et al. New Directions in 3D User Interfaces. *IJVR 5*, 2 (2006), pp. 3–14.

[27] D. A. Bowman and L. F. Hodges. An Evaluation of Techniques for Grabbing and Manipulating Remote Objects in Immersive Virtual Environments. In *Proc. of the 1997 Symp. on Interactive 3D Graphics* (1997), I3D '97.

[28] D. A. Bowman, D. B. Johnson and L. F. Hodges. Testbed Evaluation of Virtual Environment Interaction Techniques. *Presence: Teleoperators and Virtual Environments 10*, 1 (2001), pp. 75–95.

[29] D. A. Bowman, E. Kruijff, J. J. LaViola and I. Poupyrev. *3D User Interfaces: Theory and Practice*. Addison Wesley Longman Publishing Co., Inc., 2004.

[30] D. A. Bowman and R. P. McMahan. Virtual Reality: How Much Immersion Is Enough? *Computer 40*, 7 (2007), pp. 36–43.

[31] D. A. Bowman, R. P. McMahan and E. D. Ragan. Questioning Naturalism in 3D User Interfaces. *Commun. ACM 55*, 9 (2012), pp. 78–88.

[32] D. A. Bowman, C. J. Rhoton and M. S. Pinho. Text Input Techniques for Immersive Virtual Environments: An Empirical Comparison. *Proc. of the Human Factors and Ergonomics Society Annual Meeting 46*, 26 (2002), pp. 2154–2158.

[33] E. Bozgeyikli, A. Raij, S. Katkoori and R. Dubey. Point & Teleport Locomotion Technique for Virtual Reality. In *Proc. of the 2016 Annual Symp. on Computer-Human Interaction in Play - CHI PLAY '16* (Austin, Texas, USA, 2016), pp. 205–216.

[34] J. H. Brockmyer, C. M. Fox, K. A. Curtiss, E. McBroom, K. M. Burkhart and J. N. Pidruzny. The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology 45*, 4 (2009), pp. 624–634.

[35] D. C. Brogan, R. A. Metoyer and J. K. Hodgins. Dynamically Simulated Characters in Virtual Environments. *IEEE Comput. Graph. Appl. 18*, 5 (1998), pp. 58–69.

[36] J. Brooke et al. SUS - A Quick and Dirty Usability Scale. *Usability Evaluation in Industry 189*, 194 (1996), pp. 4–7.

[37] F. P. Brooks. What's Real About Virtual Reality? *IEEE Comput. Graph. Appl. 19*, 6 (1999), pp. 16–27.

[38] P. Brooks and B. Hestnes. User Measures of Quality of Experience: Why Being Objective and Quantitative is Important. *IEEE network 24*, 2 (2010), pp. 8–13.

[39] G. Bruder, F. Steinicke, K. Hinrichs and M. Lappe. Reorientation during Body Turns. In *Proc. of the 15th Joint virtual reality Eurographics conference on Virtual Environments* (Lyon, France, 2009), Eurographics Association, pp. 145–152.

[40] S. Bryson. Virtual Reality Takes on Real Physics Applications. *Computers in Physics 6*, 4 (1992), pp. 346–352.

[41] M. Buffa and J. C. Lafon. 3D Virtual Warehouse on the Web. In *Proc. of the IEEE Int. Conf. on Information Visualization* (2000), pp. 479–484.

[42] G. C. Burdea and P. Coiffet. *Virtual Reality Technology*, 2 ed. John Wiley & Sons, Inc., 2003.

[43] B. Buxton and G. W. Fitzmaurice. HMDs, Caves &Amp; Chameleon: A Human-centric Analysis of Interaction in Virtual Space. *SIGGRAPH Comput. Graph. 32*, 4 (Nov. 1998), pp. 69–74.

[44] F. Cacciola. Triangulated Surface Mesh Simplification. *CGAL User and Reference Manual 4* (2016).

[45] L.-W. Chan, H.-S. Kao, M. Y. Chen, M.-S. Lee, J. Hsu and Y.-P. Hung. Touching the Void: Direct-touch Interaction for Intangible Displays. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems* (2010), CHI '10, pp. 2625–2634.

[46] T. Chen, Z. Pan and J. Zheng. EasyMall - An Interactive Virtual Shopping System. In *Proc. of the 5th Int. Conf. on Fuzzy Systems and Knowledge Discovery* (Oct 2008), vol. 4, pp. 669–673.

[47] H.-I. Cheng and P. E. Patterson. Iconic Hyperlinks on e-Commerce Websites. *Applied Ergonomics 38*, 1 (2007), pp. 65 – 69.

[48] L. Chittaro and R. Ration. Adding Adaptive Features to Virtual Reality Interfaces for e-Commerce. In *Int. Conf. on Adaptive Hypermedia and Adaptive Web-Based Systems* (2000), pp. 86–97.

[49] J. Ciger, M. Gutierrez, F. Vexo and D. Thalmann. The Magic Wand. In *Proc. of the 18th spring conference on Computer graphics* (2003), p. 119.

[50] G. Cirio, P. Vangorp, E. Chapoulie, M. Marchal, A. Lecuyer and G. Drettakis. Walking in a Cube: Novel Metaphors for Safely Navigating Large Virtual Environments in Restricted Real Workspaces. *IEEE Transactions on visualization and computer graphics 18*, 4 (2012), pp. 546–554.

[51] D. Clark and M. J. Bailey. Virtual-virtual Haptic Feedback and Why It Wasn't Enough. In *Visualization and Data Analysis* (2002), vol. 4665, Int. Society for Optics and Photonics, pp. 308–319.

[52] R. M. S. Clifford, N. M. B. Tuanquin and R. W. Lindeman. Jedi ForceExtension: Telekinesis as a Virtual Reality Interaction Metaphor. In *2017 IEEE Symp. on 3D User Interfaces (3DUI)* (March 2017), pp. 239–240.

[53] G. Coates. Invisible Site: A Virtual Sho, 1992.

[54] S. V. G. Cobb, S. Nichols, A. Ramsey and J. R. Wilson. Virtual Reality-Induced Symptoms and Effects (VRISE). *Presence: Teleoper. Virtual Environ. 8*, 2 (Apr. 1999), pp. 169–186.

[55] A. Cockburn, C. Gutwin and S. Greenberg. A Predictive Model of Menu Performance. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems* (2007), CHI '07, pp. 627–636.

[56] A. Cockburn and B. McKenzie. 3D or Not 3D?: Evaluating the Effect of the Third Dimension in a Document Management System. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems* (2001), CHI '01, pp. 434–441.

[57] J. Cogburn and M. Silcox. Against Brain-in-a-Vatism: On the Value of Virtual Reality. *Philosophy & Technology 27*, 4 (2014), pp. 561–579.

[58] C. Comeau and J. Bryan. Headsight Television System Provides Remote Surveillance. *Electronics* (1961), pp. 86–90.

[59] N. R. Council and Others. *Virtual Reality: Scientific and Technological Challenges*. National Academies Press, 1995.

[60] F. Daiber, A. Krekhov, M. Speicher, J. Krüger and A. Krüger. A Framework for Prototyping and Evaluation of Sensor-based Mobile Interaction with Stereoscopic 3D. In *ACM ITS Workshop on Interactive Surfaces for Interaction with Stereoscopic 3D* (10 2013), ISIS3D-13, pp. 13–16.

[61] F. Daiber, M. Speicher, S. Gehring, M. Löchtefeld and A. Krüger. Interacting with 3D Content on Stereoscopic Displays. In *Proc. of The International Symp. on Pervasive Displays* (2014), PerDis '14, pp. 32:32–32:37.

[62] M. Daily, R. Sarfaty, J. Jerald, D. McInnes and P. Tinker. The 'CABANA': A Re-Configurable Spatially Immersive Display. In *3rd International Immersive Projection Technology Workshop, May* (1999), pp. 123–132.

[63] R. P. Darken, W. R. Cockayne and D. Carmein. The Omni-directional Treadmill: A Locomotion Device for Virtual Worlds. In *Proc. of the 10th Annual ACM Symp. on User Interface Software and Technology* (1997), UIST '97, pp. 213–221.

[64] D. Davies. Why One Shouldn't Make an Example of a Brain in a Vat. *Analysis 57*, 1 (1997), pp. 51–59.

[65] S. Davis, K. Nesbitt and E. Nalivaiko. A Systematic Review of Cybersickness. In *Proc. of the 2014 Conf. on Interactive Entertainment* (2014), IE '14, pp. 8:1–8:9.

[66] B. R. De Araùjo, G. Casiez and J. A. Jorge. Mockup Builder: Direct 3D Modeling on and Above the Surface in a Continuous Interaction Space. In *Proc. of Graphics Interface 2012* (2012), GI '12, pp. 173–180.

[67] O. De Troyer, F. Kleinermann, H. Mansouri, B. Pellens, W. Bille and V. Fomenko. Developing Semantic VR-Shops for e-Commerce. *Virtual Reality 11*, 2-3 (2007), pp. 89–106.

[68] M. Deering. High Resolution Virtual Reality. In *Proc. of the 19th Annual Conf. on Computer Graphics and Interactive Techniques* (1992), SIGGRAPH '92, pp. 195–202.

[69] G. Deleuze. Bergsonism, 1991.

[70] C. Demiralp, C. D. Jackson, D. B. Karelitz, S. Zhang and D. H. Laidlaw. CAVE and Fishtank Virtual-Reality Displays: A Qualitative and Quantitative Comparison. *IEEE Transactions on Visualization and Computer Graphics 12*, 03 (2006), pp. 323–336.

[71] P. R. Desai, P. N. Desai, K. D. Ajmera and K. Mehta. A Review Paper on Oculus Rift-A Virtual Reality Headset. *CoRR abs/1408.1173* (2014).

[72] A. Dillon. Reading from Paper versus Screens: A Critical Review of the Empirical Literature. *Ergonomics 35*, 10 (1992), pp. 1297–1326.

[73] W. Ding, P. Chen, H. Al-Mubaid and M. Pomplun. A Gaze-Controlled Interface to Virtual Reality Applications for Motor-and Speech-Impaired Users. *HCI International, San Diego, CA* (2009).

[74] T. Dingler, C. Giebler, U. Kunze, T. Wűrtele, N. Henze and A. Schmidt. Memory Displays: Investigating the Effects of Learning in the Periphery. In *Proc. of the 5th ACM International Symp. on Pervasive Displays* (2016), PerDis '16, pp. 118–123.

[75] A. Dix, J. Finley, G. Abowd and R. Beale. *Human-computer Interaction (2Nd Ed.).* Prentice-Hall, Inc., 1998.

[76] C. Domingues, S. Otmane and M. Mallem. 3DUI-EF: Towards a Framework for Easy Empirical Evaluation of 3D User Interfaces and Interaction Techniques. *The Int. Journal of Virtual Reality 9*, 1 (2010), pp. 73–80.

[77] A. T. Duchowski, V. Shivashankaraiah, T. Rawls, A. K. Gramopadhye, B. J. Melloy and B. Kanki. Binocular Eye Tracking in Virtual Reality for Inspection Training. In *Proc. of the 2000 Symp. on Eye tracking Research & Applications* (2000), pp. 89–96.

[78] A. Dvorak, N. L. Merrick, W. L. Dealey and G. C. Ford. Typewriting Behavior. *New York: American Book Company 1*, 6 (1936).

[79] B. M. Ens, R. Finnegan and P. P. Irani. The Personal Cockpit: A Spatial Interface for Effective Task Switching on Head-worn Displays. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems* (2014), CHI '14, pp. 3171–3180.

[80] A. S. Fernandes and S. K. Feiner. Combating VR Sickness through Subtle Dynamic Field-of-View Modification. In *Symp. on 3D User Interfaces* (2016), 3DUI '16, pp. 201–210.

[81] L. S. Figueiredo, E. Vilar Neto, E. Arruda, J. M. Teixeira and V. Teichrieb. Fishtank Everywhere: Improving Viewing Experience over 3D Content. In *Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience* (2014), Aaron Marcus, Ed., pp. 560–571.

[82] L. Findlater, K. Moffatt, J. McGrenere and J. Dawson. Ephemeral Adaptation: The Use of Gradual Onset to Improve Menu Selection Performance. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems* (2009), CHI '09, pp. 1655–1664.

[83] S. S. Fisher, M. McGreevy, J. Humphries and W. Robinett. Virtual Environment Display System. In *Proc. of the 1986 Workshop on Interactive 3D Graphics* (1987), I3D '86, pp. 77–87.

[84] S. S. Fisher, E. M. Wenzel, C. Coler and M. W. McGreevy. Virtual Interface Environment Workstations. In *Proc. of the Human Factors Society Annual Meeting* (1988), vol. 32, pp. 91–95.

[85] J. Fox, D. Arena and J. N. Bailenson. Virtual Reality: A Survival Guide for the Social Scientist. *Journal of Media Psychology: Theories, Methods, and Applications 21* (01 2009), pp. 95–113.

[86] D. Friedman. Brain-computer Interfacing and Virtual Reality. *Handbook of Digital Games and Entertainment Technologies* (2015), pp. 1–22.

[87] J. Frommel, S. Sonntag and M. Weber. Effects of Controller-based Locomotion on Player Experience in a Virtual Reality Exploration Game. In *Proc. of the International Conf. on the Foundations of Digital Games - FDG '17* (2017), pp. 1–6.

[88] J. L. Gabbard, D. Hix and J. E. Swan. User-Centered Design and Evaluation of Virtual Environments. *IEEE Comput. Graph. Appl. 19*, 6 (Nov. 1999), pp. 51–59.

[89] J. F. Garcia, A. L. Simeone, M. Higgins, W. Powell and V. Powell. Inside Looking out or Outside Looking in?: An Evaluation of Visualisation Modalities to Support the Creation of a Substitutional Virtual Environment. In *Proc. of the 2018 Int. Conf. on Advanced Visual Interfaces* (2018), AVI '18, pp. 29:1–29:8.

[90] G. Garcia-Perate, N. Dalton, R. Conroy-Dalton and D. Wilson. Ambient Recommendations in the Pop-up Shop. In *Proc. of the 2013 ACM Int. Joint Conf. on Pervasive and Ubiquitous Computing* (2013), UbiComp '13, pp. 773–776.

[91] T. A. Garner, W. Powell and V. Powell. *Everyday Virtual Reality*. Springer International Publishing, 2018, pp. 1–9.

[92] S. Gehring, M. Löchtefeld, F. Daiber, M. Böhmer and A. Krüger. Using Intelligent Natural User Interfaces to Support Sales Conversations. In *Proc. of the 2012 ACM International Conf. on Intelligent User Interfaces* (2012), IUI '12, pp. 97–100.

[93] P. J. Gianaros, E. R. Muth, J. T. Mordkoff, M. E. Levine and R. M. Stern. A Questionnaire for the Assessment of the Multiple Dimensions of Motion Sickness. *Aviation, Space, and Environmental Medicine 72*, 2 (2001), p. 115.

[94] Y. Gizatdinova, O. Špakov and V. Surakka. Comparison of Video-based Pointing and Selection Techniques for Hands-free Text Entry. In *Proc. of the International Working Conf. on Advanced Visual Interfaces* (2012), AVI '12, pp. 132–139.

[95] H. Gleitman. Basic Psychology. *Worth and Company* (1996).

[96] G. González, J. P. Molina, A. S. García, D. Martínez and P. González. Evaluation of Text Input Techniques in Immersive Virtual Environments. In *New Trends on Human–Computer Interaction*. Springer, 2009, pp. 109–118.

[97] A. Graybiel, C. D Wood, E. F Miller and D. B Cramer. Diagnostic Criteria for Grading the Severity of Acute Motion Sickness. *Aerospace Medicine 39* (06 1968), pp. 453–5.

[98] T. Grechkin, J. Thomas, M. Azmandian, M. Bolas and E. Suma. Revisiting Detection Thresholds for Redirected Walking: Combining Translation and Curvature Gains. In *Proc. of the ACM Symp. on Applied Perception* (2016), pp. 113–120.

[99] D. A. Griffith. An Examination of the Influences of Store Layout in Online Retailing. *Journal of Business Research 58*, 10 (2005), pp. 1391–1396.

[100] D. A. Griffith, R. F. Krampf and J. W. Palmer. The Role of Interface in Electronic Commerce: Consumer Involvement with Print versus On-line Catalogs. *International Journal of Electronic Commerce 5*, 4 (2001), pp. 135–153.

[101] K. Gruchalla. Immersive Well-Path Editing: Investigating the Added Value of Immersion. In *Proc. of the IEEE Virtual Reality* (2004), VR '04.

[102] J. Gugenheimer, D. Dobbelstein, C. Winkler, G. Haas and E. Rukzio. FaceTouch: Enabling Touch Interaction in Display Fixed UIs for Mobile Virtual Reality. In *Proc. of the 29th Annual Symp. on User Interface Software and Technology* (2016), UIST '16, pp. 49–60.

[103] S. Gustafson, D. Bierwirth and P. Baudisch. Imaginary Interfaces: Spatial Interaction with Empty Hands and Without Visual Feedback. In *Proc. of the 23Nd Annual ACM Symp. on User Interface Software and Technology* (2010), UIST '10, pp. 3–12.

[104] C. Gutwin and A. Cockburn. Improving List Revisitation with ListMaps. In *Proc. of the Working Conf. on Advanced Visual Interfaces* (2006), AVI '06, pp. 396–403.

[105] S. Ha, Y. Bai and C. K. Liu. Human Motion Reconstruction from Force Sensors. In *Proc. of the 2011 ACM SIGGRAPH/Eurographics Symp. on Computer Animation* (2011), SCA '11, pp. 129–138.

[106] V. Ha, J. Wallace, R. Ziola and K. Inkpen. My MDE: Configuring Virtual Workspaces in Multi-display Environments. In *CHI '06 Extended Abstracts on Human Factors in Computing Systems* (2006), CHI EA '06, pp. 1481–1486.

[107] S. G. Hart and L. E. Stavenland. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Human Mental Workload*, P. A. Hancock and N. Meshkati, Eds. Elsevier, 1988, ch. 7, pp. 139–183.

[108] K. Harwood and P. Foley. Temporal Resolution: An Insight into the Video Display Terminal (VDT) Problem. *Human Factors 29*, 4 (1987), pp. 447–452.

[109] G. Häubl and P. Figueroa. Ineractive 3D Presentations and Buyer Behavior. In *CHI '02 Extended Abstracts on Human Factors in Computing Systems* (2002), CHI EA '02, pp. 744–745.

[110] M. A. Hearst. Clustering Versus Faceted Categories for Information Exploration. *Commun. ACM 49*, 4 (Apr. 2006), pp. 59–61.

[111] M. L. Heilig. Sensorama Simulator, 1962. US Patent 3,050,870.

[112] P. J. Hezel and H. Veron. Head Mounted Displays for Virtual Reality. Technical report, Defense Technical Information Center, 1993.

[113] S. Hickson, N. Dufour, A. Sud, V. Kwatra and I. Essa. Eyemotion: Classifying Facial Expressions in VR Using Eye-tracking Cameras, 2017.

[114] K. Hinckley, M. Czerwinski and M. Sinclair. Interaction and Modeling Techniques for Desktop Two-handed Input. In *Proc. of the 11th Annual ACM Symp. on User Interface Software and Technology* (1998), UIST '98, pp. 49–58.

[115] K. Hinckley, R. Pausch, J. C. Goble and N. F. Kassell. Passive Real-world Interface Props for Neurosurgical Visualization. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems* (1994), CHI '94, pp. 452–458.

[116] K. Hinckley, R. Pausch, J. C. Goble and N. F. Kassell. A Survey of Design Issues in Spatial Input. In *Proc. of the 7th Annual ACM Symp. on User Interface Software and Technology* (1994), UIST '94, pp. 213–222.

[117] K. Hinckley, J. Tullio, R. Pausch, D. Proffitt and N. Kassell. Usability Analysis of 3D Rotation Techniques. In *Proc. of the 10th Annual ACM Symp. on User Interface Software and Technology* (1997), UIST '97, pp. 1–10.

[118] D. Hix and H. R. Hartson. *Developing User Interfaces: Ensuring Usability Through Product &Amp; Process*. John Wiley & Sons, Inc., 1993.

[119] E. R. Hoeg, K. V. Ruder, N. C. Nilsson, R. Nordahl and S. Serafin. An Exploration of Input Conditions for Virtual Teleportation. In *Virtual Reality* (2017), IEEE, pp. 341–342.

[120] L. Hoste, B. Dumas and B. Signer. SpeeG: A Multimodal Speech- and Gesture-based Text Input Solution. In *Proc. of the Int. Working Conf. on Advanced Visual Interfaces* (2012), AVI '12, pp. 156–163.

[121] Z. Huang and M. Benyoucef. From e-Commerce to Social Commerce: A Close Look at Design Features. *Electronic Commerce Research and Applications 12*, 4 (2013), pp. 246 – 259.

[122] W. IJsselsteijn, Y. De Kort, K. Poels, A. Jurgelionis and F. Bellotti. Characterising and Measuring User Experiences in Digital Games. In *Int. Conf. on Advances in Computer Entertainment Technology* (2007), vol. 620.

[123] V. Interrante, B. Ries and L. Anderson. Seven League Boots: A New Metaphor for Augmented Locomotion through Moderately Large Scale Immersive Virtual Environments. In *Symp. on 3D User Interfaces* (Charlotte, NC, USA, 2007), IEEE, pp. 167–170.

[124] T. Ishikawa and D. R. Montello. Spatial Knowledge Acquisition from Direct Experience in the Environment: Individual Differences in the Development of Metric Knowledge and the Integration of Separately Learned Places. *Cognitive Psychology 52*, 2 (2006), pp. 93–129.

[125] V. İşler, R. W. Lau and M. Green. Real-time Multi-resolution Modeling for Complex Virtual Environments. In *Proc. of the ACM Symp. on Virtual Reality Software and Technology* (1996), pp. 11–19.

[126] H. Istance, R. Bates, A. Hyrskykari and S. Vickers. Snap Clutch, a Moded Approach to Solving the Midas Touch Problem. In *Proc. of the 2008 Symp. on Eye Tracking Research & Applications* (2008), ETRA '08, pp. 221–228.

[127] R. H. Jacoby, M. Ferneau and J. Humphries. Gestural Interaction in a Virtual Environment. In *IS&T/SPIE 1994 Int. Symp. on Electronic Imaging: Science and Technology* (1994), International Society for Optics and Photonics, pp. 355–364.

[128] S. L. Jarvenpaa and P. A. Todd. Consumer Reactions to Electronic Shopping on the World Wide Web. *Int. J. Electron. Commerce 1*, 2 (Dec. 1996), pp. 59–88.

[129] J. Jerald. Scene-Motion-and Latency-Perception Thresholds for Head-Mounted Displays. Tech. rep., Department of Computer Science, University of North Carolina at Chapel Hill, 2009.

[130] J. Jerald. *The VR Book: Human-Centered Design for Virtual Reality*. ACM, 2016.

[131] J. Jerald, T. Peck, F. Steinicke and M. Whitton. Sensitivity to Scene Motion for Phases of Head Yaws. In *Proc. of the 5th symposium on Applied perception in graphics and visualization - APGV '08* (2008), p. 155.

[132] S. M. Jex. *Stress and Job Performance: Theory, Research, and Implications for Managerial Practice*. Advanced Topics in Organizational Behavior. Sage Publications, Incorporated, 1998.

[133] L. Jiang, R. Girotra, M. R. Cutkosky and C. Ullrich. Reducing Error Rates with Low-cost Haptic Feedback in Virtual Reality-based Training Applications. In *1st Joint Eurohaptics Conf. and Symp. on Haptic Interfaces for Virtual Environment and Teleoperator Systems* (2005), pp. 420–425.

[134] D. M. Johnson. Introduction to and Review of Simulator Sickness Research. Tech. rep., U.S. Army Research Institute for the Behavioral and Social Sciences, 1 2005.

[135] J. Jung, J. Yu and H. Kang. Effects of Virtual Reality Treadmill Training on Balance and Balance Self-efficacy in Stroke Patients with a History of Falling. *Journal of Physical Therapy Science 24*, 11 (2012), pp. 1133–1136.

[136] M. A. Katz and M. D. Byrne. Effects of Scent and Breadth on Use of Site-specific Search on e-Commerce Web Sites. *ACM Trans. Comput.-Hum. Interact. 10*, 3 (2003), pp. 198–220.

[137] K. Kendall and J. Kendall. *Systems Analysis and Design*. Prentice Hall PTR, 2002.

[138] R. Kennedy, J. Fowlkes, K. Berbaum and M. Lilienthal. Use of Motion Sickness History Questionnaire for Prediction of Simulator Sickness. *Aviation, Space, and Environmental Medicine 63* (08 1992), pp. 588–93.

[139] R. Kennedy, L. Frank, M. McCauley, A. Bittner and R. Root. Simulator Sickness: Reaction to a Transformed Perceptual World. VI. Preliminary Site Surveys. Tech. rep., Naval Training Equipment Center Orlando, FL, 11 1984.

[140] R. S. Kennedy, N. E. Lane, K. S. Berbaum and M. G. Lilienthal. Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology 3*, 3 (1993), pp. 203–220.

[141] V. K. Ketoma, P. Schäfer and G. Meixner. Development and Evaluation of a Virtual Reality Grocery Shopping Application Using a Multi-kinect Walking-in-Place Approach. In *Intelligent Human Systems Integration* (2018), Waldemar Karwowski and Tareq Ahram, Eds., pp. 368–374.

[142] M. Kim, M. K. Lee and L. Dabbish. Shop-i: Gaze Based Interaction in the Physical World for In-Store Social Shopping Experience. In *Proc. of the 33rd Annual ACM Conf. Extended Abstracts on Human Factors in Computing Systems* (2015), CHI EA '15, pp. 1253–1258.

[143] Y. Y. Kim, H. J. Kim, E. N. Kim, H. D. Ko and H. T. Kim. Characteristic changes in the Physiological Components of Cybersickness. *Psychophysiology 42*, 5 (2005), pp. 616–625.

[144] R. Kinkead. Typing Speed, Keying Rates, and Optimal Keyboard Layouts. *Proc. of the Human Factors Society Annual Meeting 19*, 2 (1975), pp. 159–161.

[145] Y. Kitamura, S. Ogata and F. Kishino. A Manipulation Environment of Virtual and Real Objects Using a Magnetic Metaphor. In *Proc. of the ACM Symp. on Virtual Reality Software and Technology* (2002), VRST '02, pp. 201–207.

[146] M. Kitazaki, T. Nakano, N. Matsuzaki and H. Shigemasu. Control of Eye-movement to Decrease VE-sickness. In *Proc. of the ACM Symp. on Virtual Reality Software and Technology* (2006), VRST '06, pp. 350–355.

[147] J. M. Knapp and J. M. Loomis. Limited Field of View of Head-Mounted Displays is not the Cause of Distance Underestimation in Virtual Environments. *Presence: Teleoperators and Virtual Environments 13*, 5 (2004), pp. 572–577.

[148] C. Knöpfle and G. Voß. An Intuitive VR User Interface for Design Review. In *Proc. of the Working Conf. on Advanced Visual Interfaces* (2000), AVI '00, pp. 98–101.

[149] R. Kopper, D. A. Bowman, M. G. Silva and R. P. McMahan. A Human Motor Behavior Model for Distal Pointing Tasks. *International Journal of Human-Computer Studies 68*, 10 (2010), pp. 603–615.

[150] F. Kosmalla, A. Zenner, M. Speicher, F. Daiber, N. Herbig and A. Krüger. Exploring Rock Climbing in Mixed Reality Environments. In *Proc. of the 2017 CHI Conf. Extended Abstracts on Human Factors in Computing Systems* (2017), CHI EA '17, pp. 1787–1793.

[151] P. O. Kristensson and K. Vertanen. Performance Comparisons of Phrase Sets and Presentation Styles for Text Entry Evaluations. In *Proc. of the 2012 ACM International Conf. on Intelligent User Interfaces* (2012), IUI '12, pp. 29–32.

[152] A. Kulshreshth and J.J. J. LaViola. Exploring the Usefulness of Finger-based 3D Gesture Menu Selection. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems* (2014), CHI '14, pp. 1093–1102.

[153] A. Kunert, A. Kulik, A. Huckauf and B. Fröhlich. A Comparison of Tracking- and Controller-based Input for Complex Bimanual Interaction in Virtual Environments. In *Proc. of the 13th Eurographics Conf. on Virtual Environments* (2007), EGVE'07, pp. 43–52.

[154] E. Langbehn, P. Lubos, G. Bruder and F. Steinicke. Application of Redirected Walking in Room-scale VR. In *Virtual Reality* (2017), IEEE, pp. 449–450.

[155] E. Langbehn, P. Lubos, G. Bruder and F. Steinicke. Bending the Curve: Sensitivity to Bending of Curved Paths and Application in Room-Scale VR. *IEEE Transactions on Visualization and Computer Graphics 23*, 4 (2017), pp. 1389–1398.

[156] K. Larson and M. Czerwinski. Web Page Design: Implications of Memory, Structure and Scent for Information Retrieval. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems* (1998), CHI '98, pp. 25–32.

[157] B. Laugwitz, T. Held and M. Schrepp. Construction and Evaluation of a User Experience Questionnaire. In *Symp. of the Austrian HCI and Usability Engineering Group* (2008), Springer, pp. 63–76.

[158] S. M. LaValle, A. Yershova, M. Katsev and M. Antonov. Head Tracking for the Oculus Rift. In *2014 IEEE Int. Conf. on Robotics and Automation (ICRA)* (May 2014), pp. 187–194.

[159] K. Laver, F. Lim, K. Reynolds, S. George, J. Ratcliffe, S. Sim and M. Crotty. Virtual Reality Grocery Shopping Simulator: Development and Usability in Neurological Rehabilitation. *Presence 21*, 1 (Feb 2012), pp. 183–191.

[160] J. J. LaViola. Whole-hand and Speech Input in Virtual Environments. Tech. rep., Brown University, Providence, RI, USA, 1999.

[161] J. J. LaViola, E. Kruijff, R. P. McMahan, D. Bowman and I. P. Poupyrev. *3D User Interfaces: Theory and Practice*. Addison Wesley, 2017.

[162] J.J. J. LaViola. A Discussion of Cybersickness in Virtual Environments. *SIGCHI Bull. 32*, 1 (Jan. 2000), pp. 47–56.

[163] J. J. LaViola Jr. Input and Output Devices, 2000.

[164] E. L. Law, F. Brühlmann and E. D. Mekler. Systematic Review and Validation of the Game Experience Questionnaire (GEQ) – Implications for Citation and Reporting Practice, Aug 2018.

[165] B. Lawson and A. Mead. The Sopite Syndrome Revisited: Drowsiness and Mood Changes during Real or Apparent Motion. *Acta Astronautica 43*, 3 (1998), pp. 181–192.

[166] J. Lazar. *User-Centered Web Development: Theory into Practice*, 2nd ed. Jones and Bartlett Publishers, Inc., USA, 2003.

[167] A. Lecuyer. Simulating Haptic Feedback Using Vision: A Survey of Research and Applications of Pseudo-Haptic Feedback. *Presence: Teleoperators and Virtual Environments 18*, 1 (2009), pp. 39–53.

[168] J.-Y. Lee, P.-H. Han, L. Tsai, R.-D. Peng, Y.-S. Chen, K.-W. Chen and Y.-P. Hung. Estimating the Simulator Sickness in Immersive Virtual Reality with Optical Flow Analysis. In *SIGGRAPH Asia 2017 Posters* (2017), SA '17, pp. 16:1–16:2.

[169] K. C. Lee and N. Chung. Empirical Analysis of Consumer Reaction to the Virtual Reality Shopping Mall. *Computers in Human Behavior 24*, 1 (2008), pp. 88–104.

[170] P.-W. Lee, H.-Y. Wang, Y.-C. Tung, J.-W. Lin and A. Valstar. TranSection: Hand-Based Interaction for Playing a Game Within a Virtual Reality Game. In *Proc. of the 33rd Annual ACM Conf. Extended Abstracts on Human Factors in Computing Systems* (2015), CHI EA '15, pp. 73–76.

[171] V. I. Levenshtein. Binary Codes Capable of Correcting Deletions, Insertions, and Reversals. In *Soviet Physics Doklady* (1966), vol. 10, pp. 707–710.

[172] C. J. Lin, S.-H. Ho and Y.-J. Chen. An Investigation of Pointing Postures in a 3D Stereoscopic Environment. *Applied Ergonomics 48* (2015), pp. 154–163.

[173] J. J. W. Lin, H. Abi-Rached and M. Lahav. Virtual Guiding Avatar: An Effective Procedure to Reduce Simulator Sickness in Virtual Environments. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems* (2004), CHI '04, pp. 719–726.

[174] J. J. W. Lin, H. B.-L. Duh, D. E. Parker, H. Abi-Rached and T. A. Furness. Effects of Field of View on Presence, Enjoyment, Memory, and Simulator Sickness in a Virtual Environment. In *Virtual Reality* (2002), IEEE, pp. 164–171.

[175] J.-W. Lin, H. B. L. Duh, D. E. Parker, H. Abi-Rached and T. A. Furness. Effects of Field of View on Presence, Enjoyment, Memory, and Simulator Sickness in a Virtual Environment. In *Proc. IEEE Virtual Reality* (March 2002), pp. 164–171.

[176] J.-W. Lin, P.-H. Han, J.-Y. Lee, Y.-S. Chen, T.-W. Chang, K.-W. Chen and Y.-P. Hung. Visualizing the Keyboard in Virtual Reality for Enhancing Immersive Experience. In *ACM SIGGRAPH 2017 Posters* (2017), SIGGRAPH '17, pp. 35:1–35:2.

[177] M. Liu. A Study of Mobile Sensing Using Smartphones. *Int. Journal of Distributed Sensor Networks 2013* (2013).

[178] G. Llorach, A. Evans and J. Blat. Simulator Sickness and Presence Using HMDs: Comparing Use of a Game Controller and a Position Estimation System. In *Proc. of the 20th ACM Symp. on Virtual Reality Software and Technology* (2014), VRST '14, pp. 137–140.

[179] G. L. Lohse and P. Spiller. Quantifying the Effect of User Interface Design Features on Cyberstore Traffic and Sales. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems* (1998), CHI '98, pp. 211–218.

[180] C. Lu and J. Tanaka. A Virtual Shopping System Based on Room-scale Virtual Reality. In *Proc. of the 12th Int. Conf. on Advances in Computer-Human Interactions* (2019), ACHI '19.

[181] P. Lubos, G. Bruder and F. Steinicke. Influence of Comfort on 3D Selection Task Performance in Immersive Desktop Setups. *Journal of Virtual Reality and Broadcasting (JVRB) 12*, 2 (2015).

[182] K. Lyons, T. Starner, D. Plaisted, J. Fusia, A. Lyons, A. Drew and E. W. Looney. Twiddler Typing: One-handed Chording Text Entry for Mobile Phones. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems* (2004), CHI '04, pp. 671–678.

[183] R. L. Mack and J. Nielsen. Usability Inspection Methods: Executive Summary. In *Readings in Human–Computer Interaction*, Interactive Technologies. Morgan Kaufmann, 1995, pp. 170–181.

[184] I. S. MacKenzie. KSPC (Keystrokes Per Character) As a Characteristic of Text Entry Techniques. In *Proc. of the 4th Int. Symp. on Mobile Human-Computer Interaction* (2002), Mobile HCI '02, pp. 195–210.

[185] I. S. MacKenzie. *Human-Computer Interaction: An Empirical Research Perspective*, 1st ed. Morgan Kaufmann Publishers Inc., 2013.

[186] I. S. MacKenzie and K. Tanaka-Ishii. *Text Entry Systems: Mobility, Accessibility, Universality*. Morgan Kaufmann Publishers Inc., 2007.

[187] I. S. MacKenzie and S. X. Zhang. The Design and Evaluation of a High-performance Soft Keyboard. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems* (1999), CHI '99, pp. 25–31.

[188] A. MacLean, R. M. Young, V. M. E. Bellotti and T. P. Moran. Questions, Options, and Criteria: Elements of Design Space Analysis. *Human-Computer Interaction 6*, 3 (Sept. 1991), pp. 201–250.

[189] C. Madu and A. Madu. Dimensions of e-Quality. *International Journal of Quality and Reliability Management 19*, 3 (2002), pp. 246–258.

[190] P. Majaranta and K.-J. Räihä. Text Entry by Gaze: Utilizing Eye-tracking. *Text Entry Systems: Mobility, Accessibility, Universality* (2007), pp. 175–187.

[191] S. Manjrekar, S. Sandilya, D. Bhosale, S. Kanchi, A. Pitkar and M. Gondhalekar. CAVE: An Emerging Immersive Technology – A Review. In *2014 UKSim-AMSS 16th International Conf. on Computer Modelling and Simulation* (March 2014), pp. 131–136.

[192] A. Markussen, M. R. Jakobsen and K. Hornbæk. Selection-based Mid-air Text Entry on Large Displays. In *IFIP Conf. on Human-Computer Interaction* (2013), pp. 401–418.

[193] A. Martinet, G. Casiez and L. Grisoni. The Effect of DOF Separation in 3D Manipulation Tasks with Multi-touch Displays. In *Proc. of the 17th ACM Symp. on Virtual Reality Software and Technology* (2010), VRST '10, pp. 111–118.

[194] M. McGill, D. Boland, R. Murray-Smith and S. Brewster. A Dose of Reality: Overcoming Usability Challenges in VR Head-Mounted Displays. In *Proc. of the 33rd Annual ACM Conf. on Human Factors in Computing Systems - CHI '15* (Seoul, Republic of Korea, 2015), pp. 2143–2152.

[195] M. McKenna. Interactive Viewpoint Control and Three-dimensional Operations. In *Proc. of the 1992 Symp. on Interactive 3D Graphics* (1992), I3D '92, pp. 53–56.

[196] R. P. McMahan, D. Gorton, J. Gresock, W. McConnell and D. A. Bowman. Separating the Effects of Level of Immersion and 3D Interaction Techniques. In *Proc. of the ACM Symp. on Virtual Reality Software and Technology* (2006), VRST '06, pp. 108–111.

[197] D. Medeiros, E. Cordeiro, D. Mendes, M. Sousa, A. Raposo, A. Ferreira and J. Jorge. Effects of Speed and Transitions on Target-based Travel Techniques. In *Proc. of the 22Nd ACM Conf. on Virtual Reality Software and Technology* (2016), VRST '16, pp. 327–328.

[198] E. Medina, R. Fruland and S. Weghorst. Virtusphere: Walking in a Human Size VR "Hamster Ball". *Proc. of the Human Factors and Ergonomics Society Annual Meeting 52*, 27 (2008), pp. 2102–2106.

[199] M. Meehan, B. Insko, M. Whitton and J.F. P. Brooks. Physiological Measures of Presence in Stressful Virtual Environments. In *Proc. of the 29th Annual Conf. on Computer Graphics and Interactive Techniques* (2002), SIGGRAPH '02, pp. 645–652.

[200] A. Meschtscherjakov, W. Reitberger, M. Lankes and M. Tscheligi. Enhanced Shopping: A Dynamic Map in a Retail Store. In *Proc. of the 10th International Conf. on Ubiquitous Computing* (2008), UbiComp '08, pp. 336–339.

[201] K. Meyer, H. L. Applewhite and F. Biocca. A Survey of Position Trackers. *Presence 1* (01 1992), pp. 173–200.

[202] M. Miksch, M. Steiner, M. Miksch and A. Meschtscherjakov. Motion Sickness Prevention System (MSPS): Reading Between the Lines. In *Adjunct Proc. of the 8th Int. Conf. on Automotive User Interfaces and Interactive Vehicular Applications* (2016), AutomotiveUI '16 Adjunct, pp. 147–152.

[203] P. Milgram, H. Takemura, A. Utsumi and F. Kishino. Augmented reality: A class of displays on the reality-virtuality continuum. In *Photonics for Industrial Applications* (1995), International Society for Optics and Photonics, pp. 282–292.

[204] D. P. Miller. The Depth/Breadth Tradeoff in Hierarchical Computer Menus. *Proc. of the Human Factors Society Annual Meeting 25*, 1 (1981), pp. 296–300.

[205] M. R. Mine. Virtual Environment Interaction Techniques. Tech. rep., Department of Computer Science, University of North Carolina, 1995.

[206] M. R. Mine, J.F. P. Brooks and C. H. Sequin. Moving Objects in Space: Exploiting Proprioception in Virtual-environment Interaction. In *Proc. of the 24th Annual Conf. on Computer Graphics and Interactive Techniques* (1997), SIGGRAPH '97, pp. 19–26.

[207] J. Mitchell and B. Shneiderman. Dynamic Versus Static Menus: An Exploratory Comparison. *SIGCHI Bull. 20*, 4 (1989), pp. 33–37.

[208] S. Moagar-Poladian, G.-C. Dumitrescu and I. A. Tănase. Retail e-Commerce (E-tail) – evolution, characteristics and perspectives in China, the USA and Europe. *Global Economic Observer 5*, 1 (June 2017).

[209] M. Mori, K. MacDorman and N. Kageki. The Uncanny Valley [From the Field]. *IEEE Robotics & Automation Magazine 19*, 2 (2012), pp. 98–100.

[210] M. R. Morris, K. Inkpen and G. Venolia. Remote Shopping Advice: Enhancing In-store Shopping with Social Technologies. In *Proc. of the 17th ACM Conf. on Computer Supported Cooperative Work & Social Computing* (2014), CSCW '14, pp. 662–673.

[211] R. R. Mourant and T. R. Thattacherry. Simulator Sickness in a Virtual Environments Driving Simulator. In *Proc. of the Human Factors and Ergonomics Society Annual Meeting* (2000), vol. 44, pp. 534–537.

[212] P. Murano and M. Sander. User Interface Menu Design Performance and User Preferences: A Review and Ways Forward. *International Journal of Advanced Computer Science and Applications 7*, 4 (2016), pp. 355–361.

[213] E. R. Muth, R. M. Stern, J. F. Thayer and K. L. Koch. Assessment of the Multiple Dimensions of Nausea: The Nausea Profile (NP). *Journal of Psychosomatic Research 40*, 5 (1996), pp. 511–520.

[214] J. Mütterlein and T. Hess. Immersion, Presence, Interactivity: Towards a Joint Understanding of Factors Influencing Virtual Reality Acceptance and Use. In *AMCIS2017 Proc.: Adoption and Diffusion of Information Technology (SIGADIT)* (2017).

[215] M. Nancel, J. Wagner, E. Pietriga, O. Chapuis and W. Mackay. Mid-air Pan-and-zoom on Wall-sized Displays. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems* (2011), CHI '11, pp. 177–186.

[216] M. Narayan, L. Waugh, X. Zhang, P. Bafna and D. Bowman. Quantifying the Benefits of Immersion for Collaboration in Virtual Environments. In *Proc. of the ACM Symp. on Virtual Reality Software and Technology* (2005), VRST '05, pp. 78–81.

[217] A. Nathan and S. Gao. Interactive Displays: The Next Omnipresent Technology [Point of View]. *Proc. of the IEEE 104*, 8 (Aug 2016), pp. 1503–1507.

[218] F. Nemetz. The Missing Link: Multimedia and e-Commerce. In *CHI '00 Extended Abstracts on Human Factors in Computing Systems* (2000), CHI EA '00, pp. 65–66.

[219] C. T. Neth, J. L. Souman, D. Engel, U. Kloos, H. H. Bulthoff and B. J. Mohler. Velocity-dependent Dynamic Curvature Gain for Redirected Walking. *IEEE transactions on visualization and computer graphics 18*, 7 (2012), pp. 1041–1052.

[220] D. C. Niehorster, L. Li and M. Lappe. The Accuracy and Precision of Position and Orientation Tracking in the HTC Vive Virtual Reality System for Scientific Research. *i-Perception 8*, 3 (2017).

[221] N. C. Nilsson, T. Peck, G. Bruder, E. Hodgson, S. Serafin, M. Whitton, F. Steinicke and E. S. Rosenberg. 15 Years of Research on Redirected Walking in Immersive Virtual Environments. *IEEE Computer Graphics and Applications 38*, 2 (2018), pp. 44–56.

[222] N. C. Nilsson, S. Serafin, F. Steinicke and R. Nordahl. Natural Walking in Virtual Reality: A Review. *Comput. Entertain. 16*, 2 (Apr. 2018), pp. 8:1–8:22.

[223] D. A. Norman. Cognitive Engineering. *User-centered System Design 31* (1986), p. 61.

[224] D. A. Norman and S. W. Draper. *User-centered System Design: New Perspectives on Human-computer Interaction*. CRC Press, 1986.

[225] K. L. Norman. Better Design of Menu Selection Systems Through Cognitive Psychology and Human Factors. *Human Factors 50*, 3 (2008), pp. 556–559.

[226] J. Novak-Marcincin. Hardware Devices used in Virtual Reality Technologies. *Scientific Bulletin Series C 24*, 1 (2010).

[227] J. Noyes. The QWERTY Keyboard: A Review. *International Journal of Man-Machine Studies 18*, 3 (1983), pp. 265–281.

[228] H. Ogier and J. Buchan. Exploring the Feasibility of Diegetic In-game Store User Interfaces. In *Proc. of the Australasian Computer Science Week Multiconference* (2017), ACSW '17, pp. 67:1–67:10.

[229] J. K. Parker, R. L. Mandryk, M. N. Nunes and K. M. Inkpen. TractorBeam Selection Aids: Improving Target Acquisition for Pointing Input on Tabletop Displays. In *Proc. of the 2005 IFIP TC13 Int. Conf. on Human-Computer Interaction* (2005), INTERACT'05, pp. 80–93.

[230] A. Patney, J. Kim, M. Salvi, A. Kaplanyan, C. Wyman, N. Benty, A. Lefohn and D. Luebke. Perceptually-based Foveated Virtual Reality. In *ACM SIGGRAPH 2016 Emerging Technologies* (2016), SIGGRAPH '16, pp. 17:1–17:2.

[231] R. Patterson, M. D. Winterbottom and B. J. Pierce. Perceptual Issues in the Use of Head-Mounted Visual Displays. *Human Factors 48*, 3 (2006), pp. 555–573.

[232] R. Pausch, T. Crea and M. Conway. A Literature Survey for Virtual Environments: Military Flight Simulator Visual Systems and Simulator Sickness. *Presence: Teleoperators and Virtual Environments 1*, 3 (1992), pp. 344–363.

[233] R. Pausch, J. Snoddy, R. Taylor, S. Watson and E. Haseltine. Disney's Aladdin: First Steps Toward Storytelling in Virtual Reality. In *Proc. of the 23rd Annual Conf. on Computer Graphics and Interactive Techniques* (1996), SIGGRAPH '96, pp. 193–203.

[234] T. C. Peck, H. Fuchs and M. C. Whitton. Evaluation of Reorientation Techniques and Distractors for Walking in Large Virtual Environments. *IEEE Transactions on Visualization and Computer Graphics 15*, 3 (May 2009), pp. 383–394.

[235] G. Petz and A. Greiner. First in Search – How to Optimize Search Results in E-Commerce Web Shops. In *HCI in Business* (2014), Fiona Fui-Hoon Nah, Ed., pp. 566–574.

[236] J. S. Pierce, A. S. Forsberg, M. J. Conway, S. Hong, R. C. Zeleznik and M. R. Mine. Image Plane Interaction Techniques in 3D Immersive Environments. In *Proc. of the 1997 Symp. on Interactive 3D Graphics* (1997), I3D '97, pp. 39–ff.

[237] K. Pietroszek, A. Kuzminykh, J. R. Wallace and E. Lank. Smartcasting: A Discount 3D Interaction Technique for Public Displays. In *Proc. of the 26th Australian Computer-Human Interaction Conf. on Designing Futures: The Future of Design* (2014), OzCHI '14, pp. 119–128.

[238] K. Pimentel. *Virtual Reality: Through the New Looking Glass*, 2nd ed. McGraw-Hill, Inc., 2000.

[239] M. Plechawska-Wojcik. *Heuristic Evaluation and User Experience Assessment of Online Shopping Portals Using Cognitive Walkthrough and Expert Method*. Human Capital without Borders: Knowledge and Learning for Quality of Life; Proc. of the Management, Knowledge and Learning International Conf. 2014. ToKnowPress, 2014, pp. 467–475.

[240] P. G. Polson, C. Lewis, J. Rieman and C. Wharton. Cognitive Walkthroughs: A Method for Theory-based Evaluation of User Interfaces. *Int. J. Man-Mach. Stud. 36*, 5 (May 1992), pp. 741–773.

[241] I. Poupyrev, M. Billinghurst, S. Weghorst and T. Ichikawa. The Go-go Interaction Technique: Non-linear Mapping for Direct Manipulation in VR. In *Proc. of the 9th Annual ACM Symp. on User Interface Software and Technology* (Seattle, Washington, United States, 1996), UIST '96, pp. 79–80.

[242] I. Poupyrev, T. Ichikawa, S. Weghorst and M. Billinghurst. Egocentric Object Manipulation in Virtual Environments: Empirical Evaluation of Interaction Techniques. In *Computer Graphics Forum* (1998), vol. 17, Wiley Online Library, pp. 41–52.

[243] I. Poupyrev, S. Weghorst and S. Fels. Non-isomorphic 3D Rotational Techniques. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems* (2000), pp. 540–547.

[244] H. Putnam. Brains in a Vat. *Knowledge: Critical Concepts 1* (1981), pp. 192–207.

[245] S. Razzaque. *Redirected Walking*. PhD thesis, University of North Carolina at Chapel Hill, 2005.

[246] J. T. Reason and J. J. Brand. *Motion Sickness.* Academic press, 1975.

[247] F. Rebelo, P. Noriega, E. Duarte and M. Soares. Using Virtual Reality to Assess User Experience. *Human Factors 54*, 6 (2012), pp. 964–982.

[248] G. Ren and E. O'Neill. Freehand Gestural Text Entry for Interactive TV. In *Proc. of the 11th European Conf. on Interactive TV and Video* (2013), EuroITV '13, pp. 121–130.

[249] M. L. Resnick and J. Sanchez. Effects of Organizational Scheme and Labeling on Task Performance in Product-Centered and User-Centered Retail Web Sites. *Human Factors 46*, 1 (2004), pp. 104–117.

[250] M. Ribo, A. Pinz and A. L. Fuhrmann. A New Optical Tracking System for Virtual and Augmented Reality Applications. In *IMTC 2001. Proc. of the 18th IEEE Instrumentation and Measurement Technology Conf. Rediscovering Measurement in the Age of Informatics)* (May 2001), vol. 3, pp. 1932–1936.

[251] G. E. Riccio and T. A. Stoffregen. An Ecological Theory of Motion Sickness and Postural Instability. *Ecological psychology 3*, 3 (1991), pp. 195–240.

[252] A. S. Rizzo, B. Lange, E. A. Suma and M. Bolas. Virtual Reality and Interactive Digital Game Technology: New Tools to Address Obesity and Diabetes. *Journal of Diabetes Science and Technology 5*, 2 (2011), pp. 256–264.

[253] G. Robertson, M. Czerwinski, K. Larson, D. C. Robbins, D. Thiel and M. van Dantzich. Data Mountain: Using Spatial Memory for Document Management. In *Proc. of the 11th Annual ACM Symp. on User Interface Software and Technology* (1998), UIST '98, pp. 153–162.

[254] G. Robertson, M. Czerwinski and M. van Dantzich. Immersion in Desktop Virtual Reality. In *Proc. of the 10th Annual ACM Symp. on User Interface Software and Technology* (1997), UIST '97, pp. 11–19.

[255] R. M. Ryan, C. S. Rigby and A. Przybylski. The Motivational Pull of Video Games: A Self-Determination Theory Approach. *Motivation and Emotion 30*, 4 (Dec 2006), pp. 344–360.

[256] S. Sackstein, L. Spark and A. Jenkins. Are e-Books Effective Tools for Learning? Reading Speed and Comprehension: iPad vs. Paper. *South African Journal of Education 35* (2015), pp. 01–14.

[257] A. Sanna, B. Montrucchio, P. Montuschi and C. Demartini. 3D-DVShop: A 3D Dynamic Virtual Shop. In *Multimedia 2001*. Springer, 2002, pp. 33–42.

[258] J. Sauro. Measuring Usability with the System Usability Scale (SUS), 2011.

[259] J. Scarr, A. Cockburn, C. Gutwin and A. Bunt. Improving Command Selection with CommandMaps. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems* (2012), CHI '12, pp. 257–266.

[260] J. Schalkwyk, D. Beeferman, F. Beaufays, B. Byrne, C. Chelba, M. Cohen, M. Kamvar and B. Strope. "Your Word is My Command": Google Search by Boice: A Case Study. In *Advances in Speech Recognition*. Springer, 2010, pp. 61–90.

[261] M. Schrepp, A. Hinderks and J. Thomaschewski. Construction of a Benchmark for the User Experience Questionnaire (UEQ). *International Journal of Interactive Multimedia and Artificial Intelligence 4*, 4 (06 2017), pp. 40–44.

[262] M. J. Schuemie, P. van der Straaten, M. Krijn and C. A. van der Mast. Research on Presence in Virtual Reality: A Survey. *CyberPsychology & Behavior 4*, 2 (2001), pp. 183–201.

[263] M. Schuster. Speech Recognition for Mobile Devices at Google. In *Pacific Rim International Conf. on Artificial Intelligence* (2010), pp. 8–10.

[264] J. P. Schwartz and K. L. Norman. The importance of item distinctiveness on performance using a menu selection system. *Behaviour & Information Technology 5*, 2 (1986), pp. 173–182.

[265] M. Scriven. *The Methodology of Evaluation*. Purdue University, 1966.

[266] A. Sears and B. Shneiderman. Split Menus: Effectively Using Selection Frequency to Organize Menus. *ACM Trans. Comput.-Hum. Interact. 1*, 1 (1994), pp. 27–51.

[267] A. F. Seay, D. M. Krum, L. Hodges and W. Ribarsky. Simulator Sickness and Presence in a High Field-of-view Virtual Environment. In *CHI '02 Extended Abstracts on Human Factors in Computing Systems* (2002), CHI EA '02, pp. 784–785.

[268] S. Sharples, S. Cobb, A. Moody and J. R. Wilson. Virtual Reality Induced Symptoms and Effects (VRISE): Comparison of Head Mounted Display (HMD), Desktop and Projection Display Systems. *Displays 29*, 2 (2008), pp. 58–69.

[269] S. Shim, M. A. Eastlick, S. L. Lotz and P. Warrington. An Online Prepurchase Intentions Model: The Role of Intention to Search. *Journal of Retailing 77*, 3 (2001), pp. 397–416.

[270] B. Shneiderman. *Designing the User Interface: Strategies for Effective Human-Computer Interaction*, 3rd ed. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1997.

[271] B. Shneiderman. Direct Manipulation for Comprehensible, Predictable and Controllable User Interfaces. In *Proc. of the 2nd Int. Conf. on Intelligent User Interfaces* (1997), IUI '97, pp. 33–39.

[272] G. Shoemaker, L. Findlater, J. Q. Dawson and K. S. Booth. Mid-air Text Input Techniques for Very Large Wall Displays. In *Proc. of Graphics Interface 2009* (2009), GI '09, pp. 231–238.

[273] L. Shupp, R. Ball, B. Yost, J. Booker and C. North. Evaluation of Viewport Size and Curvature of Large, High-resolution Displays. In *Proc. of Graphics Interface* (2006), GI '06, pp. 123–130.

[274] A. L. Simeone, E. Velloso and H. Gellersen. Substitutional Reality: Using the Physical Environment to Design Virtual Reality Experiences. In *Proc. of the 33rd Annual ACM Conf. on Human Factors in Computing Systems - CHI '15* (Seoul, Republic of Korea, 2015), pp. 3307–3316.

[275] E. Sinitski, A. Thompson, P. Godsell, J. Honey and M. Besemann. Postural Stability and Simulator Sickness after Walking on a Treadmill in a Virtual Environment with a Curved Display. *Displays 52* (2018), pp. 1 – 7.

[276] M. Slater. Measuring Presence: A Response to the Witmer and Singer Presence Questionnaire. *Presence: Teleoper. Virtual Environ. 8*, 5 (Oct. 1999), pp. 560–565.

[277] M. Slater. A Note on Presence Terminology. *Presence connect 3*, 3 (2003), pp. 1–5.

[278] M. Slater, M. Usoh and A. Steed. Depth of Presence in Virtual Environments. *Presence: Teleoper. Virtual Environ. 3*, 2 (1994), pp. 130–144.

[279] M. Slater, M. Usoh and A. Steed. Taking Steps: The Influence of a Walking Technique on Presence in Virtual Reality. *ACM Transactions on Computer-Human Interaction 2*, 3 (1995), pp. 201–219.

[280] R. H. Spector. Visual Fields. In *Clinical Methods: The History, Physical, and Laboratory Examinations (3rd Ed.)* (1990), HK Walker; W. Hall; and J. Hurst, Eds.

[281] M. Speicher. Shopping in Virtual Reality. In *2018 IEEE Conf. on Virtual Reality and 3D User Interfaces (VR)* (March 2018), pp. 1–2.

[282] M. Speicher, S. Cucerca and A. Krüger. VRShop: A Mobile Interactive Virtual Reality Shopping Environment Combining the Benefits of On- and Offline Shopping. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 1*, 3 (Sept. 2017), pp. 102:1–102:31.

[283] M. Speicher, F. Daiber, S. Gehring and A. Krüger. Exploring 3D Manipulation on Large Stereoscopic Displays. In *Proc. of the 5th ACM International Symp. on Pervasive Displays* (2016), PerDis '16, pp. 59–66.

[284] M. Speicher, F. Daiber, G.-L. Kiefer and A. Krüger. Exploring Task Performance and User's Preference of Mid-air Hand Interaction in a 3D Docking Task Experiment. In *Proc. of the 5th Symp. on Spatial User Interaction* (2017), SUI '17, pp. 160–160.

[285] M. Speicher, J. Ehrlich, V. Gentile, D. Degraen, S. Sorce and A. Krüger. Pseudo-haptic Controls for Mid-air Finger-based Menu Interaction. In *Proc. of the 2019 CHI Conf. Extended Abstracts on Human Factors in Computing Systems* (2019), CHI EA '19, pp. 1787–1793.

[286] M. Speicher, A. M. Feit, P. Ziegler and A. Krüger. Selection-based Text Entry in Virtual Reality. In *Proc. of the 2018 CHI Conf. on Human Factors in Computing Systems* (2018), CHI '18, pp. 647:1–647:13.

[287] M. Speicher, P. Hell, F. Daiber, A. Simeone and A. Krüger. A Virtual Reality Shopping Experience Using the Apartment Metaphor. In *Proc. of the 2018 International Conf. on Advanced Visual Interfaces* (2018), AVI '18, pp. 17:1–17:9.

[288] M. Speicher, C. Rosenberg, D. Donald, F. Daiber and A. Krüger. Exploring Visual Guidance in 360-degree Videos. In *Proc. of the 2019 ACM Int. Conf. on Interactive Experiences for TV and Online Video* (2019), TVX '19, pp. 131–133.

[289] M. Speicher, N. Rutsch, A. Krüger and M. Löchtefeld. Exploring Spatial Menu Representation and Apartment-based Categorization for Online Shopping. In *Human-Computer Interaction* (2019), INTERACT '19.

[290] M. Speicher, R. Siegel and A. Krüger. ProductFinder: A Location Aware Product Information Display for Retail Environments. In *Proc. of the 6th ACM International Symp. on Pervasive Displays* (2017), PerDis '17, pp. 23:1–23:2.

[291] S. Sridhar, A. M. Feit, C. Theobalt and A. Oulasvirta. Investigating the Dexterity of Multi-Finger Input for Mid-Air Text Entry. In *Proc. of the 33rd Annual ACM Conf. on Human Factors in Computing Systems* (2015), CHI '15, pp. 3643–3652.

[292] K. M. Stanney, K. S. Hale, I. Nahmens and R. S. Kennedy. What to Expect from Immersive Virtual Environment Exposure: Influences of Gender, Body Mass Index, and Past Experience. *Human Factors 45*, 3 (2003), pp. 504–520.

[293] F. Steinicke, G. Bruder, J. Jerald, H. Frenz and M. Lappe. Estimation of Detection Thresholds for Redirected Walking Techniques. *IEEE Transactions on Visualization and Computer Graphics 16*, 1 (2010), pp. 17–27.

[294] F. Steinicke, K. H. Hinrichs, J. Schöning and A. Krüger. Multi-touching 3D Data: Towards Direct Interaction in Stereoscopic Display Environments Coupled with Mobile Devices. In *Proc. AVI Workshop on Designing Multi-Touch Interaction Techniques for Coupled Public and Private Displays* (2008), pp. 46–49.

[295] J. Steuer. Defining Virtual Reality: Dimensions Determining Telepresence. In *Communication in the Age of Virtual Reality*, Frank Biocca and Mark R. Levy, Eds. L. Erlbaum Associates Inc., 1995, pp. 33–56.

[296] R. Stoakley, M. J. Conway and R. Pausch. Virtual Reality on a WIM: Interactive Worlds in Miniature. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems* (1995), CHI '95, pp. 265–272.

[297] J. Stott. Mechanisms and Treatment of Motion Illness. In *Nausea and vomiting: Mechanisms and Treatment*. Springer, 1986, pp. 110–129.

[298] R. Stuart. The Design of Virtual Environments. *McGraw-Hill, New York* (11 1996).

[299] E. A. Suma, G. Bruder, F. Steinicke, D. M. Krum and M. Bolas. A Taxonomy for Deploying Redirection Techniques in Immersive Virtual Environments. In *Virtual Reality* (Costa Mesa, CA, USA, 2012), IEEE, pp. 43–46.

[300] I. E. Sutherland. Sketch Pad: A Man-Machine Graphical Communication System. In *Proc. of the SHARE Design Automation Workshop* (1964), pp. 6–329.

[301] I. E. Sutherland. The Ultimate Display. *Multimedia: From Wagner to Virtual Reality* (1965), pp. 506–508.

[302] C. Swindells, A. Unden and T. Sang. TorqueBAR: An Ungrounded Haptic Feedback Device. In *Proc. of the 5th Int. Conf. on Multimodal Interfaces* (2003), ICMI '03, pp. 52–59.

[303] D. Tan and M. Czerwinski. Effects of Visual Separation and Physical Discontinuities when Distributing Information across Multiple Displays. In *Proc. of Interact 2003* (November 2003), INTERACT '03, pp. 252–255.

[304] N. Tanaka and H. Takagi. Virtual Reality Environment Design of Managing Both Presence and Virtual Reality Sickness. *Journal of Physiological Anthropology and Applied Human Science 23*, 6 (2004), pp. 313–317.

[305] R. Teather and W. Stuerzlinger. Target Pointing in 3D User Interfaces. *CEUR Workshop Proc. 588* (1 2010).

[306] J. Thomaschewski, A. Hinderks and M. Schrepp. Welcher UX-Fragebogen passt zu meinem Produkt? In *Mensch und Computer 2018 - Usability Professionals* (2018), Steffen Hess and Holger Fischer, Eds., Gesellschaft für Informatik e.V. Und German UPA e.V., pp. 437–446.

[307] R. Toth, J. Nilsson and T. Akenine-Möller. Comparison of Projection Methods for Rendering Virtual Reality. In *Proc. of High Performance Graphics* (2016), pp. 163–171.

[308] M. Treisman. Motion Sickness: An Evolutionary Hypothesis. *Science 197*, 4302 (1977), pp. 493–495.

[309] T. Tsandilas and m. c. schraefel. Bubbling Menus: A Selective Mechanism for Accessing Hierarchical Drop-down Menus. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems* (2007), CHI '07, pp. 1195–1204.

[310] A. N. Tuch, S. P. Roth, K. HornbæK, K. Opwis and J. A. Bargas-Avila. Is Beautiful Really Usable? Toward Understanding the Relation Between Usability, Aesthetics, and Affect in HCI. *Comput. Hum. Behav. 28*, 5 (2012), pp. 1596–1607.

[311] M. Usoh, K. Arthur, M. C. Whitton, R. Bastos, A. Steed, M. Slater and J.F. P. Brooks. Walking > Walking-in-place > Flying, in Virtual Environments. In *Proc. of the 26th Annual Conf. on Computer Graphics and Interactive Techniques* (1999), SIGGRAPH '99, pp. 359–364.

[312] K. Vertanen and P. O. Kristensson. A Versatile Dataset for Text Entry Evaluations Based on Genuine Mobile Emails. In *Proc. of the 13th Int. Conf. on Human Computer Interaction with Mobile Devices and Services* (2011), MobileHCI '11, pp. 295–298.

[313] A. von Kapri, T. Rick and S. Feiner. Comparing Steering-based Travel Techniques for Search Tasks in a CAVE. In *Proc. of the 2011 IEEE Virtual Reality Conf.* (2011), VR '11, pp. 91–94.

[314] S. von Mammen, A. Knote and S. Edenhofer. Cyber Sick but Still Having Fun. In *Proc. of the 22Nd ACM Conf. on Virtual Reality Software and Technology* (2016), VRST '16, pp. 325–326.

[315] A. P. Vrechopoulos, R. M. O'Keefe, G. I. Doukidis and G. J. Siomkos. Virtual Store Layout: An Experimental Comparison in the Context of Grocery Retail. *Journal of Retailing 80*, 1 (2004), pp. 13–22.

[316] K. Vredenburg, J.-Y. Mao, P. W. Smith and T. Carey. A Survey of User-centered Design Practice. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems* (2002), CHI '02, pp. 471–478.

[317] J. Walker, B. Li, K. Vertanen and S. Kuhl. Efficient Typing on a Visually Occluded Physical Keyboard. In *Proc. of the 2017 CHI Conf. on Human Factors in Computing Systems* (2017), CHI '17, pp. 5457–5461.

[318] F. Waller. The Archeology of Cinerama. *Film History 5*, 3 (1993), pp. 289–297.

[319] K. R. Walsh and S. D. Pawlowski. Virtual Reality: A Technology in Need of IS Research. *Communications of the Association for Information Systems 8*, 1 (2002), p. 20.

[320] D. Wan. Magic Home: Exploiting the Duality Between the Physical and the Virtual Worlds. In *CHI '00 Extended Abstracts on Human Factors in Computing Systems* (2000), CHI EA '00, pp. 53–54.

[321] Y. Wang, C. L. Mackenzie, V. A. Summers and K. S. Booth. The Structure of Object Transportation and Orientation in Human-Computer Interaction. In *Proc. of ACM CHI'98 Conf. on Human Factors in Computing Systems* (1998), pp. 312–319.

[322] C. Ware, K. Arthur and K. S. Booth. Fish Tank Virtual Reality. In *Proc. of the INTERACT and CHI Conf. on Human Factors in Computing Systems* (1993), CHI '93, pp. 37–42.

[323] F. Weichert, D. Bachmann, B. Rudak and D. Fisseler. Analysis of the Accuracy and Robustness of the Leap Motion Controller. *Sensors 13*, 5 (2013), pp. 6380–6393.

[324] S. Werner, B. Krieg-Brückner, H. A. Mallot, K. Schweizer and C. Freksa. Spatial Cognition: The Role of Landmark, Route, and Survey Knowledge in Human and Robot Navigation. In *Informatik als Innovationsmotor: 27. Jahrestagung der Gesellschaft für Informatik* (1997), pp. 41–50.

[325] E. Whitmire, H. Benko, C. Holz, E. Ofek and M. Sinclair. Haptic Revolver: Touch, Shear, Texture, and Shape Rendering on a Reconfigurable Virtual Reality Controller. In *Proc. of the 2018 CHI Conf. on Human Factors in Computing Systems* (2018), CHI '18, pp. 86:1–86:12.

[326] D. Whittinghill, B. Ziegler, T. Case and B. Moore. Nasum Virtualis: A Simple Technique for Reducing Simulator Sickness. In *Games Developers Conf. (GDC)* (2015).

[327] R. T. Wigand, R. I. Benjamin and J. L. H. Birkland. Web 2.0 and Beyond: Implications for Electronic Commerce. In *Proc. of the 10th International Conf. on Electronic Commerce* (2008), ICEC '08, pp. 7:1–7:5.

[328] D. Wigdor and D. Wixon. *Brave NUI World: Designing Natural User Interfaces for Touch and Gesture*. Morgan Kaufmann Publishers Inc., 2011.

[329] B. Williams, G. Narasimham, T. P. McNamara, T. H. Carr, J. J. Rieser and B. Bodenheimer. Updating orientation in large virtual environments using scaled translational gain. In *Proc. of the 3rd Symp. on Applied Perception in Graphics and Visualization - APGV '06* (Boston, Massachusetts, 2006), p. 21.

[330] A. D. Wilson and M. Agrawala. Text Entry Using a Dual Joystick Game Controller. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems* (2006), CHI '06, pp. 475–478.

[331] C. A. Wingrave, B. Williamson, P. D. Varcholik, J. Rose, A. Miller, E. Charbonneau, J. Bott and J. J. LaViola. The Wiimote and Beyond: Spatially Convenient Devices for 3D User Interfaces. *IEEE Computer Graphics and Applications 30*, 2 (March 2010), pp. 71–85.

[332] J. O. Wobbrock. *Measures of Text Entry Performance*. San Francisco: Morgan Kaufmann, 2007.

[333] J. O. Wobbrock, D. H. Chau and B. A. Myers. An Alternative to Push, Press, and Tap-tap-tap: Gesturing on an Isometric Joystick for Mobile Phone Text Entry. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems* (2007), CHI '07, pp. 667–676.

[334] S. Yoo and C. Parker. Controller-less Interaction Methods for Google Cardboard. In *Proc. of the 3rd ACM Symp. on Spatial User Interaction* (2015), SUI '15, pp. 127–127.

[335] C. Yu, Y. Gu, Z. Yang, X. Yi, H. Luo and Y. Shi. Tap, Dwell or Gesture: Exploring Head-Based Text Entry Techniques for HMDs. In *Proc. of the 2017 CHI Conf. on Human Factors in Computing Systems* (2017), CHI '17, pp. 4479–4488.

[336] R. Yu, W. S. Lages, M. Nabiyouni, B. Ray, N. Kondur, V. Chandrashekar and D. A. Bowman. Bookshelf and Bird: Enabling real walking in large VR spaces. In *3D User Interfaces (3DUI), 2017 IEEE Symp. on* (2017), IEEE, pp. 116–119.

[337] P. Zaphiris, S. H. Kurniawan and R. D. Ellis. Age Related Differences and the Depth vs. Breadth Tradeoff in Hierarchical Online Information Systems. In *Proc. of the User Interfaces for All 7th International Conf. on Universal Access: Theoretical Perspectives, Practice, and Experience* (2003), ERCIM'02, pp. 23–42.

[338] J. C. Zeng and Z. H. Yu. Virtual Reality Technology and its Application. *Tsinghua University Press, Beijing, PRC* (1996).

[339] A. Zenner and A. Krüger. Shifty: A Weight-Shifting Dynamic Passive Haptic Proxy to Enhance Object Perception in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics 23*, 4 (2017), pp. 1285–1294.

[340] S. Zhai. Human Performance in Six Degree of Freedom Input Control, 1995.

[341] F. Zhang, S. Lin, X. Li, Y. Shuai, H. Jiang, C. Yao, F. Ying and F. Gao. Navigation Configuration and Placement Influences the Visual Search Efficiency and Preference. In *Proc. of the 26th International Conf. on World Wide Web Companion* (2017), WWW '17 Companion, pp. 871–872.

[342] L. Zhao and N. Zhang. The Virtual Reality System in Electronic Commerce. In *IEEE Symp. on Robotics and Applications* (June 2012), ISRA '12, pp. 833–835.